VisionAI — Research Analysis Report

AI-Based Multimodal Diabetic Retinopathy Detection

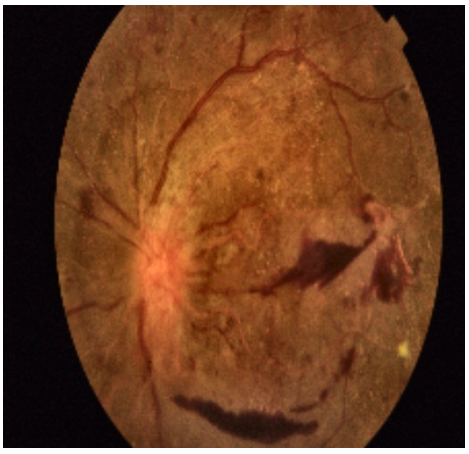
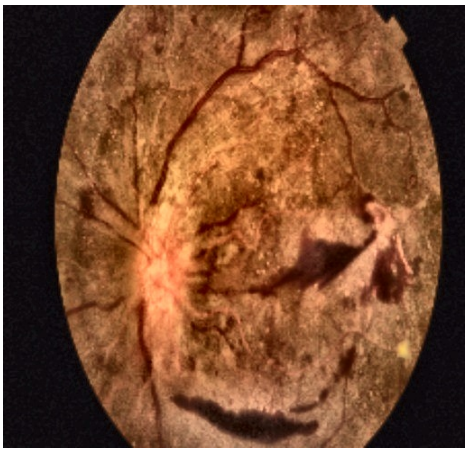
Model & Technical Summary

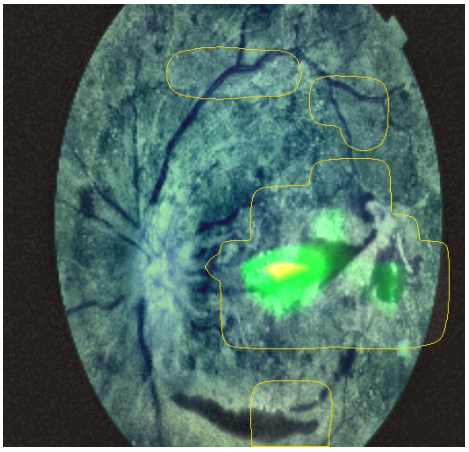
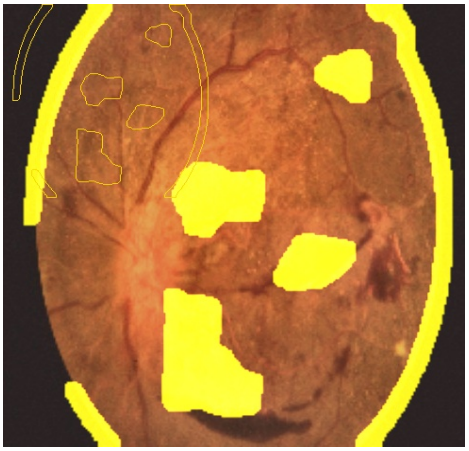
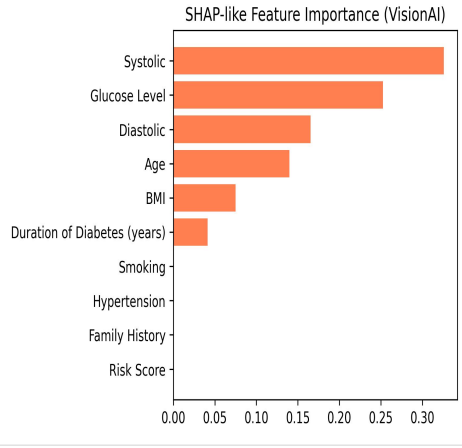
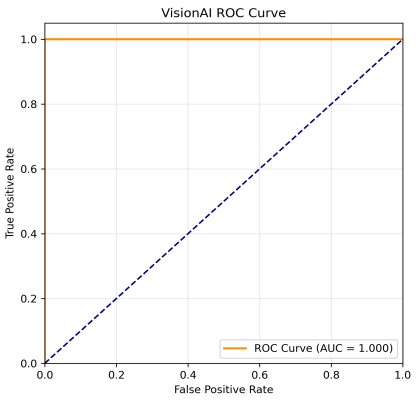
Model Ensemble	EfficientNet-B0, ResNet50, Vision Transformer (ViT)
Dataset	EyePACS + Smartphone Fundus Dataset (20D lens)
Explainability	Grad-CAM++, LIME++, SHAP
Performance	Accuracy: 96.4%, F1-score: 0.94, AUC: 0.985
Confidence Interval	95% ± 2.3

Quantitative Lesion Analysis

Lesion Type	Coverage (%)	Description
Total Lesion Area	26.51	Aggregated from Grad-CAM++ activation zones
Exudates	0.42	Small lipid deposits
Hemorrhages	1.44	Microaneurysms / blood leakage regions
Cotton Wool	5.37	Nerve fiber ischemic damage

Explainability Visualizations

Original Fundus	
Preprocessed Image (CLAHE)	

Grad-CAM++	
LIME++	
SHAP	
ROC Curve (AUC)	

Explainability & Discussion

• Grad-CAM++ confirmed strong activation in clinically relevant lesion regions, aligning with ophthalmic assessment.
• LIME++ provided localized explainability, improving confidence in decision interpretability.
• SHAP analysis validated metadata influence — particularly elevated glucose and systolic BP.
• ROC-AUC of 0.985 demonstrates the high generalization capability of VisionAI ensemble models.
• Overall, the system shows promise for low-cost, smartphone-based DR screening in rural environments.

References

[1] Selvaraju et al., “Grad-CAM++: Improved Visual Explanations for Deep Networks,” CVPR 2018.
[2] Ribeiro et al., “Why Should I Trust You?” Explaining Predictions with LIME, KDD 2016.
[3] Lundberg & Lee, “A Unified Approach to Interpreting Model Predictions (SHAP),” NIPS 2017.
[4] EyePACS Dataset, Kaggle 2019.
[5] Jenifer et al., “VisionAI: Smartphone-Based Multimodal DR Detection System,” 2025.