**Project Title:**

**Diabetic Retinopathy Classification Using Deep Learning Techniques**

**Project code:**

ZN-CS-2025-146

**Project Advisors**

Dr Hassam Ali (Internal)

Dr Muhammad Fayyaz (External)

**Project Manager**

Prof Dr Muhammad Ilyas

**Project Team**

Member 1 : Zarnain Javed (Group Leader)

Roll no : BSCS51F22S022

Member 2 : Narmeen Ayub

Roll no : BSCS51F22S011

**Table of Contents**

[1. Abstract 3](#_Toc210398093)

[2. Background and Justification 4](#_Toc210398094)

[3. Research Methodology 4](#_Toc210398095)

[4. Project Scope 5](#_Toc210398096)

[5. High Level Project Plan 6](#_Toc210398097)

[6. References 6](#_Toc210398098)

**Diabetic Retinopathy Classification Using Deep Learning Techniques**

# 1. Abstract

Diabetic Retinopathy (DR) is a leading cause of vision impairment worldwide, demanding timely diagnosis. While deep learning has achieved promising accuracy in DR detection, imbalanced datasets, lack of lesion-specific interpretability, and reliance on fundus-only data hinder clinical deployment. This project proposes a multi-modal deep learning framework that integrates fundus images with patient clinical metadata (e.g., age, diabetes duration, HbA1c levels) to improve DR classification. A knowledge-guided attention module will be introduced to focus on clinically significant lesions (microaneurysms, exudates, hemorrhages). To mitigate data imbalance, synthetic retinal images will be generated using diffusion models, ensuring fair training across DR severity levels.

**The objectives are:**

1. To develop a multi-modal fusion network that combines fundus imaging with clinical metadata (like age, gender etc )
2. To incorporate a knowledge-guided attention mechanism for lesion-localized model interpretability.
3. To eliminate class imbalance by generating a balanced dataset using a two-stage augmentation strategy: primary data creation via a fast, single-step Consistency/Diffusion Model, supplemented by secondary classical transformations for enhanced robustness.
4. To conduct a comprehensive evaluation on benchmark datasets (APTOS, EyePACS) using robust performance metrics and explainability analysis.

The expected outcome is an explainable, clinically reliable, and lightweight model deployable in real-time DR screening, especially in low-resource settings.

# 2. Background and Justification

* **Problem:** Diabetic Retinopathy (DR) grading is highly imbalanced (most images belong to “No DR”), making automated detection unreliable. Current CNN-based systems (DenseNet, EfficientNet, MobileNet with attention) achieve strong accuracy [[1](#_ENREF_1)] but still:
  + Fail in under-represented classes (mild, moderate DR).
  + Provide weak clinical interpretability (black-box outputs).
  + Depend only on fundus images, ignoring valuable clinical metadata [[2](#_ENREF_2)].
* **Gap in Literature:**
  + Dual-attention CNNs (e.g., GAB + CAB) improve imbalance handling but remain single-modal [[3](#_ENREF_3)], [ [4](#_ENREF_4)].
  + GAN-based augmentation exists, but high-generative AI models (e.g., Consistency and Diffusion) enable superior synthetic data creation [[5](#_ENREF_5)] but remain underexplored in DR classification.
  + Very few works integrate fundus + patient metadata for multi-modal learning [[2](#_ENREF_2)].
* **Justification:**  
  This research fills the gap by combining multi-modal data fusion, synthetic data generation, and knowledge-guided attention to deliver an interpretable, generalizable DR classifier.

# 3. Research Methodology

**Step 1**: Dataset Preparation

* Datasets: APTOS (3,669 images) + EyePACS (35,000 images).
* Metadata Integration: Age, gender, diabetes duration, HbA1c (if available in subsets).
* Preprocessing: Rescaling (512×512), CLAHE, normalization.
* Augmentation:Flips, rotations, and synthetic DR image generation using Consistency Models or Diffusion Models (e.g., DDPM / Stable Diffusion fine-tuned on DR) for class balancing and dataset diversity [[5](#_ENREF_5)].

