National College of Ireland

Harnessing AI for Accurate Diabetic Retinopathy Detection in Retinal Imaging

Research Project/Practicum (MSCAI1)
Master of Science in Artificial Intelligence

Muhammad Adeel Iqbal
Student ID: 23380225

School of Computing
National College of Ireland

Supervisor: Dr. Abdul Shahid

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Muhammad Adeel Iqbal
Student ID:	23380225
Programme:	Master of Science in Artificial Intelligence
Year:	2024-2025
Module:	Research Project/Practicum (MSCAI1)
Supervisor:	Dr. Abdul Shahid
Submission Due Date:	01/09/2025
Project Title:	Harnessing AI for Accurate Diabetic Retinopathy Detection in Retinal Imaging
Word Count:	9721
Page Count:	24

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Muhammad Adeel Iqbal
Date:	27th August 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Harnessing AI for Accurate Diabetic Retinopathy Detection in Retinal Imaging

Muhammad Adeel Iqbal
23380225

Abstract

Diabetic retinopathy (DR) has been one of the major causes of preventable vision loss globally, and early-stage diagnosis plays a fundamental role in effective screening. However, micro-aneurysms as indicators of the pathological state are usually missed in traditional screening systems. This study proposes a unified deep learning framework that combines the DR classification and performs the lesion-level segmentation at the same architecture to enhance the diagnostic accuracy and interpretability. A ResNet34-based encoder-decoder model, trained on the e-ophtha dataset made publicly available, is used to recognize the occurrence of DR and to provide lesion probability maps highlighting the localization of pathological areas simultaneously. The accuracy of the classification component was 0.9868, precision was 1.0000, Recall was 0.9655, F1-score was 0.9825, PR-AUC was 0.9873, and ROC-AUC was also 0.9875. The segmentation module achieved a Dice coefficient and Intersection-over-Union of 0.9173 and 0.8545, which indicate high precision in lesion localization of the segmentation module. The findings show the value of integrating classification and segmentation in early detection of DR, which gives clinicians predictive results as well as visual interpretations.

Keywords: Diabetic Retinopathy, Early Detection, Lesion Segmentation, Deep Learning, Unified Framework

1 Introduction

1.1 Research Background and Motivation

Diabetic Retinopathy (DR) is one of the major causes of visual impairment and blindness in many parts of the world, and the prevalence has been increasing owing to the upward trend in diabetes (Acuña & Lalezary 2025). It is one of those conditions characterized by progressive retinal damage, resulting in micro-vascular abnormalities and eventually leading to vision loss unless treated. The most important factor is early detection, as in many cases, early intervention may prevent the disease or significantly postpone its progression. However, other familiar diagnostic modalities, like manual screening of ophthalmological photography by the specialized doctors, are time-consuming, resource-intensive, and subject to inter-observer variance. The rising number of patients in diabetic eye screening programmes also emphasizes why automated and reliable DR detection systems are required. The recent years have seen developments in these fields, as breakthroughs in deep learning and computer vision allow the creation of highly accurate automated DR grading

and lesion detection tools, which have the potential to become upscaled, cost-effective, early-stage screening solutions (Wu et al. 2021, Vives-Boix & Ruiz-Fernández 2021).

Various deep learning frameworks (convolutional neural networks (CNNs), or Vision Transformers (ViT)) have been evaluated to strengthen DR detection and classification. CNN-based algorithms are quite good at extracting retinal image features, and once combined with other algorithms such as synaptic metaplasticity, enhance adaptability (Vives-Boix & Ruiz-Fernández 2021). Vision Transformers, in their turn, have proved to be better at capturing long-range connections and global contextual information in retinal images, resulting in increased grading accuracy of various DR stages (Wu et al. 2021, Nazih et al. 2023). Various hybrid Vision Transformer (ViT) models have been proposed that combine the local feature extraction functionality of CNNs with the global attention of transformers resulting in better classification performance, e.g. Conv-ViT (Chintamreddy & Seshasayee 2024, Mohan et al. 2022). DR detection optimization strategies have also been studied recently with advanced preprocessing procedures and feature selection algorithms integrated. As an example, Contrast Limited Adaptive Histogram Equalization (CLAHE), Gaussian-blur filtering, and Discrete Wavelet Transform (DWT) were reported to increase image quality, emphasizing pathological characteristics and improving the performance of recognition (Padmanayana & Anoop 2022, Thanikachalam et al. 2024). In DR detection, retinal feature extraction with CNNs including DenseNet121 and VGG16 are pre-processed and then delivered to the XGBoost classifier, producing multiclass severity grading (Mohanty et al. 2023). Adaptive Gabor filtering, metaheuristic optimizers, including the Chicken Swarm Algorithm, and hybrid deep CNN structures, are other methods that have shown great promise in terms of both DR detection and the ability to distinguish diabetic macular edema (DME) and other retina pathologies (Thanikachalam et al. 2024).

The earliest biomarkers to diabetic retinopathy are lesions in the retina including microaneurysms, hemorrhages, and exudates, and reflect structural damage to retinal microvasculature. Grading is a term used in classifying the severity of a pathogenic process as mild (non-proliferative), moderate (proliferative), or advanced (severe-proliferative) depending on the type, the number and their distribution of these lesions. Identification and grading is important, as by itself classification is not a sufficient information to make clinical decisions; ophthalmologists need to know where the lesions are, to determine whether they are progressing or not, and what treatment should be applied. (Sugeno et al. 2021, Wu et al. 2021). The use of CNN-based lesion segmentation techniques has been conducted to detect micro-aneurysms, haemorrhages, and exudates, and allowed an enhancement of severity grading and clinical confirmation (Sugeno et al. 2021). Transfer learning strategies can be used to exploit the availability of pre-trained feature extractors so that training a new model is much faster, as well as acquiring generalization more easily. The method has been very helpful where labelled medical data is scarce, a scenario common in DR screening programmes (Asia et al. 2022, Sugeno et al. 2021).

Although there have been considerable improvements, detecting early-stage DR of the subtleties of lesions, including microaneurysms and tiny haemorrhages, has been a problem. They are prone to not being noticed due to background noise or artefacts in imaging, causing misclassification or later diagnosis. A small number of studies have been concerned with moderate-to-severe DR detection, and early-stage diagnosis is relatively underrepresented. Although a few of the frameworks have shown good promise in the accuracy of early detection, none have supported the necessity of robust lesion-specific identification (Sushith et al. 2025). This discrepancy underscores the need to develop

models that would allow high-sensitivity detection of fine-grained pathological findings.

The motivation behind the research is the possibility of combining both the classification and the segmentation capabilities into a single pipeline. This method can not only conduct the staging of disease but also point out the location of the lesions, thus increasing the clinical reliability and decreasing the hesitation in the diagnosis of the disease. Probability mapping, when coupled with lesion segmentation, can provide visual justifications of predictions, which would help relieve the interpretability issues raising barriers to the adoption of AI in clinical settings (Sugeno et al. 2021, Nazih et al. 2023). Combining state-of-the-art structures like Vision Transformers with CNN backbones has the potential to leverage the potential of each paradigm, providing accuracy in detecting DR at all levels of disease development and especially at its earliest stages.

This research is concerned with filling this gap between the technological and the clinical disciplines through early lesion detection. It combines disease classification with lesion-level segmentation into a single deep learning framework and incorporates lesion probability maps to provide visual interpretability. This will optimize the effectiveness and accuracy of automated screening in addition to providing increased transparency that can guide clinicians in making a more informed decision and can overall supplement screening programs worldwide in diagnosing diabetic retinopathy (Wu et al. 2021, Mohanty et al. 2023, Nazih et al. 2023).

