THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

ICPR 2020

Robust Localization of Retinal Lesions via Weakly-supervised Learning

Ruohan Zhao, Qin Li, Jane You

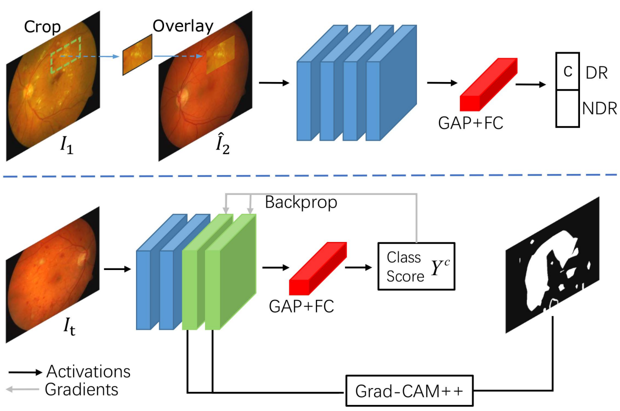
Background

Diabetic retinopathy (DR) is a common microvascular complication and one of the most common cause of vision loss and blindness among diabetics.

Motivation

Propose a localization framework that leverages solely image-level labels, achieving the localization of various lesions in a weakly-supervised manner.

Method



Algorithm 1: Training and localizing procedure

Training procedure:
Input: Training Data $I = \{(I_i, c_i)\}_{i=1}^{N}$
Output: Network parameters θ
While training is not convergent:
Use mix-training $\hat{I}_2, \hat{c} \leftarrow$ Mix-Training(I_1, I_2, c_1, c_2);
Get classification score $Y \leftarrow$ Classifier$(\hat{I}_2)$;
Update $\theta \leftarrow$ BinaryCrossEntropy$(Y, \hat{c})$;

Localization procedure:
Input: The inference image I_t
Output: localization map L
$A_n^k, Y \leftarrow$ Feed-forward(I_t);
For different layers n:
$L_n \leftarrow$ Grad-CAM$^{++}(A_n^k, Y)$;
$L \leftarrow$ Aggregation(L_n)

Mix-training

The classifier would attend to less distinct lesions e.g. tiny circular and deep-red dots.

Crop a random patch from image I_1 and overlaid the cropped patch to the corresponding region of I_2 to synthesize a new training sample $\hat{I}_2$.

$$\hat{I}_2 = \mathbf{M} \odot I_1 + (\mathbf{1} - \mathbf{M}) \odot I_2$$

where $\odot$ represents the element-wise multiplication and $\mathbf{M} \in \{0,1\}^{W \times H}$ is the binary mask that determines the regions to bind two samples. The label $\hat{c}$ of the synthetic sample $\hat{I}_2$ is defined as:

$$\hat{c} = \begin{cases} \lambda c_1 + (1-\lambda) c_2 & c_1 \in \mathrm{DR}, c_2 \in \mathrm{NDR} \\ c_2 & others \end{cases}$$

Inference

Gradients of multiple layers are calculated to derive several localization maps, denoted as L_n:

$$L_n = \mathrm{ReLU}\left(\sum_k w_k \cdot A_n^k\right)$$

Denote the feature map at the n^{th} convolution layer of unit k as A_n^k. w_k is the weight to reflect the significance of feature map.

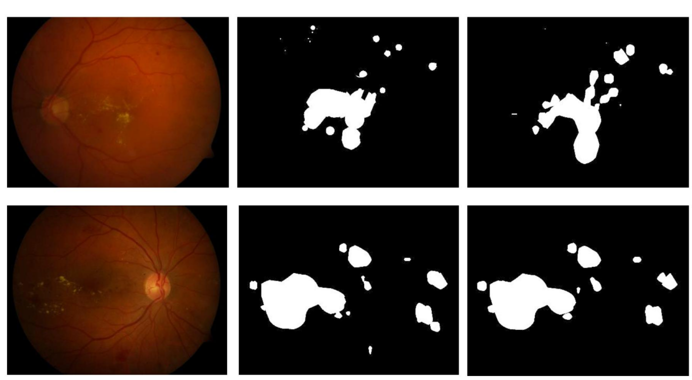
Result

TABLE I: Performance evaluation at lesion-level with other methods on DIARETDB1.

Method	Red lesion				Bright lesion			
	Microaneurysm		Hemorrhages		Hard Exudates		Soft Exudates	
	Sen%	FPI	Sen%	FPI	Sen%	FPI	Sen%	FPI
Chudzik *et al.*	64.1	8	-	-	-	-	-	-
Seoud *et al.*	63.9	8	-	-	-	-	-	-
Quellec *et al.*	61	10	71	10	80	10	90	10
Gondal *et al.*	52	1.5	91	1.5	87	1.5	89	1.5
Ours	63.4	4.2	95.9	2.6	94.2	3.0	93.1	2.3

TABLE II: Sensitivity% at image-level on the DIARETDB1 dataset. The best is shown in bold.

Method	MAs	HEs	Soft Exudates	Hard Exudates
Liu *et al.*	-	-	83.0	83.0
Zhou *et al.*	-	94.4	-	
Zhao *et al.*	-	**98.1**	-	-
Quellec *et al.*	-	94.7	-	-
Gondal *et al.*	50	97.2	90.9	**100**
Ours	**68.9**	97.5	**92.2**	98.5



Fundus image   Ground-truth   Localization result