# Early Detection of Diabetic Retinopathy Using Hybrid Deep Learning Models

Author : Charulatha R

Domain : Machine Learning

Sub-domain : Deep Learning

Specialization: Image Processing

Programming Language : Python 3.7+

Deep Learning Frameworks : TensorFlow

Notebook : PYCharm

Year : 2025

## ABSTRACT

## This project proposes a hybrid deep learning model integrating EfficientNetV2, Swin Transformer, and CBAM Attention for robust feature extraction and classification, followed by MLP for improved prediction accuracy. The model classifies retinal images into multiple DR stages and achieved high performance on the APTOS 2019 Blindness Detection dataset.

## METHODOLOGY / WORKFLOW

The workflow of the proposed model is as follows:

1. Dataset: Used APTOS dataset containing retinal images labeled into 5 DR severity levels.

2. Preprocessing: Applied resizing, normalization, and contrast enhancement to improve image clarity.

3. Feature Extraction: Used EfficientNetV2 and Swin Transformer for deep feature extraction.

4. Attention Mechanism: Integrated CBAM to enhance key spatial and channel features.

5. Classification: Used MLP for final DR stage prediction based on fused features.

## MODEL ARCHITECTURE

The hybrid model architecture combines deep feature extraction with advanced attention and boosting mechanisms. The extracted features from EfficientNetV2 and Swin Transformer are concatenated and passed through CBAM attention. The resulting feature vector is classified by MLP into DR stages (0–4).

## RESULTS AND EVALUATION

The model achieved strong performance on the APTOS dataset, summarized as follows:

|  |  |
| --- | --- |
| Metric | Value |
| Accuracy | 92.8% |
| Precision | 90.4% |
| Recall | 89.7% |
| F1-score | 89.9% |
| AUC | 0.94 |

The integration of attention and hybrid features improved accuracy and interpretability compared to standalone CNN models.

## KEY FINDINGS

* Hybrid model improved performance over standalone CNN architectures.
* Attention module enhanced lesion localization accuracy.
* Swin Transformer boosted feature generalization on complex retinal images.

## FUTURE SCOPE

* Deploy model into a web or mobile application for real-time diagnosis.
* Extend dataset with multi-source retinal images for better generalization.
* Integrate explainable AI methods to enhance clinical interpretability.

## REFERENCES

APTOS Blindness Detection Dataset – Kaggle