Automated Diabetic Retinopathy Identification via Lesion Guided Network

Yijin Huang
Department of Electrical and Electronic Engineering, Southern University of Science and Technology
Shenzhen, China

Li Lin
Department of Electrical and Electronic Engineering, Southern University of Science and Technology
Shenzhen, China

Meng Li
State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Centre, Sun Yat-sen University
Guangzhou, China

Jin Yuan
State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Centre, Sun Yat-sen University
Guangzhou, China

Xiaoying Tang*
Department of Electrical and Electronic Engineering, Southern University of Science and Technology
Shenzhen, China
tangxy@sustech.edu.cn

ABSTRACT

In this paper, we propose and validate a novel neural network named Lesion Guided Network (LGN) for automatic diagnosis of Diabetic Retinopathy (DR) from fundus images. RetinaNet is first adopted and trained on a coarsely-annotated dataset for rough lesion detection. Lesion-Aware Module (LAM) in LGN is proposed to highlight regions of interest in fundus images utilizing the coarse lesion maps. Then, the outputs of LAM are fed into a covolutional neural network (CNN) for DR identification. The proposed method is evaluated on a private dataset consisting of 4465 fundus images. Experimental results demonstrate the superiority of the proposed LGN, achieving comparable performance with ophthalmologists.

CCS CONCEPTS

• **Computing methodologies** → *Computer vision*; *Neural networks.*

KEYWORDS

Diabetic retinopathy, lesion guided network

ACM Reference Format:
Yijin Huang, Li Lin, Meng Li, Jin Yuan, and Xiaoying Tang. 2020. Automated Diabetic Retinopathy Identification via Lesion Guided Network. In *The Fourth International Symposium on Image Computing and Digital Medicine (ISICDM 2020), December 5–7, 2020, Shenyang, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3451421.3451452

*Corresponding author is Dr. Xiaoying Tang (tangxy@sustech.edu.cn)

ISICDM 2020, December 5–7, 2020, Shenyang, China

ACM ISBN 978-1-4503-8968-6/20/12…$15.00
https://doi.org/10.1145/3451421.3451452

1 INTRODUCTION

Diabetic Retinopathy (DR) is one of the microvascular complications of diabetes, causing vision impairment and blindness [1]. In DR, an early diagnosis and timely intervention is vital for protecting a patient's visual function. Therefore, an automated, accurate, and efficient screening algorithm based on fundus images is crucial and urgent to reduce the number of untreated patients in a large population.

Recently, with the development of deep learning techniques, convolutional neural networks (CNNs) have been proposed for DR grading making use of fundus images. For example, Ning et al. [2] designed a 13-layer CNN to classify fundus images into normal and referable DR. Pratt et al. [3] applied transfer learning by adopting a pre-trained CNN as a feature extractor and retraining the last fully connected layer for DR detection. Yang et al. [4] improved the DR detection performance by introducing an imbalanced weighting map to give move attention to patches with lesions.

Although deep learning based DR identification approaches have achieved promising performance, existing methods still have some limitations. First, the type and amount of some tiny lesions (eg., small hemorrhages and microaneurysms) are considered as important pathological signs in the early stage of DR [5]. However, these lesions are difficult to be detected by CNNs because their features are prone to vanish in deep layers of vanilla CNNs. Therefore, data-driven methods that take only fundus images as input depend significantly on the amount of data to learn discriminative features for DR detection. Second, although some attention-based methods [4] utilize lesion information to highlight regions of interest, they require high precision of lesion detection, and thus a suitable dataset with lesion annotations is essential. Nevertheless, the process of lesion annotating is time-consuming and tedious, even error-prone, and hence a well-annotated dataset is very limited.

1.1 Dataset

To address the aforementioned issues, we propose a Lesion Guided Network (LGN) to refine coarse lesion maps and then discriminate DR from fundus images. Our pipeline consists of two parts:

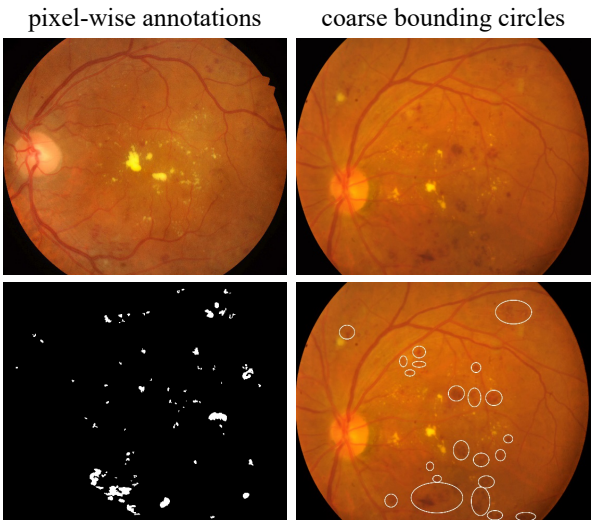

Figure 1: Representative fundus images with two different types of hemorrhage annotations.

