17th Asian Conference on Computer Vision

Cross Feature Fusion of Fundus Image and Generated Lesion Map for Referable Diabetic Retinopathy Classification

Dahyun Mok, Junghyun Bum, Duc-Tai Le, Hyunseung Choo
Sungkyunkwan University, Korea

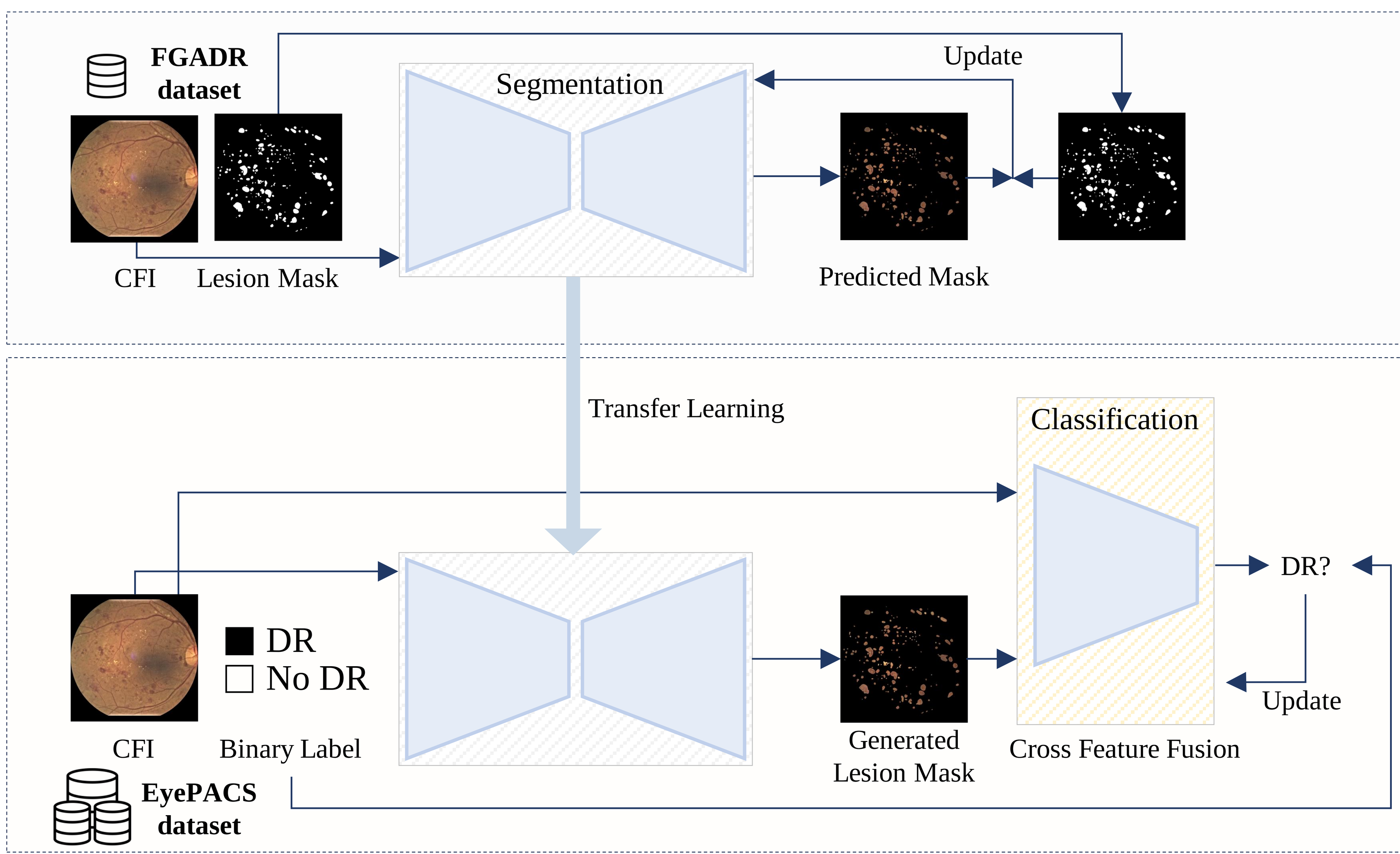
Motivation

- Diabetic Retinopathy (DR) is a primary cause of blindness, necessitating early detection and diagnostic
- This work leverages advanced AI-driven approaches to enhance diagnostic precision and reduce resource burdens in medical imaging applications