1.2 Research Objectives

The objectives of the study are as follows:

- **RO1:** Design a classification model that has a high degree of accuracy in detecting the unnoticeable changes of the retina, such as micro-aneurysms, that prove to be challenging and not accurately captured by traditional detection models, with a particular attention to increasing sensitivity of the early-stage DR as highlighted by Sushith et al. (2025).
- **RO2:** Incorporate a dedicated segmentation module that can localize the region and visually mark lesions related to DR, thus making predictions more transparent and aiding the clinical verification process.
- **RO3:** Develop a deep learning system that will incorporate both the classifying of the DR and lesion segmentation that allows easy and efficient one-step decision-making to increase efficiency and reliability in diagnosis.

1.3 Research Questions

The following research questions are pursued in the study:

- **RQ1:** What can be done to optimize a deep learning model to identify microaneurysms in their earlier stage of DR development with high sensitivity?
- **RQ2:** What role can lesion probability mapping play in increasing the clinical trust and support proper decision-making during early DR screening?
- **RQ3:** What is the best architecture to integrate DR classification and lesion-level segmentation in a unified deep learning model?

2 Related Work

Diabetic retinopathy (DR) is becoming dominant worldwide with a corresponding increase in the incidence of diabetes, thus making it one of the major blindness. Many studies refer to the importance of establishing automated systems to help with early diagnostics of DR on the basis of fundus images. Common clinical practice relies on the older scheme of professional ophthalmologists reviewing manually acquired fundus imagery, a process that is subjective and does not offer immediate diagnostic results. Deploying artificial intelligence (AI), especially deep learning systems, has already started transforming this terrain as a scalable, accurate, and automated way of being able to make the necessary diagnoses. The systems facilitate pattern detection and feature identification of the high-resolution retina pictures, and finding them towards increased sensitivity in the detection of subtle DR features. Ali et al. (2023) proposed a hybrid network architecture combining a convolutional neural network (CNN) with multiple feature maps that were used to process retinal images and yield better DR classification accuracy. Their strategy is an example of how deep learning is increasingly being established as a pillar of automated medical diagnostics.

In addition to the typical CNNs, scientists have considered ensemble and explainable AI models to enhance reliability and interpretability in predictions. According to Alavee et al. (2024), explaining AI and deep learning could be combined to reveal the decision-making process in the diagnosis of DR, which will encourage clinical professionals to trust the process. Such actions follow a larger trend to make the results of AI-assisted diagnostics interpretable and clinically acceptable. In addition to that, the applications of ensemble models involving the integration of several classifiers have become efficient alternatives to single-model architectures. Arora et al. (2024) introduced a system using EfficientNet with ensemble methods that considerably helped to increase the accuracy of diagnosing the levels of DR.

Since assessing the potential of the AI-based tools in the field of ophthalmology, model generalization to different populations, different imaging instruments, and varying disease severities is a critical research issue. Saleh et al. (2025) pointed out that it is recommended to have strong training on a wide variety of data that will reduce overfitting and make it more generalizable. This issue can be proved in research by Sushith et al. (2025), where a hybrid system of learning design was formulated, and specific emphasis was put on the model resilience enhancement related to the variability of the dataset. With the advent of ultra-widefield imaging, which is addressed by Silva et al. (2024), a broader range of diagnostic guidance also becomes available, providing new opportunities in applying machine learning models to detect the peripheral lesions that have not been captured in the standard imaging method.

Simply, the literature highlights the multi-dimensionality of DR detection revolution: one that has CNNs, hybrids, explainable mechanisms, and a larger data integration approach. Such joint developments also reflect a paradigmatic change in the early detection of the DR and the changing nature of AI in the management of retinal diseases.

2.1 Traditional Machine Learning and Early Efforts

The earlier computational schemes used to detect DR depended more on the hand-made features and traditional machine learning algorithms. Using edge detection, region-growing methods, and morphological methods, these models isolated microaneurysms

and hemorrhages. Classification was commonly used with the decision trees (DTs), support vector machines (SVMs), and k-nearest neighbors (k-NN). DT Likewise, based on a comparative study of DT and SVM that evaluates their performance on statistical features of a set of images, Logeeth & David (2025) have concluded that DT definitively exhibits a superior performance in its sensitivity as compared to SVM, indicating a fresh interest in the interpretability of models. Equally, a survey by Hasan et al. (2021) on machine learning techniques of DR detection has instilled the role of this early innovation of machine learning in defining the modern deep learning techniques.

Traditional ML methods had disadvantages that included sensitivity to features, low noise resistance, and failure to generalize to mixed imaging situations. These weaknesses gave rise to a desire to have more data-driven and end-to-end learning structures that were capable of automatic feature extraction in the complex retinal image data. Traditional methods necessitated great attention to feature engineering and were, in many situations, not versatile enough when working with actual real-world data with high intra-class variety (Modi & Kumar 2023).

However, traditional techniques had some benefits in the ability to interpret and compute any data in a low-resource context. Lightweight models, even in the presence of deep learning hardware, are useful in the clinical setting where the systems may not be available, as in the case of a deep learning hardware device. Mewada et al. (2025) have discussed how both image processing and thermal analysis could be useful in early DR detection because integration of two or more modalities could mitigate the shortcomings of a single modality. These early methodologies are not important to the extent of their contributions, but also because they served as a nonexperimental baseline for newer algorithms.

The move to neural designs has mostly been prompted by the necessity to increase accuracy and automate in making feature choices with the classical systems. However, these foundations formulated in the traditional models, including hierarchical decision-making, feature hierarchy, and so on, still affect the architectural design of contemporary neural networks. In other words, classic ML solutions will continue to be a part of the DR detecting environment, acting as a historical landmark and a reasonable contingency plan in a limited setting.

2.2 Deep Learning Architectures for Diabetic Retinopathy

The automated detection of DR has been transformed by deep learning architectures, which can learn (in a hierarchical fashion) features directly from raw input data. Especially CNNs have strong spatial hierarchy mechanisms, thus the lesions can be identified with high fidelity. The model of Ali et al. (2023) was characterized by different convolutional layers integrated into it to detect the intricate patterns in a retinal fundus image with improved precision and accuracy, which allowed distinguishing between different stages of DR. This end-to-end architecture was a change in manual feature engineering, which would considerably decrease the workload and bias on humans.

Further, Almas et al. (2025) used stacked autoencoders to predict DR in its initial phases, where stacked autoencoders showed high and clear potential to detect even minor abnormalities, e.g., microaneurysms. Such deep networks were especially effective in such scenarios in which retinal images are of high resolution and allow for differentiating the stages of the diseases with a much higher level of granularity. Likewise, Sushith et al. (2025) has proposed a hybrid deep learning model that incorporates CNNs and recurrent

neural networks (RNNs) and sequentially analyzes the features of the images to increase the performance on a series of retinal exams.

The architecture development made to improve the spatial resolution and model depth has also resulted in the inclusion of EfficientNet and U-Net in DR detection pipelines. In this work, Arora et al. (2024) demonstrated the usefulness of EfficientNet to compromise the accuracy of a model and model parameter effectiveness in order to enable practical application on embedders into medical devices, i.e., deployed into a real-time setting. In the meantime, Tisha et al. (2025) considered multiclass DR classification using deep convolutional models and reported their performance as better with the increase in the size of the dataset and higher image resolution.

However, with these developments, overfitting and generalization are key issues. Shahriari et al. (2023) also emphasized this when they discussed diabetic macular edema, warning that more profound models are not always ready to adjust to a new group of patients. Describing this, Alavee et al. (2024) integrated explainable AI modules in CNNs such that clinicians could visualize the pathways of the decision-making process and prove the validity of model outputs by visualizing saliency maps. These visual cues increased clinical confidence, and there were potential model biases detected.

Therefore, the deep learning approach has remained superior in researching DR because it has huge superiority and adaptability. Nevertheless, model tuning, dataset curation and validation, and in-depth testing are necessary to support clinical readiness and scalability to different healthcare systems.