Object Detection Network: As mentioned above, a well-annotated dataset such as IDRiD [6] is difficult to obtain. Thus, as shown in Fig. 1, instead of using pixel-wise annotations, we adopt manually-delineated bounding circles to roughly label lesions, and then construct a dataset of 590 samples with coarse lesion annotations [7]. This coarsely-annotated dataset nevertheless may be inaccurate to some degree. RetinaNet [8] with Resnet50 [9] backbone is adopted as the lesion detection network and trained on our dataset to produce coarse lesion maps for subsequent DR identification.
Lesion Guided Network: For existing attention-based methods, coarse lesion maps are of limited help to DR identification, and may even degrade the performance. In this work, a Lesion-Aware Module (LAM) is proposed to refine the coarse lesion maps, which takes a fundus image and the corresponding coarse lesion map as input and outputs a refined lesion map. The refined lesion map is then fused with the original fundus image and fed in a CNN for DR identification. We evaluate the proposed pipeline with LGN on a private dataset. Experimental results show that LAM is capable of refining the coarse lesion map, which guides the classification network to achieve a better DR identification performance.

2 METHOD

A private dataset consisting of 4465 fundus images with category labels (normal versus DR) was collected at Zhongshan Ophthalmic Center, Sun Yat-sen University and Gaoyao People's Hospital, which were acquired using a TOPCON-50DX digital fundus camera with a resolution of 2880×2136. For a evaluation purpose, 3293/390/782 images are respectively used for training/validation/test. For lesion detection, three ophthalmologists labeled exudates, hemorrhages, and microaneurysms in 590 samples using manually-delineated bounding circles. All images with lesion annotations were randomly selected from the training set to avoid data leakage.

2.1 Coarse Lesion Map Generation

RetinaNet with Resnet50 backbone is adopted as our lesion detection CNN. The detection CNN takes a fundus image as input and outputs a probability map of lesions. Then a Gaussian filter is applied to smooth the probability map and generate the lesion map, which is one of the inputs of LGN. Since the lesion detection training dataset is coarsely-annotated, which may contains imprecise bounding circles and miss-annotated lesions, the trained RetinaNet has a relatively limited capability of very accurately detecting lesions from fundus images. As a result, numerous false positives and false negatives exist in the lesion map generated by the trained object detection network.

2.2 Lesion Guided Network

As depicted in Fig. 2, the proposed LGN is comprised of two components: a Lesion-Aware Module and a classification CNN (e.g., ResNet50). As stated in section 2.2, directly feeding a coarse lesion map with numerous errors and uncertainties into a classification network can be harmful to the DR identification. To alleviate this issue, we design a straightforward sub-network, namely Lesion-Aware Module. LAM consists of one 7×7 convolutional layer, three 3×3 convolutional layers and one 1×1 convolutional layer. Except for the 1×1 convolutional layer, each layer is followed by a batch normalization layer, leaky rectified linear unit [10]. The last 1×1 convolutional layer is used to generate a single-channel feature map, followed by a sigmoid layer to produce the final lesion probability map. Taking the original image $I_{ori} \in \mathbb{R}^{w \times h \times c}$ and the corresponding coarse lesion map $L_{coa} \in \mathbb{R}^{w \times h}$ as input, LAM outputs a refined lesion map $L_{ref} \in \mathbb{R}^{w \times h}$ to highlight the regions of interest as follows:

$$L_{ref} = \sigma(W_{lam} \cdot L_{coa} + b_{lam}), \tag{1}$$

$$I_{wei}(k) = I_{ori}(k) \circ L_{ref}, k \in [1, c], \tag{2}$$

where W_{lam} and b_{lam} respectively denote the weights and biases of convolutional layers in LAM, symbol $\circ$ denotes element-wise product, $I_{ori}(k)$ is the k^{th} channel of the original image I_{ori}, and I_{wei} is the weighted fundus image produced by LAM. The weighted image I_{wei} then serves as the input of the classification network for diagnosing DR.

2.3 Implementation Details

For the lesion detection model, we initialize it using pre-trained parameters from a model trained on the COCO dataset [11]. Adam optimizer with 0.01 initial learning rate is adopted to train the network. During training, the learning rate decays by multiplying 0.1 when the training loss does not decrease. For LGN, the weights of LAM are randomly initialized and a ResNet50 model pre-trained on ImageNet is used as the initial classification model. Stochastic gradient descent with 0.003 initial learning rate and cosine learning rate decay [12] is used to train the LGN. The batch size is 32 and the maximum number of training epochs is 20.

