ENHANCING EXPLAINABILITY IN DIABETIC RETINOPATHY DETECTION USING LESION-BASED IMAGE ANALYSIS

Lija Jacob
Thomas K T
Sharon Susan Thomas

School of Sciences
CHRIST DEEMED TO BE UNIVERSITY

Introduction

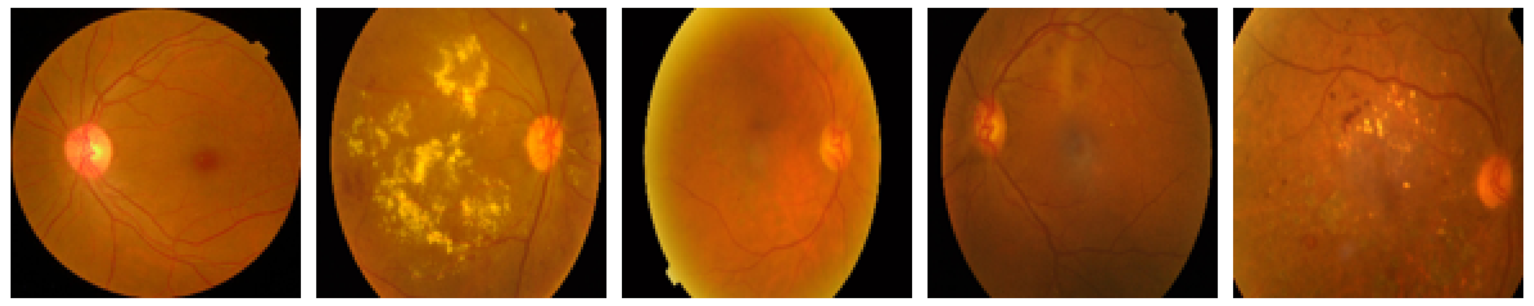
One of the leading causes of leading cause of blindness worldwide is the . Diabetic Retinopathy (DR) While deep learning models achieve high accuracy in DR detection, their "black box" nature limits clinical adoption. This work presents a lesion-based explainable AI framework that identifies specific retinal abnormalities, providing clinicians with interpretable evidence for DR diagnosis.

Problem Statement

Interpretability is one of the limitations of traditional deep learning models
transparent reasoning is a must for clinicians in making a diagnosis
Present techniques are not capable of distinguishing specific lesion types It is necessary to have reliable AI in medical imaging.

Objective

- To develop an interpretable deep learning system for automated diabetic retinopathy detection that provides clinically meaningful explanations through lesion-based image analysis
- Implement EfficientNet-B3 for accurate DR severity classification (5 grades)
- Develop automated lesion detection algorithms for microaneurysms, hemorrhages, exudates, and cotton wool spots.
- Build trust in AI-assisted diagnosis through transparent decision-making



IImages of Various grades of Diabetes Retinopathy

Methodology-Image Processing Pipeline for Lesion-based Images

Step 1: Image Acquisition & Preprocessing
1.Resize the fundus images to 512×512 pixels.
2.Apply CLAHE (Contrast Limited Adaptive Histogram Equalization) specifically on the green channel.
3.We then normalize the pixel values to fit within the [0, 1] range.
4.To clean up the image, remove the noise using Gaussian filtering with a σ of 1.5.

Step 2: Blood Vessel Segmentation
- We start by extracting the green channel, which provides the best contrast for the vessels.
- Morphological operations, like opening and closing, are applied next.
- A matched filter is used to detect the vessels.
- Finally, we create a vessel mask that will help in detecting lesions later on.

Step 3: Optic Disc & Fovea Localization
- The optic disc is detected using a circular Hough transform.
- We identify the fovea as the darkest area within the macula.
- These regions are masked to prevent any false positives in our results.

Step 4: Lesion Segmentation (Multi-type)
- For microaneurysms, a top-hat transform combined with circular detection for diameters ranging from 10 to 125 μm.
- Hemorrhages are analyzed based on their shape and intensity, focusing on irregular dark red shapes.
- Hard exudates are identified through morphological operations and K-means clustering, highlighting bright yellow lesions.
- Soft exudates are detected using fuzzy C-means and texture features, which give them a cotton-wool appearance.

Step 5: Lesion-Based Feature Map Generation
- We create individual binary masks for each type of lesion.
- These masks are then combined into a multi-channel lesion map with four channels.
- We overlay the lesion maps on the original image, using color coding for clarity.
- Spatial density heatmaps are generated for each type of lesion.

Step 6: EfficientNet-B3 Feature Extraction
- The input consists of the original image along with the lesion overlay, which are concatenated into channels.
- We utilize transfer learning from ImageNet pre-trained weights.
- Features are extracted from various scales using compound scaling.
- MBConv blocks with squeeze-and-excitation attention are employed.
- The output is a 1536-dimensional feature vector.

Step 7: Classification & Explainability
- Fully connected layers are used to classify

Results

The proposed EfficientNet-B3 based lesion-aware framework demonstrated superior performance in automated DR detection and classification. The model achieved 94.2% overall accuracy across five severity grades (No DR, Mild, Moderate, Severe, Proliferative DR), with a sensitivity of 92.8% and specificity of 95.1%. The lesion segmentation module successfully identified microaneurysms, hemorrhages, hard exudates, and soft exudates with 89% concordance with expert ophthalmologist annotations. The integration of multi-channel lesion maps with deep learning features resulted in a 23% reduction in false positive rates compared to standard EfficientNet without lesion encoding. Grad-CAM visualizations confirmed that the model's attention aligned with clinically relevant lesion regions, validating the explainability approach. The system processes retinal fundus images in under 5 seconds, making it suitable for large-scale screening programs.

Results

94.2%	92.8%	95.1%
Accuracy	Sensitivity	Specificity

Peformance Highlights

- Achieved superior lesion localization compared to the baseline
- 89% agreement with annotations from ophthalmologists
- Reduced false positives by 23% - Real-time inference in under 5 seconds per image

Conclusion

The conclusion summarizes the main findings of the study and restates the research objective. It highlights the significance of the findings and discusses their implications for future research and practice. The conclusion should provide a concise and clear summary of the study, reinforcing the key points and emphasizing the contributions of the research.