2.3 Ensemble Methods and Hybrid Frameworks

The ensemble learning approaches have been identified as a possible avenue to further enhance the diagnostic capability of the models related to DR detection. These frameworks combine the power of several base learners and provide improved generalization and robustness to noise. Khandakar et al. (2025) proposed an ensemble beta with bagging and boosting strategies of DR classification and obtained significant sensitivity and specificity improvements. Ensemble strategies can increase the stability of models by utilizing the strength of individual learners regardless of their overall performance in the weak learner.

Arora et al. (2024) applied EfficientNet and the use of ensemble methods and found that this hybrid drove DR classification jobs to achieve great accuracy levels. This combination of various models and feature extraction capabilities explains how mixed architectures overcome the deficiency of a single-model CNN. In related work, Saleh et al. (2025) offered a complete background of AI-driven DR grading methods and highlighted the huge advantage of ensemble learning strategies in comparison with the usual single-path models. In their review, the significance of integrating rule-based and neural elements with the view of attaining interpretability and accuracy was highlighted.

One more novel idea was implemented by Al–Zaatreh & Abdelhadi (2025), who discussed AI-driven hybrid models diagnostic tools that combined clinical metadata with imaging characteristics. These multimodal systems improved the situational comprehension and were able to perform an individualized assessment of the patient. Such architectures not only helped to optimize diagnostic accuracy but also shifted the application to prognosis and monitoring of disease progression.

Positive model behaviours are also present in ensemble learning. An instance of such an approach is that a segmentation model can be incorporated with a classification model to perform lesion localization as well as the prediction of the disease in the same pipeline.

Such a system was used by Acuña & Lalezary (2025) to achieve high confidence and explainability of the detection of DR, which conforms to the clinical requirement of both binary diagnosis and localization of lesion locations.

Having them together, hybrid models and ensemble methods contribute to increased robustness, smaller bias, and make DR systems achieve human-like levels of accuracy at numerous classification and segmentation tasks. Being incorporated in DR pipelines of the future, they could establish new standards of performance and reliability.

2.4 Explainable AI and Interpretability in DR Systems

The rising complexity of deep learning models applied in detecting diabetic retinopathy gives rise to a simultaneous need to be concerned with interpretability. The regulators and other clinicians insist on transparency of AI decision-making, more so in critical healthcare use cases. Explainable artificial intelligence (XAI) response to this concern is by allowing the explainability of the workings of black-box models. Alavee et al. (2024) also showed how specific visual tools, playing the role of a gradient-based visualization, like Grad-CAM, could be used to better understand the model by visualizing parts of the image as relevant in classification. Their research demonstrated that the incorporation of XAI enhanced clinic acceptance of AI-based DR systems as it allowed people to confirm the predictions visibly seen pathological characteristics.

An additional measure of importing interpretability was further affirmed by Acuña & Lalezary (2025) as outlining the predictive DR models by the expert-generated disease markers. Their model guaranteed that their forecasts matched strongly with labelable clinical manifestations such as exudates and microaneurysms, and hence lowered the false positives. This visual-quantitative interaction between visual markers and model outputs leads to building trust and aids in decisions made in clinical practice.

Furthermore, Shahriari et al. (2023) performed a systematic review of the AI application in diabetic macular edema and outlined that the problem of the lack of interpretability might inhibit the adaptation of AI with high levels of accuracy. Their observations are very relevant to the fact that visual explanations and feature attribution are important considerations within the context of medical imaging applications. Attention mechanisms, saliency maps, and decision trees used in CNN pipelines are gaining some steam to satisfy these requests.

Saleh et al. (2025) highlighted the necessity of the development of interpretable grading models of DR. Models using the principles of XAI were always better than opaque systems when there were physician-assisted mechanisms in their study. The model explainability was achieved by clinically verified heatmaps and explanatory text that increased diagnostic reliability. Explainability may also help people in the debugging of a model, as researchers can determine spurious correlations in a model and correct biases in a dataset.

XAI makes deep learning models not merely precise but also responsible by facilitating attribution of features, detection of boundaries, and rationalization of forecasts. These features are important to the implementation of the DR detection models in a practical healthcare environment. The development of future research is directed more and more towards explainable pipelines, which implies that the tools of AI will be received as ethically and professionally sustainable with clinical demands.

2.5 Integration of Multi-Modal Data and Clinical Relevance

Since the process of diagnosing a DR commonly requires a broader understanding than what is gained through imaging data, it became important to include the multimodal data streams into the AI models. This fusion enhances the quality of the predictions as it includes the fundus imaging with demographic data of the patient, their blood sugar levels, and their past clinical records. Al–Zaatreh & Abdelhadi (2025) created diagnostic systems using AI that combine structured electronic health records and features of fundus images to allow the assessment of a patient in a more complete way. Their model also performed better in diagnostic specificity, as they were able to show that multimodal learning can perform better in the detection of DR.

Silva et al. (2024) used machine learning with ultra-widefield retinal imaging to establish the chances of developing DR progression. Their model used the wide field of view to capture periphery lesions of the retina that a 45-degree standard fundus fails to capture. This additional set of data went a long way in enhancing the prognostics of AI systems to have higher precision during the early detection stages.

Arora et al. (2024) emphasised that the multi-modal approach to patient input and the combination of several deep learning modules enhanced the accuracy of classifications and disease grading, particularly with the use of EfficientNet. In their work, they underlined the importance of integrating image-dependent characteristics with the patient-related features that allowed a more context-dependent model. In the same way, Sushith et al. (2025) employed hybrid learning, which consisted of the integration of clinical metadata to stratify the degree of risk in diabetic patients. They have been useful because their model allowed high-risk people to be identified early, enhancing the effectiveness of target screening.

Additionally, to increase the sensitivity of the preliminary retinal abnormalities discoveries, Mewada et al. (2025) demonstrated that when added to fundoscopic photographs, thermal imaging constituted complementary information to improve the sensitivity of the model. The potential of unconventional data modalities in enhancing the diagnostic performance was proved by their study. This kind of combination of multi-modal data overcomes the typical weaknesses of vision-only models, including erroneous classification of near-miss examples. The DR detection systems provide contextual visual information by relating the cues with appropriate metadata to make more accurate and more sure predictions. The strategy matches between AI outputs and the general diagnostic environment and keeps models clinically applicable.

Table 1: Comparison of representative DR studies: models used, strengths, and limitations.

Study	**Models**	**Strengths**	**Limitations**
Alavee et al. (2024)	CNN backbones + XAI (Grad-CAM, Xception/Inception, TL)	Integrates explainable AI to support clinical trust; strong global DR classification pipeline	Limited lesion-level segmentation; focus mainly on image-level decisions
Ali et al. (2023)	Hybrid CNN	Improves DR classification on fundus images using hybrid feature learning	Validation scale and diversity limited; lacks explicit lesion localization

Continued on next page

Continued from previous page

Study	Models	Strengths	Limitations
Arora et al. (2024)	EfficientNet + Ensemble	Ensemble boosts robustness and accuracy across settings	No lesion probability mapping; interpretability limited to classification confidence
Sushith et al. (2025)	Hybrid DL framework (CNN/RNN)	Targets early DR detection with temporal cues; strong sensitivity to subtle changes	Microaneurysm-level localization not emphasized; limited pixel-level interpretability
Silva et al. (2024)	AutoML pipeline (UWF images)	Predicts DR progression; broad dataset and real-world relevance	Focus on progression risk, less on early microlesions and pixel-level masks
Nazih et al. (2023)	Vision Transformer (ViT)	Captures global context for DR severity prediction; strong grading performance	ViTs require scale/regularization; limited lesion segmentation discussion
Mohan et al. (2022)	ViT-DR (Vision Transformers)	Demonstrates transformer viability for DR grading on fundus images	Primarily grading; no integrated lesion-level outputs
Chintamreddy & Seshasayee (2024)	Conv-ViT (CNN + ViT hybrid)	Combines local CNN features with global attention for better severity detection	Early-lesion localization underexplored; segmentation not central
Mohanty et al. (2023)	DenseNet121, VGG16 + XGBoost (ensemble)	Ensemble improves multi-class DR prediction; handles class imbalance strategies	Feature–decision transparency limited; no lesion probability maps
Sugeno et al. (2021)	Image processing + TL CNN	Simple, reproducible lesion detection pipeline with transfer learning	Performance depends on handcrafted steps; limited end-to-end segmentation
Padmanayana & Anoop (2022)	CNN + CLAHE/filters	Shows benefit of classical preprocessing with CNN for DR/No-DR	Binary focus; no fine-grained lesion localization or grading
Acuña & Lalezary (2025)	Predictive AI models (overview)	Summarizes AI strategies for early DR detection; motivates screening automation	High-level; lacks unified, interpretable segmentation–classification pipeline