3 RESULTS AND DISCUSSION

3.1 Evaluation Method and Metrics

To evaluate the performance of LGN, three experiments are conducted:
ResNet50: In this setting, only fundus images were used as the input of the classification network.
ResNet50 + coarse lesion map: To show the superiority of LAM

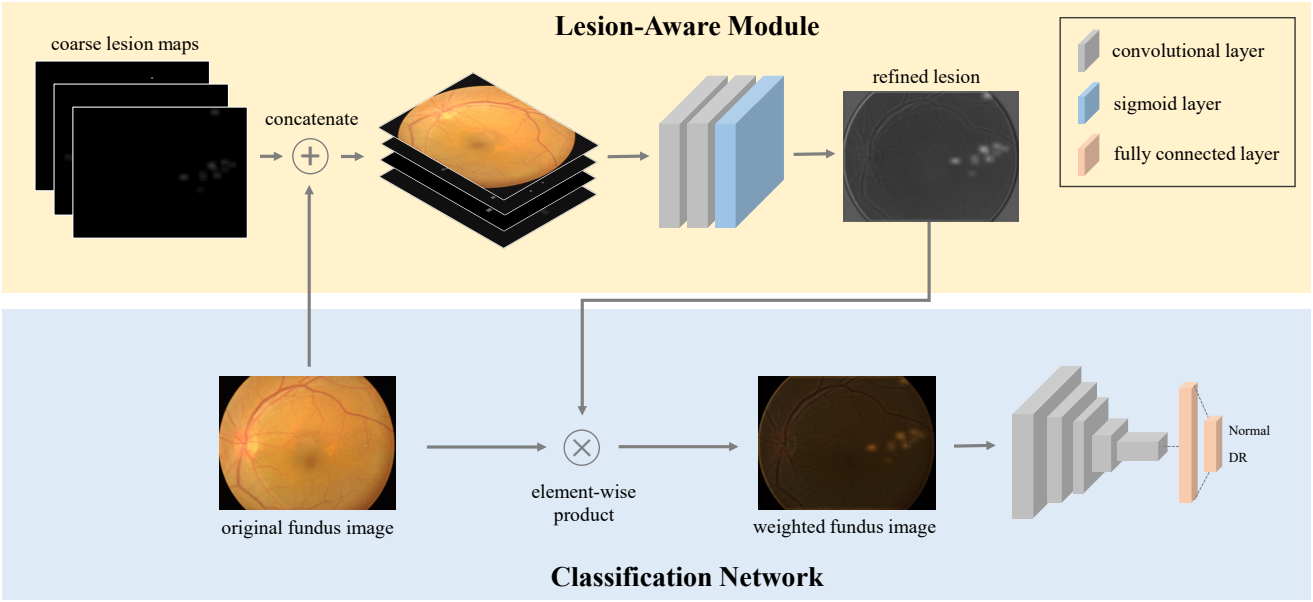

Figure 2: Lesion Guided Network. Lesion-Aware Module (top) takes a concatenation of the fundus image and the corresponding coarse lesion map as input and outputs a refined lesion map. And the weighted image obtained by element-wise production of the fundus image and the refined lesion map is fed into the classification network (bottom) for DR identification.

in producing lesion probability map, a weighted image is generated by directly multiplying the coarse lesion map and the original image, and is then used to identify DR.

LGN (proposed): In this experiment, LAM is applied to refine the coarse lesion map and generate a weighted image for DR identification. ResNet50 is adopted as the classification network, same as the other two settings. As for evaluation metrics, we employ sensitivity (Sen), specificity (Spec), and accuracy (ACC). They are defined as below:

$$Sen = \frac{TP}{TP + FN}, \tag{3}$$

$$Spec = \frac{TN}{TN + FP}, \tag{4}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \tag{5}$$

where TP denotes true positive, (DR is correctly classified), TN denotes true negative, FP denotes false positive, and FN denotes false negative. The performance of different methods is tabulated in Table 1. The results show that directly fusing coarse lesion maps with images does not help the DR identification, while LGN improves accuracy by 1.15% compared over the plain ResNet50.

Some representative examples of with corresponding refined lesion maps and weighted images are shown in Fig. 3, which demonstrate that LGN is capable of reducing noises in lesion maps and identifying easily-overlooked lesions, as well as generating discriminative weighted images for DR identification.