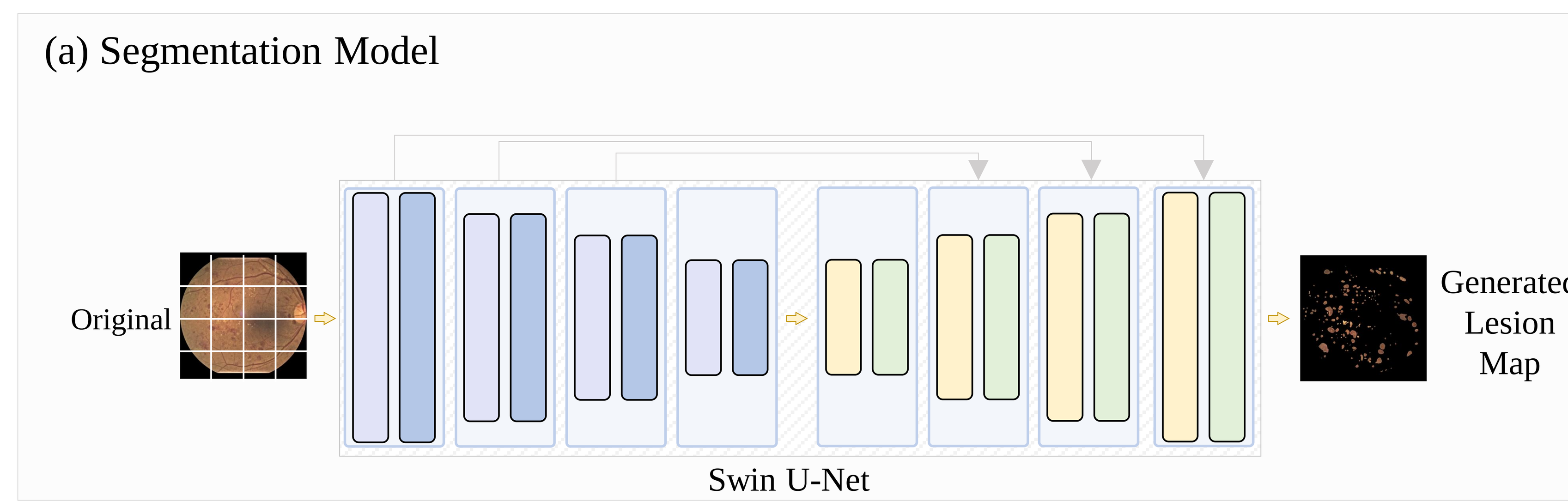
Contribution

- Proposes a novel approach to address the limitation of pixel-level annotations by generating pseudo lesion maps for large-scale unlabeled datasets
- Achieves superior classification performance by integrating cross-feature fusion of fundus images and generated lesion maps using a cross-attention mechanism