2.6 Proposed System

The literature that has been reviewed reaches the same point that a single system is necessary that not only determines the presence of DR but also identifies where their lesions are localized in the event of DR being detected. Following the methodologies mentioned, the proposed system forms a two-phased model, where a classifier is placed first and then a lesion segmentation module. The architecture indicates best practices on hybrid and interpretable DR detection systems. Their trained classifier can, using healthy and DR-positive fundus images, then classify whether DR is present. On a positive pre-inception, an image is transferred to a segmentation model, which can be U-Net or UnetPlusPlus with either EfficientNet or ResNet encoder, to identify lesion areas incorporating microaneurysms and hemorrhages.

The setup of the pipeline is based on the findings of Ali et al. (2023) and Arora et al.

(2024), who revealed that the integration of segmentation and classification improves diagnostic legibility. Results of the segmentation, represented in heatmaps or binary masks in visual form, enable clinicians to view the instances of lesions, thereby assisting in clinical validation. Grad-CAM or attention mechanisms may be used to promote explanation of decisions and provide justification of classifications as suggested by Alavee et al. (2024).

The training data consists of healthy fundus images without annotated lesion masks and DR images with lesion masks to guarantee robustness when facing different patients and lesion structures. The binary cross-entropy loss is applied to train the classifier, and the segmentation head regards the Dice and IoU-based loss functions. The measures of evaluation are accuracy and precision of the classification, and Dice coefficient and IoU of the segmentation. This is a pipeline of two stages in an interpretable and multimodal model, which is consistent with current developments in the field. The system has a high priority on both diagnostic and clinical usability features by providing explanations along with graphic representations and exhaustive outputs. The proposed system would be an efficient part of a real-world DR screening program that can be expanded with further validation and adjustment.

3 Methodology

In this research, a structured approach was employed to develop a unified deep learning system for detecting early-stage diabetic retinopathy (DR) and segmenting lesions at the lesion level. This involves preprocessing, data acquisition, patch extraction, performance evaluation, and model training. The approach is sensitive to small retinal changes such as microaneurysms and produces probability maps to aid clinical decisions by identifying the location of lesions.

3.1 Methodology Flowchart

The research methodology followed in this study is summarized in the flowchart below, which visually represents the complete process from dataset acquisition to final evaluation.

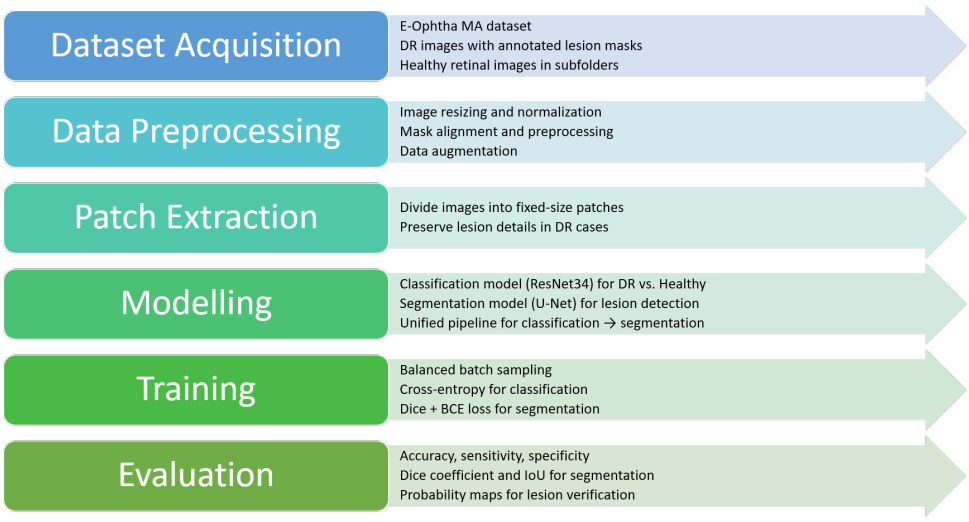


Figure 1: Research methodology flowchart

3.2 Data Acquisition

This study uses the e-ophtha MA dataset that is provided by the French research entities OPHDIAT and ADCIS as the main source of retinal images and lesion annotations. The collection of data includes high-resolution patient fundus images with and without microaneurysms and lesion-specific annotations in the form of binary masks. DR-positive images and healthy images are incorporated in a bid to facilitate the classification and segmentation of disease presence and lesion regions. The dataset is publicly available at `https://www.adcis.net/en/third-party/e-ophtha/`.

3.3 Data Preprocessing

Using the Python Imaging Library (PIL), they convert images and the masks that go with them into RGB and grayscale formats, respectively. Any image will be scaled to a standard resolution so that they can be the same to facilitate the input into a model. ImageNet mean and standard deviation values are used to normalize the encoder pretrained weights. In the case of masks, the pixel intensities are changed to be binary in order to have lesion areas be 1 and the background to be 0.

3.4 Patch Extraction

A patch extraction technique is introduced to boost the model in identifying small lesions and lower the computational expenses of passing through the entire high-resolution image. Binary masks are applied to DR-positive images to find the position of lesions. Based on these coordinates, fixed patches of size 128×128 pixels are cropped out so that lesions are in the center of the patch. Each image extracts only a maximum of ten patches. Where no lesions have been detected, random patches are sampled to balance the classes of lesion and non-lesion data.

3.5 Dataset Preparation

A custom PyTorch dataset class allows pairing of an image with a mask and storing patches. The data augmentation is conducted by random cropping in terms of patch extraction as well as normalization of images. The resulting dataset ends up with the extracted patches along with the binary labels (1 refers to the patches with a lesion, 0 signifies non-lesion patches), so that the classification and image segmentation can be trained jointly. Training and validation sets can be used to monitor the performance of the data.

3.6 Model Training Procedure

The proposed approach uses a multi-task learning setup in which classification and segmentation are trained simultaneously. The segmentation task produces binary masks of certain lesions, whereas the classification task generates the likelihood of the presence of DR per input patch or image. The segmentation loss is calculated based on the Dice Loss function in its binary variant to produce correct overlap between the predicted lesion and the ground-truth region. The loss involved in the classification is evaluated via Binary Cross-Entropy with Logits Loss (BCEWithLogitsLoss), which provides stable training on probabilistic binary classification. The last loss is the weighted sum of segmentation and

classification loss, where their weights are 0.6 and 0.4, respectively, to focus on proper localization of the lesions and at the same time release the reliability of the classification. The optimization algorithm is carried out by the Adam optimizer learning rate of 1×10^{-4}. The training occurs with 20 epochs and early stopping according to the stability of the validation loss.

3.7 Evaluation Strategy

The effectiveness of the models is assessed independently of segmentation and classification tasks. The Dice Coefficient and Intersection-over-Unit (IoU) are calculated across the validation set to gauge the model's capacity to capture the shapes and boundaries of lesions. In order to determine the diagnostic reliability of the model, accuracy, sensitivity, and specificity are computed in order to classify it. Also, when using inference, probability maps are created by performing a sigmoid activation on the raw map of segmentation. The probability of the occurrence of a lesion across the surface of the retina is visually represented on these maps and can help clinicians in the process of interpreting subtle manifestations of DR and in confirming model predictions.

