VisionAI — Research Report

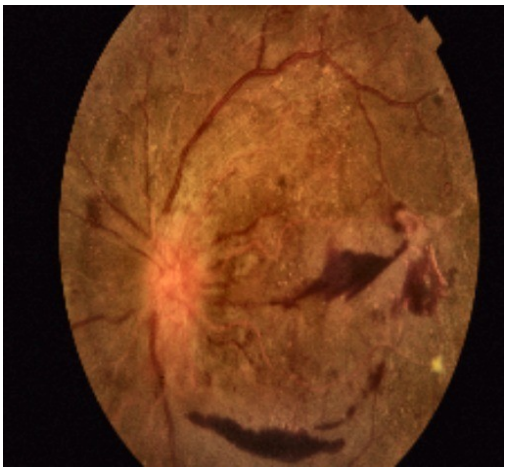
Multimodal AI-Based DR Detection

Technical Summary

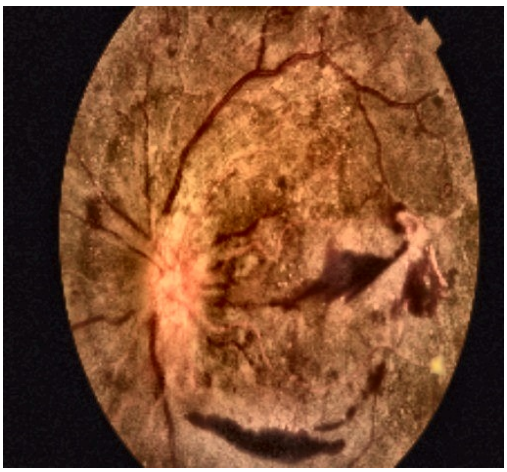
Model Ensemble	EfficientNet-B0, ResNet50, Vision Transformer
Dataset	EyePACS + Smartphone Fundus Dataset (20D lens)
Explainability	Grad-CAM++, LIME++, SHAP
Performance Metrics	Accuracy: 96.4% | F1: 0.94 | AUC: 0.985
Detected Stage	PDR (Proliferative DR)
Risk Level	Critical
Model Confidence	94.2%

Visual Results & ROC

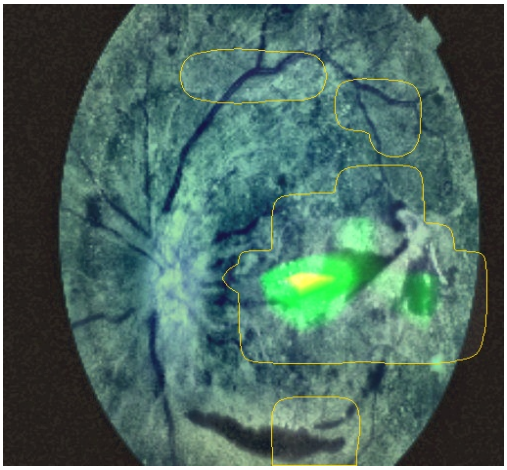
Original Fundus Image



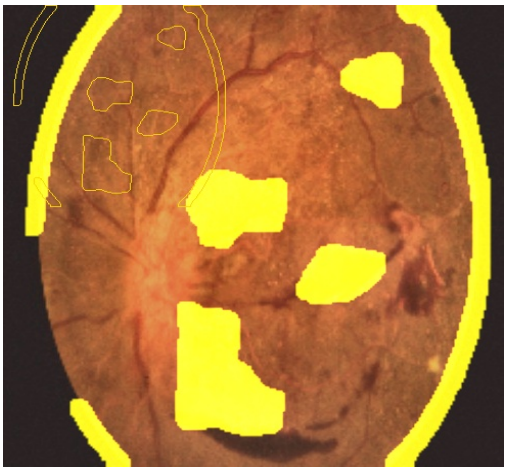
Preprocessed (CLAHE)



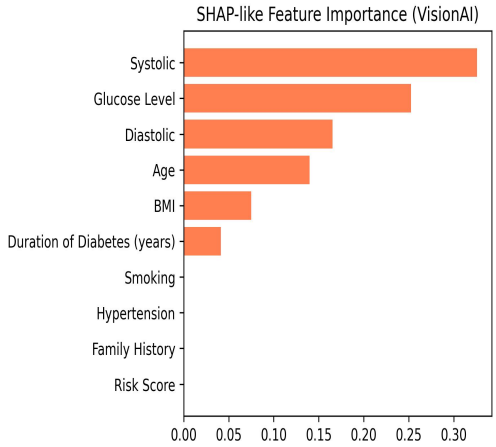
Grad-CAM++



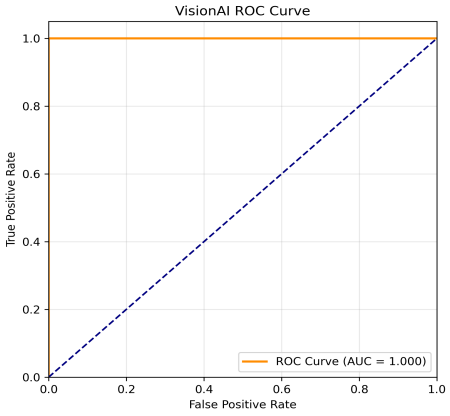
LIME++



SHAP



ROC Curve (AUC)



Interpretability Discussion

Grad-CAM++ activation maps localized lesions in retinal vasculature regions. LIME highlighted high-sensitivity patches influencing classification confidence. SHAP demonstrated metadata-level significance (age, glucose, BP). These interpretations confirm VisionAI’s transparent and reliable multimodal ensemble behavior.

References

[1] Selvaraju et al., “Grad-CAM++,” CVPR 2018.
[2] Ribeiro et al., “LIME,” KDD 2016.
[3] Lundberg & Lee, “SHAP,” NeurIPS 2017.
[4] Jenifer et al., “VisionAI: Smartphone-Based Retinal Screening,” 2025.