**Step 2**: Model Design

* Backbones: DenseNet-169, EfficientNet-b0 [[1](#_ENREF_1)].
* Fusion: CNN + MLP + Transformer.
* Knowledge-guided attention.
* Classifier.

**Step 3:** Feature Engineering

* **Fundus Features:** Extracted by CNN backbones, representing lesions (microaneurysms, hemorrhages, exudates, vessel changes).
* **Metadata Features:** Extracted by MLP from patient data (age, gender, HbA1c, duration).
* **Multi-Modal Features:** Fusion via concatenation + transformer encoder to capture cross-relations between fundus and metadata [[2](#_ENREF_2)].

**Step 4**: Training & Evaluation

* Optimizer: Adam(Adaptive Moment Estimation), learning rate schedule.
* Metrics: Accuracy, F1, Precision, Sensitivity, Specificity, Quadratic Weighted Kappa.
* Explainability: Grad-CAM + lesion-specific saliency maps.

**Step 5 :** Validation

* Cross-dataset evaluation (train on APTOS, test on EyePACS, and vice versa).
* Comparison with existing methods (dual-attention CNNs, GAN-augmented DR classifiers).

# 4. Project Scope

In Scope:

* Multi-modal DR classification (fundus + metadata).
* Attention-guided lesion interpretability.
* Synthetic dataset generation using diffusion/consistency models.
* Validation on APTOS &EyePACS datasets.

Out of Scope:

* Clinical deployment in hospitals (prototype only).
* Real-time lesion segmentation/localization.
* Integration with telemedicine platforms (future work).

**5. High Level Project Plan**

| **Phase** | **Activities** | **Timeline** | **Resources** |
| --- | --- | --- | --- |
| Month 1–2 | Literature review, finalize proposal, dataset acquisition | 2 months | Papers, Kaggle datasets |
| Month 3–4 | Preprocessing & metadata integration | 2 months | Python, OpenCV, Ophthalmologist input |
| Month 5–6 | Synthetic image generation via diffusion models | 2 months | Stable Diffusion, GPUs |
| Month 7–8 | Multi-modal model design (CNN + MLP fusion + attention) | 2 months | PyTorch/TensorFlow |
| Month 9–10 | Training, fine-tuning, evaluation on APTOS/EyePACS | 2 months | GPU (Colab/Server) |
| Month 11 | Explainability (Grad-CAM, lesion priors), ablation studies | 1 month | Explainability toolkits |
| Month 12 | Report writing, paper drafting, final submission | 1 month | MS Word/LaTeX |

# 6. References

[1] M. M. Farag, M. Fouad, and A. T. Abdel-Hamid, "Automatic severity classification of diabetic retinopathy based on densenet and convolutional block attention module," *IEEE Access,* vol. 10, pp. 38299-38308, 2022.

[2] M. Z. Atwany, A. H. Sahyoun, and M. Yaqub, "Deep learning techniques for diabetic retinopathy classification: A survey," *IEEE Access,* vol. 10, pp. 28642-28655, 2022.

[3] A. He, T. Li, N. Li, K. Wang, and H. Fu, "CABNet: Category attention block for imbalanced diabetic retinopathy grading," *IEEE Transactions on Medical Imaging,* vol. 40, pp. 143-153, 2020.

[4] A. Hannan, Z. Mahmood, R. Qureshi, and H. Ali, "Enhancing diabetic retinopathy classification accuracy through dual-attention mechanism in deep learning," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization,* vol. 13, p. 2539079, 2025.

[5] A. Goel, P. Rao, P. Bhat, R. Ramesh, and S. Natarajan, "An exploration of diffusion models in the context of data augmentation," *2024 10th International Conference on Applied System Innovation (ICASI),* pp. 232-234, 2024.