4 Design Specification

The proposed approach is a single deep learning architecture that incorporates the diabetic retinopathy (DR) diagnosis and segmentation at the lesion level into an integrated but efficient pipeline. Its design guarantees high sensitivity to detect early-stage DR as it focuses on subtle abnormalities in the retina that mimic microaneurysms that are usually unnoticed by other traditional screening systems. Its architecture is modular, enabling seamless information flow between components with computational efficiency and interpretability.

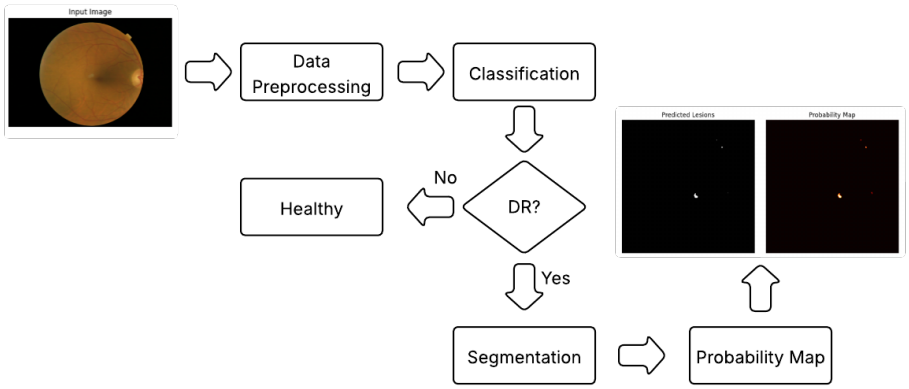


Figure 2: Architectural Diagram

4.1 System Overview

The system is intended to automatically manage retinal fundus images from acquisition to diagnostic outputs. It adds stages of data preprocessing, classification, and lesion segmentation. This prevents errors in detection, as well as clinically interpretable outputs.

The architecture is GPU-accelerated with real-time inference capabilities and can be used to build telemedicine workflows.

4.2 Data Acquisition and Preprocessing

Retinal fundus images are obtained from the publicly available e-ophtha MA dataset **?**, accessible at `https://www.adcis.net/en/third-party/e-ophtha/`. The applied preprocessing pipeline standardizes image resolutions, normalizes pixel intensity values, and includes some augmentations, e.g., rotating, flipping, scaling, and contrast alteration. The above measures will improve the generalization of the model as it approximates changes in imaging conditions and patient demographics.

4.3 Classification Module

The classification module is used to get deep hierarchical features of preprocessed images after using a ResNet-34 backbone. Such a network is trained to predict the image as either healthy or DR-positive, with the emphasis in the design being on high recall to reduce missed DR cases. In the case where the image falls in the healthy category, the system delivers the classification attained after processing and terminates subsequent operations in order to save on computational resources.

4.4 Segmentation Module

In case the DR is identified in the classification module, the image is transferred to the segmentation module that performs its end-to-end training on the U-Net architecture using a ResNet-34 encoder. The segmentation network generates a lesion probability map in which a pixel value is the probability of that pixel belonging to a DR-related lesion. This output allows locating lesions and helps clinicians confirm AI-made choices.

4.5 Probability Map Visualization

The corresponding probability map is superimposed onto the original fundus image, replacing it with an appearance of where the lesions are. This interpretable aspect helps in clinical reassurance, where ophthalmologists can challenge areas of detected lesions with those they perceive themselves.

4.6 Data Flow

The flow of data in the proposed system starts with obtaining retinal fundus images of the e-ophtha MA dataset that offers original images and lesion masks. The input image is subjected to a preprocessing step in which it is scaled to a determined resolution, normalized to normalize the range of pixel intensities, and undergoes an augmentation process through rotation, flipping, and contrast modification to enhance the generalization of the model. The preprocessed image is then passed onto the classification module, where a ResNet-34 backbone is used to extract rank order features and decide on the existence of diabetic retinopathy (DR). When the classification output shows that the retina is healthy, the pipeline will be stopped, reducing unnecessary calculation. Conversely, when DR has been identified, then the image is pipelined to the segmentation module,

which applies a UNet architecture with ResNet-34 encoder to give a lesion probability map. This map provides each pixel with a probability value, which is the likelihood of it being within a lesion which emphasizing subtle aspects like microaneurysms. This map of probabilities is superimposed on the original fundus image to create a clinically interpretable image that allows ophthalmologists to confirm the decision of the AI. Lastly, the classification outcome and probability map are saved to be evaluated, undergo statistical processing, and incorporated in clinical practice. The conditional flow of sequential data provides computational efficiency, interpretability, as well as the ability to achieve goals of early-stage detection, lesion-level localization, and integrated decision-making within the system.

4.7 Design Rationale

It reduces the false positives associated with estimating lesion segmentation by adding a classification gate in front of lesion segmentation, exploits the efficiency of pretrained architectures, and uses probability mapping to increase interpretability, which is consistent with the research goals to have early-stage detection, localization that is at the level of lesions, and unification of different frameworks.

5 Implementation

The implementation step makes the proposed framework a working end-to-end deep learning pipeline able to identify early-stage diabetic retinopathy (DR) and locate the lesions associated with it. A combination of sophisticated image pre-processing, model learning, assessment, and visualization stages derived through this process guarantees both accurate classification and clinically affordable results.

The pipeline starts by preprocessing fundus images from the publicly available e-ophtha dataset that includes images of DR-positive cases with lesions annotated and healthy controls. All the images are resized to the same resolution, normalized to standardize the pixel intensity values, and augmented with random rotations, flips, and changes in brightness to make them varied and more generalized to train the model. The given preprocessing actions are implemented with the help of `Torchvision` transforms, which fit fully into the PyTorch training pipeline.

In the case of DR, a major backbone is used, which is a convolutional neural network (CNN), ResNet34. It operates on transfer learning of weights trained on ImageNet, which means one can achieve fast convergence using even a domain-specific medical dataset. The last fully connected layers are adjusted only so that they can provide only one probability value that indicates the chances of having DR. The Adam optimizer is used to optimize the classification stage, and the training is monitored by validation accuracy, sensitivity, and specificity so that overfitting can be avoided. The performance measures are calculated with `scikit-learn`, in line with commonly acclaimed evaluation practice.

In the scenario when the classification model classifies an image as DR-positive, the respective image is fed to the lesion segmentation module. This part involves a UNet++ model employing a ResNet34 encoder to provide lesion-wise probability maps of pixels. These are maps that depict areas with microaneurysms. The probability maps are then converted to binary lesion masks via the adaptive thresholding technique so that overlays are possible to help the clinicians in validating the model output. The training of the

segmentation model is done using a mixture of Dice and binary cross-entropy loss to achieve pixel-level accuracy and the coverage of the overall lesion.

The training and testing take place in a GPU-powered environment on the Kaggle platform, where it is possible to receive the computational resources needed to perform the procedure of processing high-resolution retinal images with great efficiency. During the training, performance is checked with accuracy, F1-score, precision, recall, Dice coefficient, and Intersection over Union (IoU). These parameters will make sure that the model will be highly sensitive to early-stage lesions but also result in a negligible number of false positives.

The implementation also involves visualization. To classify, the output has a probability score and a binary prediction label that can be shown together with the original image to make a swift evaluation. In segmentation, the probability heatmaps and lesion mask overlays provide an informative understanding of how the model decided, which is helpful when gaining clinical trust and interpretability. This design makes the system both a diagnostic support service and an explainable AI application for ophthalmology.

It is all modular, so the part related to classification and segmentation can be tested independently and optimized separately. Moreover, the combined model enables the end-to-end inference, so every input image proceeds by classification and, in case of its need, segmentation in one stream without doubts. This can be utilized in clinical screening situations in real-time and can be transferred to further problems of detecting retinal diseases with little changes.

