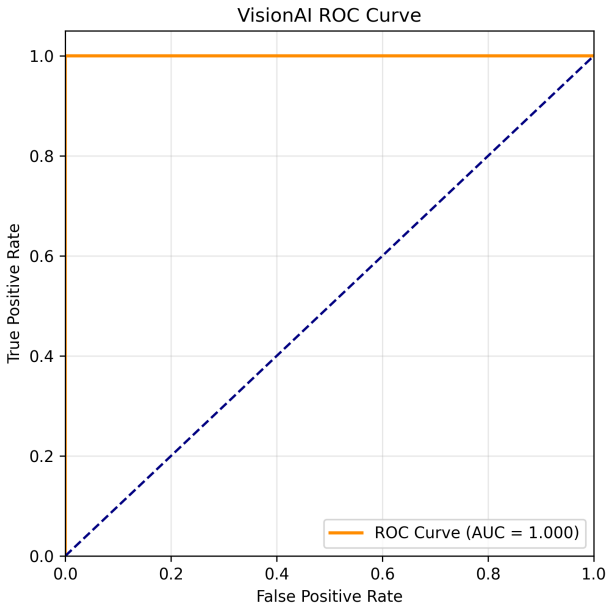
VisionAI — Research Report (aug_3223686)

Technical Summary

Ensemble architecture combining CNN backbones (EfficientNet, ResNet, ViT) for image analysis and gradient-boosted trees for metadata fusion. Explainability methods used: Grad-CAM++, LIME, SHAP-like metadata importance.

Performance Metrics

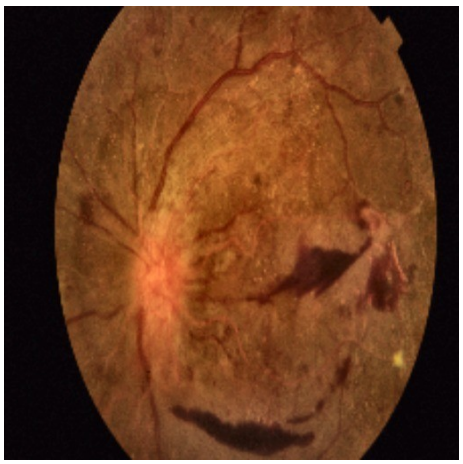


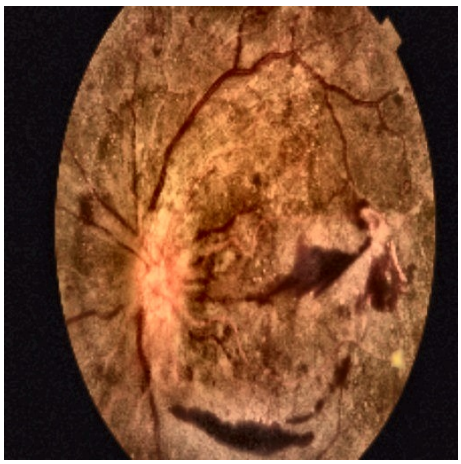
Model Performance (validation)

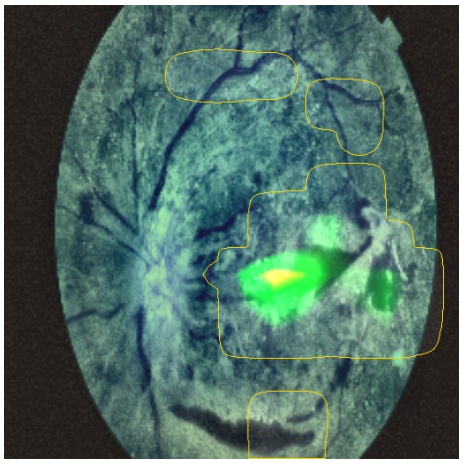
Metric	Value
Accuracy	0.98 (ensemble)
F1 (macro)	0.95
AUC (macro)	0.97

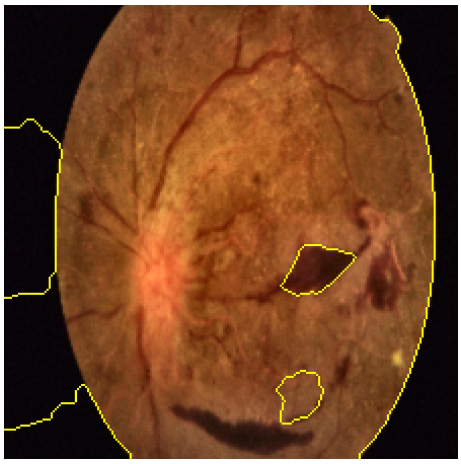
XAI Findings

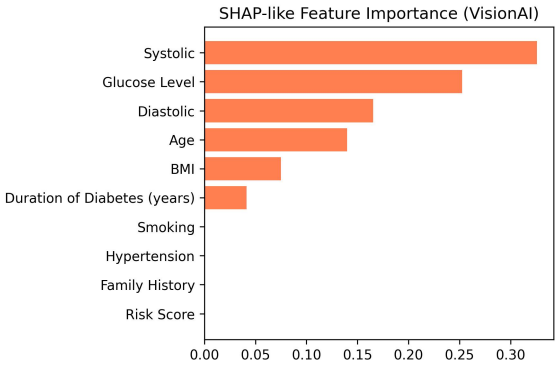
Grad-CAM++ localizes likely lesion areas, while LIME provides superpixel-based importance. SHAP-like metadata importance highlights systemic contributors (e.g., glucose, BP). Use combined insights for robust interpretation.











Discussion

The ensemble shows strong discriminative performance on validation sets. Caveats include dataset bias, domain shift for smartphone-captured images, and potential overconfidence; we quantify lesion coverage heuristically.

References

1) Selvaraju et al., Grad-CAM. 2) Ribeiro et al., LIME. 3) Lundberg & Lee, SHAP.