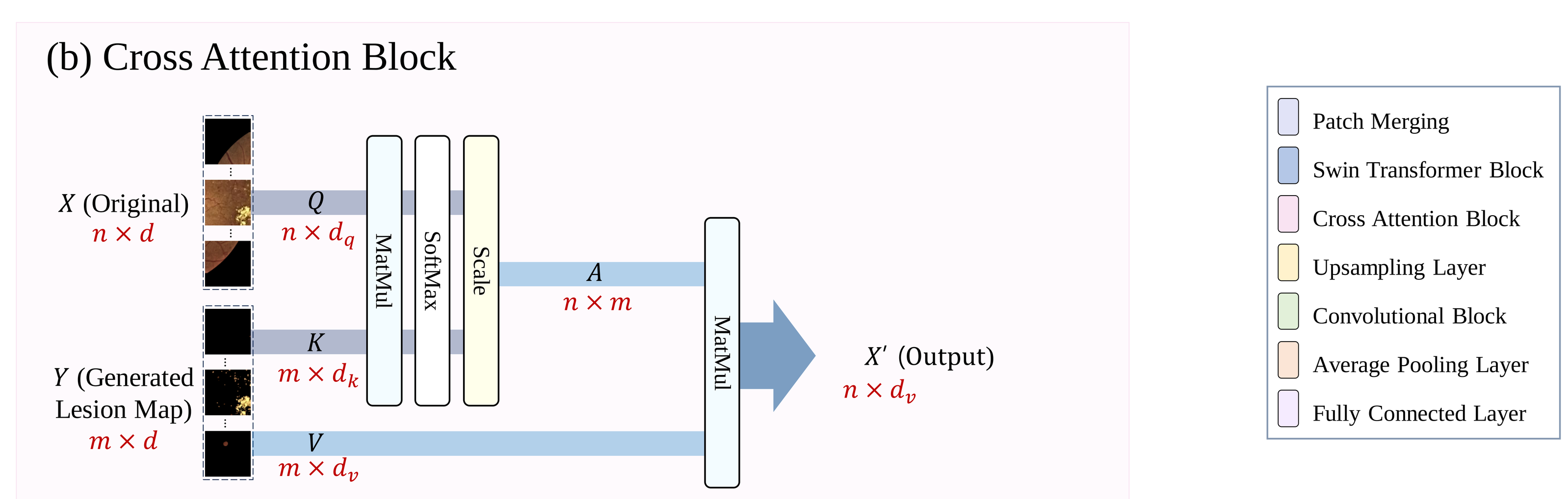
Cross Feature Fusion



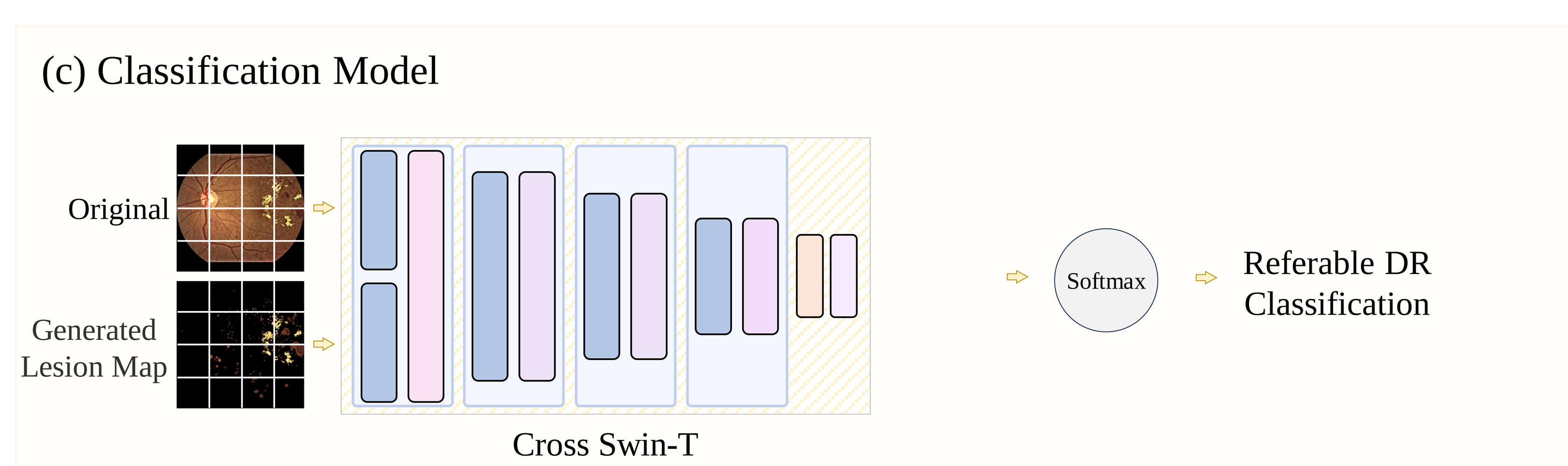
- Dataset:
  - FGADR (pixel level): 1,842 images
  - EyePACS (image-level): 88,702 images
- The architecture of the proposed method consists of two steps
- Segmentation
  - Train lesion segmentation by comparing predictions with ground truth and updating via backpropagation
- Classification
  - Train a model for disease detection using cross-feature fusion of original fundus images and generated lesion maps



Utilize Swin U-Net to generate lesion maps from input



Computes attention weights to align relevant features from both inputs: original image X and generated lesion map Y



The cross-attention mechanism integrates features to enhance lesion focus and improve classification performance

Performance Evaluation

Table 1 Proposed method outperforms SOTA methods

Model	ACC	AUC	TPR	TNR
ResNet50[10]	87.8	88.1	67.2	93.1
Swin-T(Tiny)[26]	91.6	91.7	81.3	92.2
Swin-S(Small)[26]	91.7	91.8	81.7	93.1
Swin-B(Base)[26]	91.7	91.7	81.8	93.4
Swin-L(Large)[26]	91.9	91.9	81.7	91.7
ConvNeXt[27]	87.5	85.1	76.4	81.6
CrossViT[4]	88.0	85.6	76.8	83.2
CoATNet[24]	81.4	79.4	72.6	77.3
Ensemble of Plain and Robust model[21]	90.2	91.0	79.2	88.1
Lesion-Aware Contrastive Learning[16]	84.6	83.8	73.4	85.2
Ours (Proposed Method)	**94.6**	**96.2**	**82.0**	**97.6**

Table 2 Combining cross-feature fusion and transfer learning significantly enhances performance

Cross Feature Fusion of Original Image and Generated Lesion Map	Transfer Learning from Swin U-Net Encoder	ACC	AUC
		91.6	91.5
	✓	93.4	95.0
✓		94.0	95.9
✓	✓	**94.6**	**96.2**