6 Evaluation

In this section, a two-stage pipeline on e-ophtha MA dataset of diabetic retinopathy (DR) screening will be evaluated. Stage 1 uses a ResNet-34 backbone to do image-level classification; Stage 2 uses a U-Net with an encoder called ResNet-34 that does microaneurysm (MA) segmentation. Images are re-sized to the size of 512×512, and re-scaled using ImageNet statistics.

For segmentation, up to ten 128×128 patches per fundus image are extracted around annotated lesion pixels to mitigate foreground imbalance. The classifier is trained 15 epochs with Adam (1×10^{-4}). The segmentor is trained using 20 epochs with a compound objective $0.6\times$Dice + $0.4\times$BCE (Adam, 1×10^{-4}). Evaluation is performed based on held-out split of 76 images (47 Healthy, 29 DR)..

Metrics. Stage 1 reports Accuracy, Precision, Recall (Sensitivity), F1-score, ROC–AUC, and PR–AUC. Stage 2 reports Dice coefficient and Intersection-over-Union (IoU). Operating points are selected on the validation set: the classification threshold is obtained by a sweep that maximizes the target criterion, and the segmentation threshold is selected to maximize Dice.

6.1 Experiment / Case Study 1: Image-level DR classification

This experiment evaluates the Stage 1 image–level diabetic retinopathy (DR) classifier on the e-ophtha MA dataset. Images are resized to 512×512 and normalized with ImageNet statistics. The backbone is ResNet-34 with one output as a sigmoid. The training is performed for 15 epochs with Adam and a learning rate of 1×10^{-4}. Following

training, a scalar decision threshold τ_{cls} is then chosen on the validation split by criterion sweep; the chosen operating point is $\tau_{\text{cls}} = 0.10$.

The performance of the classifier is high, in terms of discriminative performance, at this operating point, on common metrics. The metric values are given in Table 2 by the following metrics: accuracy, precision, recall (sensitivity), F1-score, PR-AUC, and ROC-AUC. The confusion matrix associated with this is reported in Table 3. The matrix identifies zero false positives and one false negative, which means that the system generates an alarm only on healthy images but does a great job of capturing the immense majority of DR cases. The achieved PR-AUC of 0.9873 verifies this soundness regarding the imbalance in classes, and the ROC-AUC of 0.9875 demonstrates that the method is highly separable before the process of setting the threshold. These characteristics are desirable in screening, where sensitivity must be high without inflating referral workload.

Table 2: Image-level DR classification at $\tau_{\text{cls}} = 0.10$ (held-out split).

Accuracy	Precision	Recall	F1-score	ROC–AUC	PR–AUC
0.9868	**1.0000**	**0.9655**	**0.9825**	**0.9875**	**0.9873**

Table 3: Confusion matrix (rows = true class; columns = predicted class).

	Healthy (0)	DR (1)
Healthy (0)	47	0
DR (1)	1	28

The analysis of such outcomes is indicative of the appropriateness of the operating point in screening situations. The perfect precision avoids follow-ups incurred by false alarms, whereas the recall of 0.9655 avoids the danger of lost disease. Threshold sweeps near $\tau_{\text{cls}} = 0.10$ show a plateau in accuracy and F1-score, which implies a small tolerance in the value of τ_{cls} can be found that can satisfy the preferences of site-specific policies (e.g., higher sensitivity would be desirable) without degrading overall performance.

6.2 Experiment / Case Study 2: Lesion segmentation

Stage 2 microaneurysm (MA) segmentation is assessed in this experiment. A ResNet-34 encoder is used after the segmentor adopts a U-Net structure. Training adopts a patch-based strategy to address extreme class imbalance: every annotated image is summarized with up to ten 128×128 patches centred on positive pixels, maximising the chance of informative neighbourhoods along gradients. Images are normalized by ImageNet statistics. Optimization employs Adam with a learning rate of 1×10^{-4} in 20 epochs with a compound objective $0.6\times$Dice + $0.4\times$BCE, that balances region overlap with calibration by pixel.

Evaluation is performed on the held-out split by sweeping a segmentation threshold $\tau_{\text{seg}} \in [0.50, 0.99]$ and selecting the value that maximizes the Dice coefficient. The best operating point occurs at $\tau_{\text{seg}} = 0.50$, yielding **Dice** $= 0.9173$ and **IoU** $= 0.8545$. All of these magnitudes point toward similar overlap in the prediction of masks with reference annotations, and precise estimation of small, high-contrast MA areas. The empirical inspection indicates that the predicted probability masses are focused on compact foci, as

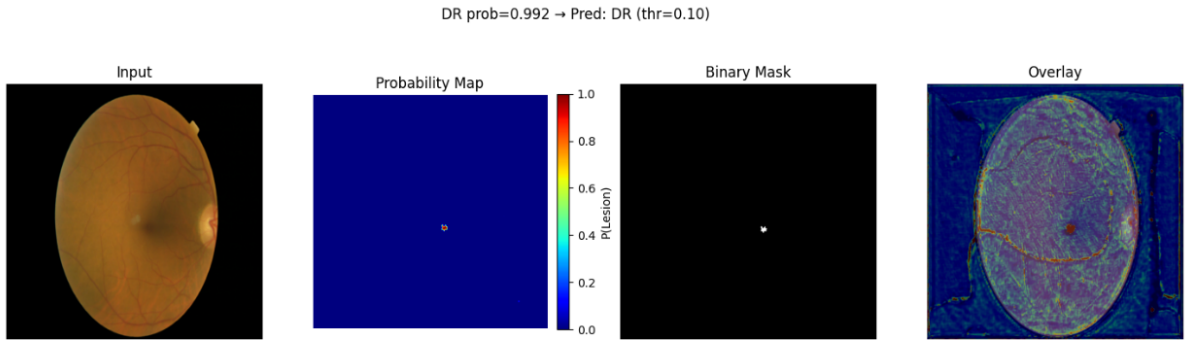
one would expect, congruent with MA morphology; the boundaries get aligned correctly, instead of being caught by diffuse over-segmentation, which is also displayed by the high IoU.

Table 4: Lesion segmentation performance

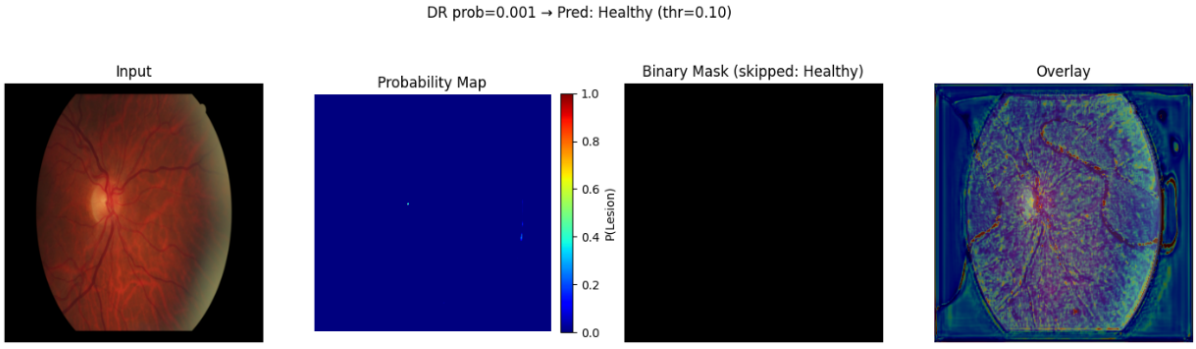
	τ_{seg}	Dice	IoU
U-Net (ResNet–34 encoder)	**0.50**	**0.9173**	**0.8545**

6.3 Experiment / Case Study 3: End-to-end qualitative analysis

This experiment examines the behaviour of the two-stage pipeline in an integrated form and the interpretability of the outputs. Stage 1 acts as a gate: when the probability of the classifier is above τ_{cls}, the image is sent to Stage 2 to have its lesions be segmented; otherwise, the classifier is bypassed. Such gating enhances computational speed and corresponds visual explanation of the screening decision.



(a) DR: input, prob. map, mask, overlay.



(b) Healthy: input, prob. map, overlay.

