

# Diabetic Retinopathy Detection Using a Hybrid CNN-Transformer Model

## Introduction

Introducing this project to develop an automatic system to detect diabetic retinopathy from the retina image System classification of images following:

1. No\_Dr
2. Mild
3. Moderate
4. Severe
5. Proliferative DR

## Problem Statement

Diabetes retinopathy is a major complication of diabetes that can lead to blindness if it is not detected early The challenge is to build a model that can accurately classify the retinal image in the corresponding diabetes retinal stages

## Approach

### Data Preparation

Read and convert images read by OpenCV and convert from BGR to RGB Each image is changed in size at  $224 \times 224$  pixels .

### Model Architecture

A hybrid model based on Resnet-50, a combined neural network (CNN) used to extract deep visual information Transformer Vision (VI): self-management mechanisms are used to process the same information while capturing the overall context Fusion: Two features of networks (each feature with a vector at 1000 dimensions) is crushed and introduced in a completely connected class, creating predictions for five classes CNN combination and transformation creates a modern hybrid technique, taking advantage of the characteristics of the two systems It is increasingly used in research and applications that require reliable image classification

### Process of Training

The organize was prepared with cross-entropy misfortune.

AdamW was chosen as an optimiser due to its viability in preparing profound systems.

Preparing Circle:

The show was prepared utilizing 20 ages and a clump measure of 16.

Validation accuracy was tracked, and the model with the highest validation accuracy was saved.

To simplify deployment, the top-performing model (.pth file) was converted to TorchScript (.pt file) using PyTorch's `torch.jit.trace`.

## Benchmarking Results

- **Validation Performance:**
  - The model's accuracy was consistently increased over time.  
The best model was saved based on the highest validation accuracy recorded during training.
- **Modern Approach:**
  - The hybrid CNN-Transformer architecture is a relatively new image categorisation method.  
The usage of pre-trained models, such as ResNet-50 and ViT, ensures that the model benefits from transfer learning.  
Data augmentation and normalisation can help to improve model performance.  
Overall, the method is both contemporary and effective for the intended purpose.  
Conclusion  
The researchers successfully developed a hybrid model for detecting diabetic retinopathy. The model leverages the strengths of CNNs and Transformers to deliver a cutting-edge solution. This strategy is popular in recent research and useful for practical applications due to its validation performance and the usage of new approaches.

## Conclusion

The researchers successfully developed a hybrid model for detecting diabetic retinopathy. The model leverages the strengths of CNNs and Transformers to deliver a cutting-edge solution. This strategy is popular in recent research and useful for practical applications due to its validation performance and the usage of new approaches.