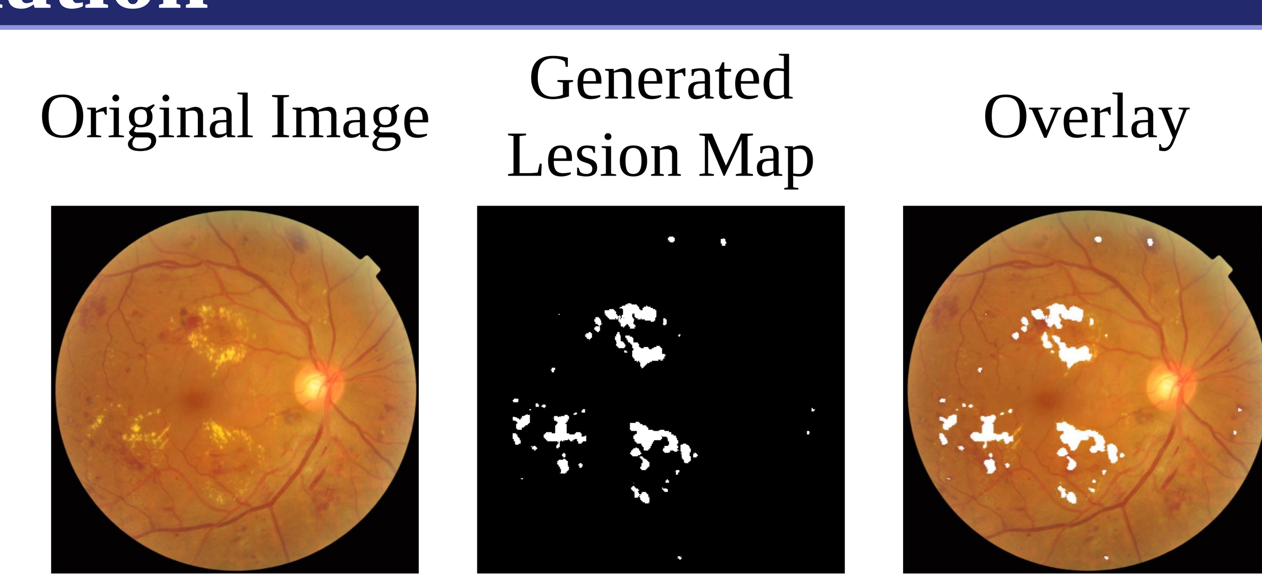


Figure 1 Generated lesion maps demonstrate the segmentation capability to accurately localize lesions

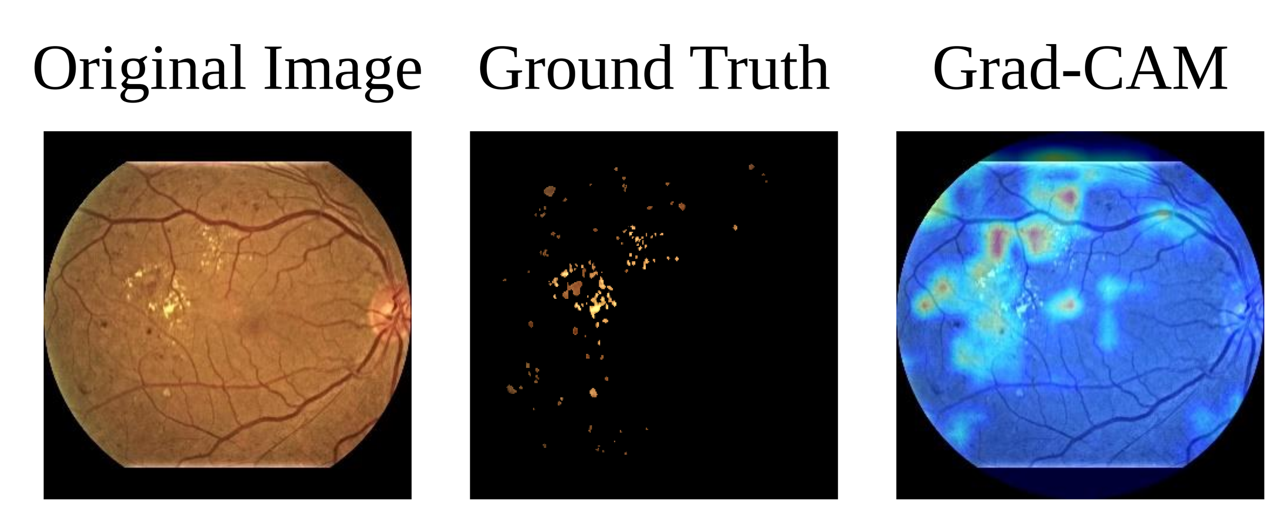


Figure 2 Proposed classification heatmaps closely match ground truth, demonstrating precise lesion localization