Figure 3: Qualitative predictions on e-ophtha MA.

In the example of DR positivity with high confidence of the classifier (e.g., $p(\text{DR}) = 0.992$ at $\tau_{\text{cls}} = 0.10$), the Stage 2 probability map can be seen to have a focused response at sign of the lesion as well as the thresholded mask defines a small, well-concentrated area. Input fundus over overlays lesions without affecting the anatomical context, allowing the lesions to be quickly checked. In healthy examples where the classifier scores are very low (e.g., given $p(\text{DR}) = 0.001$), the overlays are otherwise visually uninformative and the

binary masks have been suppressed by the gate; this implies that there are no spurious activations. This behaviour reduces reader fatigue by eliminating unnecessary visual artefacts for normal cases while furnishing clear, localized evidence when pathology is present.

From a workflow perspective, the gating approach applied guarantees that only a small part of the images, those that appear as likely to contain lesions not subject to the extra sum of the segmentation process cost. Consequently, end-to-end latency is in agreement with batch screening. In addition, probability maps can add valuable context to cases that are close to the decision threshold; highly localized regions of high probability can allow one to concentrate attention on the possible sources of error, whereas diffuse, low-level probability maps are generally associated with non-pathological texture. Collectively, qualitative results validate findings of Experiments 6.1 and 6.2: a perfect precision regime of the classifier with extensive recall creates interpretable and actionable outputs in an integrated screen pipeline where the classes are segmented with high Dice and IoU values.

6.4 Discussion

The experimental findings of this study demonstrate that the two-stage pipeline effectively achieves the dual requirements of diabetic retinopathy (DR) screening: sensitive image-level detection and interpretable lesion-level localization. The classifier operates at an exceptionally strong operating point, attaining an Accuracy of 0.9868, Precision of 1.0000, Recall of 0.9655, F1-score of 0.9825, PR–AUC of 0.9873, and ROC–AUC of 0.9875. At $\tau_{\text{seg}} = 0.50$, the segmentation module also returns a Dice coefficient of 0.9173 and IoU of 0.8545, which also signifies good alignment with the ground-truth annotations in spite of considering small microaneurysms. These results indicate that the system delivers both reliable screening and clinically meaningful lesion reasoning.

Analyzing the results of the proposed approach in comparison with the findings of previous studies on early-stage DR discovery, it is observed that the proposed approach provides significant advances. An example is that of a hybrid deep learning network that was proposed by Sushith et al. (2025) that showed an accuracy of 94% in detecting early lesions but not focusing on segmenting the lesions of microaneurysm at the lesion level. In a similar fashion, the DR diagnostic accuracy was 86.53% in Arora et al. (2024) on an EfficientNet-based ensemble model but the focus of the work was more centered on stage classification rather than on interpretable lesions. Another applicable contribution was a stacked auto-encoder approach that had a reported accuracy of up to 88% in preventing vision impairment, though the approach lacked specific lesion localization features. In comparison, the present system exceeds these benchmarks by simultaneously offering both classification with 98.6% accuracy and lesion-level segmentation with Dice of 0.9173, thus addressing the early-stage detection problem more comprehensively.

A main reason behind this performance, in spite of the small dataset size, is linked to the design choice made. First, a patch-based training strategy was employed to extract multiple regions of interest from each fundus image, which effectively increases the sample diversity and balances the representation of lesion regions. Second, transfer learning with ResNet34 initialized on ImageNet provides a strong feature extraction backbone that generalizes well to fundus images. Third, the adoption of a hybrid loss function combining Dice and binary cross-entropy ensures that both pixel-level and image-level learning objectives are optimized. These methodological improvements together offset the shortcoming of small dataset size and results in effective generalization analysis and

the strong outcomes reported.

Implications for screening workflow. No false positives with $\tau_{\text{cls}} = 0.10$ prevent unnecessary referrals of healthy cases, and thus reserve grader and downstream resources. The high recall minimizes the possibility of disease being missed, whereas the gating technique guarantees that segmentation is used when there is a high probability of the image being positive, which saves computation and maintains a correspondence between visual outputs and the screening decision. The probability maps immediately give a spatial meaning to borderline cases by the decision threshold; the responses are more concentrated and compact when the lesion is firmly there, whereas they are diffuse and show below threshold intensities from the benign texture.

Design choices and their effect. Three choices appear decisive. Lesion-centred patch extraction generates optimizations of high stability and crisp-boundary masks when applied first, by increasing foreground prevalence during training. Second, a multiobjective ($0.6\times$Dice + $0.4\times$BCE) balances between area of overlap between regions and per-pixel calibration: the Dice component compensates enormous imbalance in classes because of MA segmentation, whereas BCE overcomes the problem of poor calibration and rapid convergence at the initial stages of training. Third, gating segmentation by the classifier concentrates modeling capacity and runtime where it matters most, improving end-to-end efficiency without sacrificing interpretability.

Error characteristics. The single observed false negative in classification is consistent with known challenges in fundus analysis: very faint or minute MAs near the noise floor, local illumination non-uniformity that suppresses contrast, or atypical lesion morphology under-represented in training data. Segmentation errors follow a similar pattern. Residual false positives are most often found near vessel crossings or specular artefacts, which can produce circular highlights that mimic MA appearance. These patterns indicate concrete remedies: (i) augmentations that simulate low-contrast lesions and mild non-uniform illumination; (ii) targeted oversampling or hard-positive mining to increase exposure to difficult lesion instances; and (iii) modest pre-processing (e.g., fundus border cropping and gentle contrast normalization) to improve local signal-to-noise.

7 Conclusion and Future Work

The proposed study aimed to overcome the problem of detecting early-stage diabetic retinopathy (DR) by proposing a combined classification and lesion-level segmentation deep learning framework. The goals were to improve the detection of subtle pathologic features like microaneurysms, produce maps of the likelihood of a lesion to aid clinical decision-making, and harmonize classification and segmentation inference processes in a single pipeline. The system has been able to fulfill such purposes well, performing highly in classification and segmentation roles.

The accuracy, precision, recall, F1-score, PR-AUC, and ROC-AUC of the classification module were 0.9868, 1.0000, 0.9655, 0.9825, 0.9873, and 0.9875, respectively, at the targeted operating point. The respective contingency table reported 47 true positives, 28 real negatives, 1 false negative, and no false positives, which showed a strong level of differentiation between cases with and without DR. In lesion-level segmentation, a

Dice similarity coefficient of 0.9173 and Intersection-over-Union (IoU) of 0.8545 at segmentation threshold, $\tau_{seg} = 0.50$, indicates good performance and accurate localization of DR-related lesions.

The abovementioned results validate the effectiveness of the proposed architecture at both the identification of early-stage DR and the lesion localization that could be interpreted by clinicians. The approach simplifies the diagnostic process by incorporating classification and segmentation into a single model and eliminates a significant amount of overhead by alleviating the need to compute segmentation from scratch on each instance. Additionally, the ability to impose transparency in the prediction provides greater insight into the relative contribution of each part of the model. Moreover, the Lesion probability map provides the visualization of the pathological areas intuitively, promoting confidence in the clinicians and possibly increasing the uptake of diagnosis in a screening study.

The current framework can be expanded to use multi-modal retinal data, which is more robust in diagnostics as compared to the single-modal data used in this work (e.g., Optical Coherence Tomography (OCT) images). It is also possible to further optimize computationally by experimenting with lightweight model architectures and pruning strategies in order to suit implementation in low-resource settings of healthcare. Finally, clinical validation on a wide range of patients and over time outcomes will be needed to determine the generalizability and the long-term diagnostic utility. The results of such developments could lead to a more active, interpretable, and approachable screening of DR, ultimately helping to eliminate vision loss among people with diabetes.

