Proposal: DR-KANTreeNet for Diabetic Retinopathy Classification

Web Link:

https://github.com/ychen2537-chen/CS566-Project-DR-KANTreeNet-for-Diabetic-Retinopathy-Classification

Problem Statement

Diabetic Retinopathy (DR) is one of the leading causes of preventable blindness worldwide. Early and accurate detection is essential for timely treatment, but current automated screening systems face challenges such as class imbalance (many more normal images than severe DR), subtle lesion variations, and difficulty in capturing fine-grained vascular changes.

Conventional convolutional neural networks (CNNs) or vision transformers (ViTs) alone struggle to model both local lesion features and global vessel-tree structures effectively.

Motivation

The goal of this project is to develop an enhanced computer vision model that can robustly classify retinal fundus images into the five standard DR grades.

By combining lesion-focused attention, vascular structure modeling, and multi-scale fusion, this work aims to achieve higher diagnostic accuracy, especially for rare but clinically critical classes (moderate to severe DR).

This would directly benefit ophthalmology screening programs and could be extended to other medical imaging domains.

Proposed Approach

We will implement **DR-KANTreeNet**, a novel model with the following innovations (as described in the uploaded source code):

1. **Lesion-Aware Attention**

   An attention mechanism applied to ResNet-50 features, guided by lesion-specific cues and severe DR detectors.

   Enhances sensitivity to subtle pathological signs.

2. **Vessel-Tree Modeling (VesselTreeNet)**

   A KAN-based CNN designed to extract and analyze vascular “tree-like” structures from fundus images.

   Uses OpenCV or PIL preprocessing for vessel segmentation, followed by a convolutional pipeline with KAN layers.

3. **Quad-Modal Fusion**

   Combines local CNN features, global ViT features (small and large), and vascular features.

   Employs KAN-based gating networks and graph convolution for refined decision-making.

4. **Improved Training Techniques**

Advanced augmentation (random rotation, color jitter, perspective transforms, severe-case augmentation).

Class-balanced focal loss and mixup regularization to address imbalance.

Multi-GPU support with memory-efficient KAN layers.

Expected Outcomes

1. Improved weighted F1 score and Cohen's Kappa compared to baseline CNN/ViT models.
2. More balanced per-class performance, particularly on severe DR.
3. A robust and scalable training pipeline suitable for multi-GPU environments.

Rough Timeline

Week	Tasks
1	Literature review; dataset preparation (APTOS/Kaggle DR).
2	Implement baseline ResNet50 + ViT classifiers; establish metrics.
3	Integrate **Enhanced Lesion Attention**; run initial experiments.
4	Develop and test **VesselTreeNet** module with vessel segmentation.
5	Implement **Quad-Modal Fusion** with KAN layers; tune hyperparameters.
6	Train full **DR-KANTreeNet** on multi-GPU; monitor metrics and debug.
7	Validation with cross-validation; generate confusion matrices and per-class F1.
8	Write final report, prepare visualizations (ROC, confusion matrices, ablation studies).

Reference:

Wei, X., Liu, Y., Zhang, F., Geng, L., Shan, C., Cao, X., & Xiao, Z. (2025). MSTNet: Multi-scale spatial-aware transformer with multi-instance learning for diabetic retinopathy classification. *Medical Image Analysis*, *102*, 103511.

Batten, J., Schaap, M., Sinclair, M., Bai, Y., & Glocker, B. (2025). Vector Representations of Vessel Trees. *arXiv preprint arXiv:2506.11163*.

Tao, Q., Ge, Z., Cai, J., Yin, J., & See, S. (2019, October). Improving deep lesion detection using 3D contextual and spatial attention. In *International conference on medical image computing and computer-assisted intervention* (pp. 185-193). Cham: Springer International Publishing.

Karthik, Maggie, and Sohier Dane. APTOS 2019 Blindness Detection. https://kaggle.com/competitions/aptos2019-blindness-detection, 2019. Kaggle.