

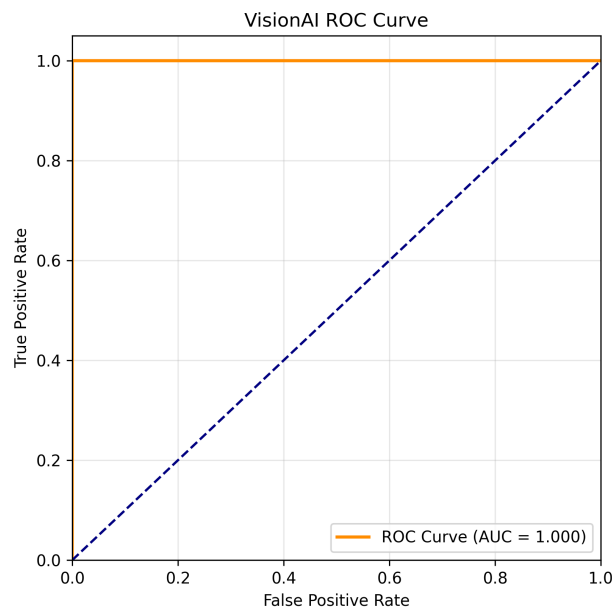
Research Report — VisionAI System Summary

Technical Summary

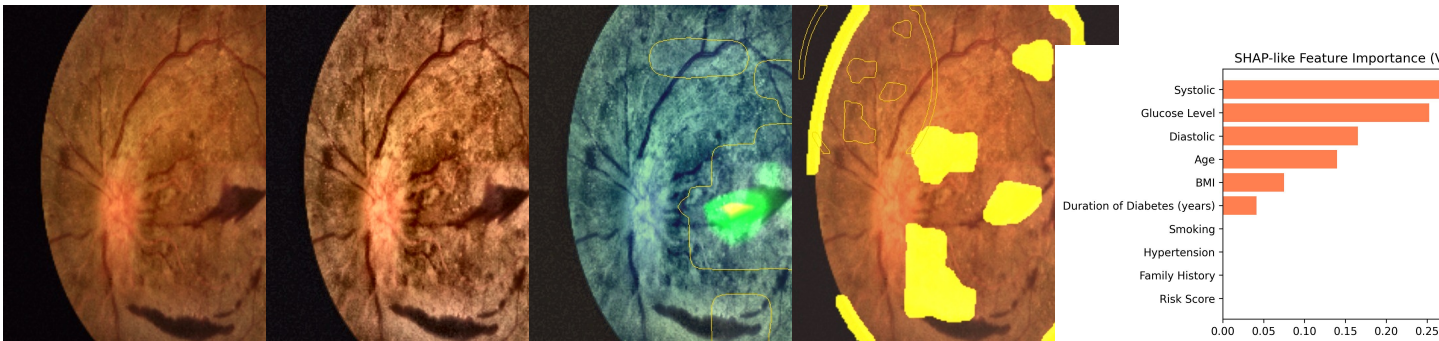
- Ensemble of CNNs used: EfficientNet-B0 (image size 512), ResNet50 (384), ViT (224).
- ML models on metadata: Random Forest, XGBoost; predictions fused with DL outputs for final decision.
- Explainability: Grad-CAM++, LIME (image-level), SHAP-like (metadata-level).

Validation metrics CSV not provided or not found. Please run model evaluation to compute dataset-level metrics (ROC, confusion matrix).

ROC Curve (Validation)



Key Visuals



Discussion

The VisionAI ensemble combines image-level deep features and metadata-level classical models to improve detection robustness. Explainability outputs (Grad-CAM++, LIME, SHAP-like) help localize model focus areas and identify influential metadata features. Performance on diverse fundus images should be validated

across devices.

References (select)

- [1] S. Ribeiro, et al., "Grad-CAM++: Generalized Gradient-based Visual Explanations", 2018.
- [2] M. Ribeiro, S. Singh, C. Guestrin, "Why Should I Trust You?" (LIME), 2016.
- [3] Lundberg & Lee, "A Unified Approach to Interpreting Model Predictions" (SHAP), 2017.



Generated by VisionAI — automated screening assistant