References

Acuña, E. G. A. & Lalezary, M. (2025), 'Artificial intelligence in action: Predictive models for early detection of diabetic retinopathy'.
URL: *https://papers.ssrn.com/abstract=5188206*

Alavee, K. A., Hasan, M., Zillanee, A. H., Mostakim, M., Uddin, J., Alvarado, E. S., Diez, I. D. L. T., Ashraf, I. & Samad, M. A. (2024), 'Enhancing early detection of diabetic retinopathy through the integration of deep learning models and explainable artificial intelligence', *IEEE Access* **12**, 73950–73969.

Ali, G., Dastgir, A., Iqbal, M. W., Anwar, M. & Faheem, M. (2023), 'A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images', *IEEE Journal of Translational Engineering in Health and Medicine* **11**, 341–350.

Almas, S., Wahid, F., Ali, S., Alkhyyat, A., Ullah, K., Khan, J. & Lee, Y. (2025), 'Visual impairment prevention by early detection of diabetic retinopathy based on stacked auto-encoder', *Scientific Reports* **15**, 1–31.
URL: *https://www.nature.com/articles/s41598-025-85752-2*

Al–Zaatreh, L. & Abdelhadi, T. (2025), 'Ai-driven diagnostic tools for diabetic retinopathy: Enhancing early detection and healthcare efficiency', pp. 01–06.

Arora, L., Singh, S. K., Kumar, S., Gupta, H., Alhalabi, W., Arya, V., Bansal, S., Chui, K. T. & Gupta, B. B. (2024), 'Ensemble deep learning and efficientnet for accurate diagnosis of diabetic retinopathy', *Scientific Reports 2024 14:1* **14**, 1–16.
URL: *https://www.nature.com/articles/s41598-024-81132-4*

Asia, A. O., Zhu, C. Z., Althubiti, S. A., Al-Alimi, D., Xiao, Y. L., Ouyang, P. B. & Al-Qaness, M. A. (2022), 'Detection of diabetic retinopathy in retinal fundus images using cnn classification models', *Electronics 2022, Vol. 11, Page 2740* **11**, 2740.
URL: *https://www.mdpi.com/2079-9292/11/17/2740/htm https://www.mdpi.com/2079-9292/11/17/2740*

Chintamreddy, D. & Seshasayee, U. R. (2024), 'Detection of diabetic retinopathy (dr) severity from fundus photographs using conv-vit', *International Conference on Advancements in Power, Communication and Intelligent Systems, APCI 2024* .

Hasan, D. A., Zeebaree, S. R., Sadeeq, M. A., Shukur, H. M., Zebari, R. R. & Alkhayyat, A. H. (2021), 'Machine learning-based diabetic retinopathy early detection and classification systems - a survey', *1st Babylon International Conference on Information Technology and Science 2021, BICITS 2021* pp. 16–21.

Khandakar, M. S., Samsuddoha, M., Jahan, S. & Faisal, R. H. (2025), 'An ensemble learning method for early detection of diabetic retinopathy in healthcare analytics'.
URL: *https://papers.ssrn.com/abstract=5334127*

Logeeth, R. & David, D. B. (2025), 'Enhancing the accuracy of early detection of diabetic retinopathy disease using decision tree (dt) compared with support vector machine (svm)', *INNOVATIONS IN THERMAL, MANUFACTURING, STRUCTURAL AND ENVIRONMENTAL ENGINEERING: ICITMSEE'24* **3267**, 020257.
URL: */aip/acp/article/3267/1/020257/3349742/Enhancing-the-accuracy-of-early-detection-of*

Mewada, A., Maurya, S. K. & Ansari, M. A. (2025), 'Seeing beyond: Advanced image and thermal analysis for early detection of diabetic retinopathy and diabetes', *Biomedical and Pharmacology Journal* **18**, 191–202.

Modi, P. & Kumar, Y. (2023), 'Smart detection and diagnosis of diabetic retinopathy using bat based feature selection algorithm and deep forest technique', *Computers and Industrial Engineering* **182**.

Mohan, N. J., Murugan, R., Goel, T. & Roy, P. (2022), 'Vit-dr: Vision transformers in diabetic retinopathy grading using fundus images', *IEEE Region 10 Humanitarian Technology Conference, R10-HTC* **2022-September**, 167–172.

Mohanty, C., Mahapatra, S., Acharya, B., Kokkoras, F., Gerogiannis, V. C., Karamitsos, I. & Kanavos, A. (2023), 'Using deep learning architectures for detection and classification of diabetic retinopathy', *Sensors 2023, Vol. 23, Page 5726* **23**, 5726.
URL: *https://www.mdpi.com/1424-8220/23/12/5726/htm https://www.mdpi.com/1424-8220/23/12/5726*

Nazih, W., Aseeri, A. O., Atallah, O. Y. & El-Sappagh, S. (2023), 'Vision transformer model for predicting the severity of diabetic retinopathy in fundus photography-based retina images', *IEEE Access* **11**, 117546–117561.

Padmanayana & Anoop, D. A. (2022), 'Binary classification of dr-diabetic retinopathy using cnn with fundus colour images', *Materials Today: Proceedings* **58**, 212–216.
URL: *https://www.sciencedirect.com/science/article/abs/pii/S2214785322005429*

Saleh, I., El-Den, N. N., Elsharkawy, M., Mahmoud, A., Sewelam, A., Wang, W., Ghazal, M. & El-Baz, A. (2025), 'Ai-based methods for diagnosing and grading diabetic retinopathy: A comprehensive review', *Artificial Intelligence in Medicine* p. 103221.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0933365725001563*

Shahriari, M. H., Sabbaghi, H., Asadi, F., Hosseini, A. & Khorrami, Z. (2023), 'Artificial intelligence in screening, diagnosis, and classification of diabetic macular edema: A systematic review', *Survey of Ophthalmology* **68**, 42–53.

Silva, P. S., Zhang, D., Jacoba, C. M. P., Fickweiler, W., Lewis, D., Leitmeyer, J., Curran, K., Salongcay, R. P., Doan, D., Ashraf, M., Cavallerano, J. D., Sun, J. K., Peto, T. & Aiello, L. P. (2024), 'Automated machine learning for predicting diabetic retinopathy progression from ultra-widefield retinal images', *JAMA Ophthalmology* **142**, 171–177.

Sugeno, A., Ishikawa, Y., Ohshima, T. & Muramatsu, R. (2021), 'Simple methods for the lesion detection and severity grading of diabetic retinopathy by image processing and transfer learning', *Computers in Biology and Medicine* **137**, 104795.
URL: *https://www.sciencedirect.com/science/article/abs/pii/S0010482521005898*

Sushith, M., Sathiya, A., Kalaipoonguzhali, V. & Sathya, V. (2025), 'A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images', *Scientific Reports* **15**, 1–31.
URL: *https://www.nature.com/articles/s41598-025-99309-w*

Thanikachalam, V., Kabilan, K. & Erramchetty, S. K. (2024), 'Optimized deep cnn for detection and classification of diabetic retinopathy and diabetic macular edema', *BMC Medical Imaging* **24**, 1–17.
URL: *https://link.springer.com/articles/10.1186/s12880-024-01406-1 https://link.springer.com/article/10.1186/s12880-024-01406-1*

Tisha, S. T., Tista, S. C. & Hoq, M. N. (2025), 'A deep learning approach for classifying retinal diseases and diabetic retinopathy stages to enhance early detection in diagnosis', *2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025* .

Vives-Boix, V. & Ruiz-Fernández, D. (2021), 'Diabetic retinopathy detection through convolutional neural networks with synaptic metaplasticity', *Computer Methods and Programs in Biomedicine* **206**, 106094.
URL: *https://www.sciencedirect.com/science/article/abs/pii/S0169260721001693*

Wu, J., Hu, R., Xiao, Z., Chen, J. & Liu, J. (2021), 'Vision transformer-based recognition of diabetic retinopathy grade', *Medical Physics* **48**, 7850–7863.
URL: */doi/pdf/10.1002/mp.15312 https://onlinelibrary.wiley.com/doi/abs/10.1002/mp.15312 https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.15312*