Table 1: Performance comparisons between the proposed method and some other alternatives.

Method	Sen	Spec	Acc
ResNet50	0.8618	**0.9693**	0.9156
ResNet50 + coarse lesion map	0.8645	0.9514	0.9079
LGN (proposed)	**0.9079**	0.9463	**0.9271**

4 CONCLUSION

In this paper, we propose and validate a DR identification pipeline based on RetinaNet and a novel Lesion Guided Network using fundus images. A RetinaNet model is trained on a coarsely-annotated dataset and is then employed to produce coarse lesion maps. The proposed LGN can further refine the lesion maps using Lesion-Aware Module. Then, utilizing lesion information, weighted fundus images are fed into the classification network in LGN for DR identification. Experimental results identify the importance of LGN in terms of refining coarse lesion maps. The effectiveness of the proposed LGN has been successfully validated on a private dataset containing 4465 fundus images. With that being said, LAM still has limitations in processing coarse lesion maps with too many errors. Moreover, although the annotating process using coarse bounding circles is much easier than standard pixel-wise processing, constructing such a dataset with coarse annotations is still a burden for ophthalmologists. Addressing such limitations, plausibly employing weakly-supervised or unsupervised techniques, is one of our future research directions.

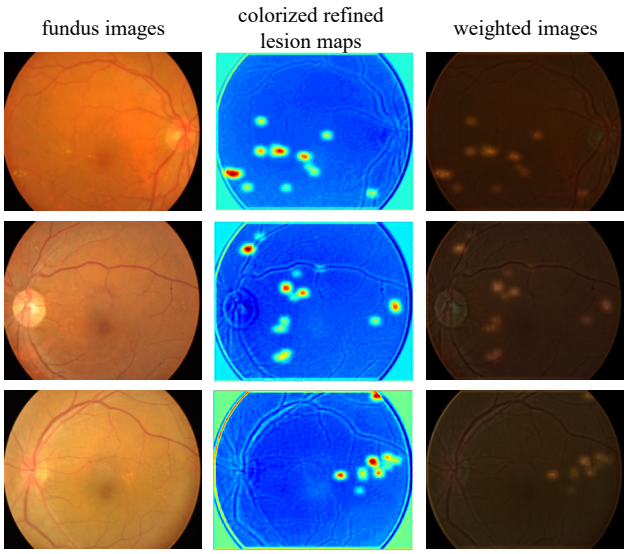

Figure 3: Representative fundus images with the corresponding refined lesion maps and weighted images.

ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (62071210), the Shenzhen Basic Research Program (JCYJ20190809120205578), the National Key R&D Program of China (2017YFC0112404), and the High-level University Fund (G02236002). The authors declare that they have no competing financial interests.

REFERENCES

[1] Cheung Ning, Mitchell Paul, and Yin Wong Tien, "Diabetic retinopathy," *The Lancet*, vol. 376, no. 9735, pp. 124 – 136, 2010.

[2] Harry Pratt, Frans Coenen, Deborah M Broadbent, Simon P Harding, and Yalin Zheng, "Convolutional neural networks for diabetic retinopathy," *Procedia Computer Science*, vol. 90, pp. 200–205, 2016.

[3] Daniel S Kermany, Michael Goldbaum, Wenjia Cai, Carolina CS Valentim, Huiying Liang, Sally L Baxter, Alex McKeown, Ge Yang, Xiaokang Wu, Fangbing Yan, et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[4] Yehui Yang, Tao Li, Wensi Li, Haishan Wu, Wei Fan, and Wensheng Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 533–540.

[5] Yijin Huang, Li Lin, Meng Li, Jiewei Wu, Pujin Cheng, Kai Wang, Jin Yuan, and Xiaoying Tang, "Automated hemorrhage detection from coarsely annotated fundus images in diabetic retinopathy," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2020, pp. 1369–1372.

[6] Prasanna Porwal, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, and Fabrice Meriaudeau, "Indian diabetic retinopathy image dataset (idrid): a database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, pp. 25, 2018.

[7] Li Lin, Meng Li, Yijin Huang, Pujin Cheng, Honghui Xia, Kai Wang, Jin Yuan, and Xiaoying Tang, "The SUSTech-SYSU dataset for automated exudate detection and diabetic retinopathy grading," *Scientific Data*, vol. 7, no. 1, pp. 1–10, 2020.

[8] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[10] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint arXiv:1505.00853*, 2015.

[11] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.

[12] Ilya Loshchilov and Frank Hutter, "Sgdr: stochastic gradient descent with restarts. corr abs/1608.03983 (2016)," *arXiv preprint arXiv:1608.03983*, 2016.