

# Diabetic Retinopathy Lesion Identification and Multiple Instance Learning

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Renyu Li , Yunchao Gu , Xinliang Wang , Junjun Pan

Accurate identification of lesions and their use across different medical institutions are the foundation and key to the clinical application of automatic diabetic retinopathy (DR) detection. Existing detection or segmentation methods can achieve acceptable results in DR lesion identification, but they strongly rely on a large number of fine-grained annotations that are not easily accessible and suffer severe performance degradation in the cross-domain application.

diabetic retinopathy identification

multiple instance learning

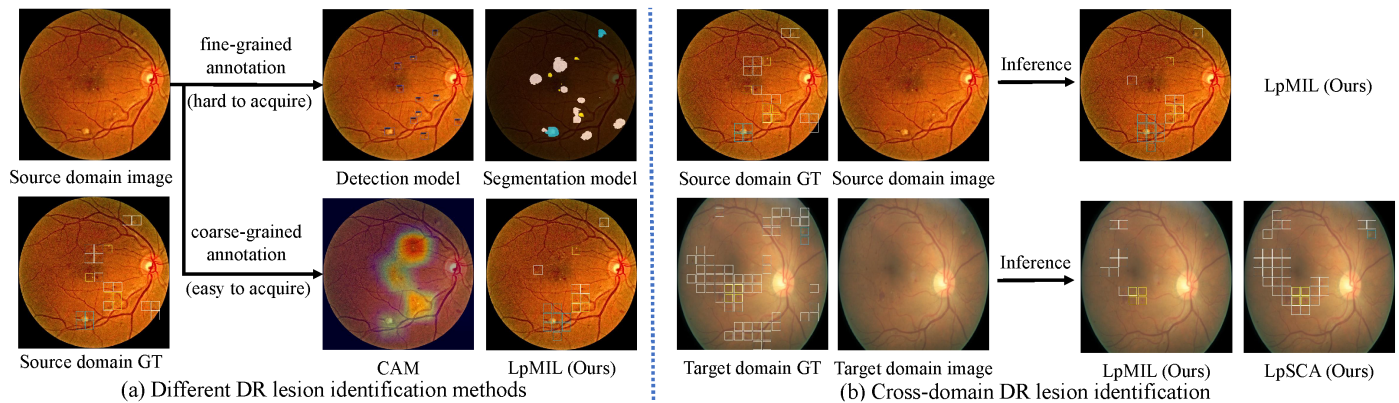
supervised learning

## 1. Introduction

Diabetic retinopathy (DR) is one of the most common complications of diabetes and one of the leading causes of visual impairment in the working-age population. Fortunately, timely diagnosis can prevent further deterioration of the lesions, thus reducing the risk of blindness. During the diagnosis of DR, the ophthalmologist completes the comprehensive diagnosis by identifying the lesion attributes on the fundus image, such as microaneurysm (MA), hemorrhage (HE), exudate (EX), cotton wool spots (CWS), neovascularization (NV), and intraretinal microvascular abnormalities (IRMA). However, due to the difficulty in identifying certain lesions, this process can be time-consuming and labor-intensive. Automatic DR-aided diagnosis methods use deep learning models to extract features from the fundus image to complete the location of the lesions, and the results can be provided to ophthalmologists for further diagnosis. At the same time, with the maturity of automatic DR-assisted diagnosis technology, the requirements for deep learning models in clinical applications are also increasing. For example, it is expected to have the ability to be used across medical institutions. In conclusion, cross-domain localization of DR lesions is becoming a concern of both academia and industry.

In recent years, with the development of deep learning, as shown in **Figure 1a**, several lesion identification models have been proposed to assist ophthalmologists in the diagnosis of DR. Models [\[1\]](#)[\[2\]](#)[\[3\]](#)[\[4\]](#) trained with fine-grained annotations such as pixel-level annotations or bounding box annotations have been proposed and have achieved acceptable results in DR lesion identification. However, the application of these models is limited due to the time-consuming manual annotation. Therefore, some methods [\[5\]](#)[\[6\]](#)[\[7\]](#) attempt to accomplish both DR grading and lesion identification using only coarse-grained annotations such as grading labels or lesion attribute labels. However, due to the limited supervision provided by coarse-grained annotations, these methods tend to be biased on the most important lesion regions while ignoring trivial lesion information. In addition, in clinical applications, image quality and imaging performance vary due to the different image acquisition equipment used in different healthcare

facilities, the direct application of models on other datasets will suffer huge performance losses (**Figure 1b**), which greatly limits the flexibility and scalability of these deep learning methods.



**Figure 1.** (a) Different DR lesion identification methods, including models trained with fine-grained annotations represented by detection models and segmentation models, and models trained with coarse-grained annotations represented by CAM and LpMIL. Using only coarse-grained annotations, the LpMIL not only achieves better lesion identification performance than CAM, but also achieves results that are competitive with detection and segmentation models. (b) Directly applying the LpMIL trained on the source domain to the target domain results in severe performance degradation, while the LpSCA improves cross-domain lesion identification performance through the semantic constrained adaptation method. “GT” denotes ground truth.

## 2. Related Work

### 2.1. Diabetic Retinopathy Lesion Identification

To complete automatic DR lesion identification, many DR lesion identification methods based on pixel-level or bounding box annotations have been proposed. Yang et al. [8] propose a two-stage Convolutional Neural Network (CNN) for DR grading and lesion detection, which uses the lesion detection results to assign different weights to the image patch to improve the performance of DR grading. Li et al. [9] adopt the object detection model to extract lesion features from fundus images for DR grading. Using a small number of pixel-level annotations, Foo et al. [10] propose a multi-task learning approach to simultaneously complete the tasks of DR grading and lesion segmentation. Zhou et al. [2] propose the FGADR dataset, on which DR lesion segmentation is performed. However, the application of these methods is limited due to the difficulty of obtaining fine-grained annotations. Therefore, researchers attempt to accomplish both DR grading and lesion identification using only coarse-grained grading labels. Wang et al. [5] utilize the attention map to highlight suspicious areas, and complete DR grading and lesion localization at the same time. Sun et al. [6] formulate lesion identification as a weakly supervised lesion localization problem through a transformer decoder, which jointly performs DR grading and lesion detection. Different from previous methods, researchers define DR lesion identification as a multi-label classification problem for patch-level lesion identification and use a cross-domain approach to enable the model to be used across healthcare institutions.

## 2.2. Multiple Instance Learning

Multiple instance learning has become a widely adopted weakly supervised learning method [11][12][13][14][15][16]. Many works have studied different pooling functions combined with instance embedding or instance prediction to accomplish bag-level prediction [17][18][19][20]. However, the pooling function itself cannot provide sufficient information, and supervision can only be retained at the bag level, which severely limits the effect of instance-level prediction. To address this problem, some approaches introduce artificial instance labels by specifying thresholds to provide both bag-level and instance-level supervision. Zhou et al. [18] use specified thresholds to directly assign positive or negative labels based on prediction scores, providing supervision for all instances. Morfi et al. [21] propose MMM loss for audio event detection. This loss function provides supervision for instances of extreme prediction scores according to the specified threshold and obtains bag level prediction through average pooling aggregation for bag level supervision. Seibold et al. [22] use a more customized way to create soft instance labels, flexibly providing supervision for all instances, and apply it to the pathological localization of chest radiographs pathologies.

To the best of our knowledge, these methods have not been applied to automatic DR detection.

## 2.3. Domain Adaptation

Domain adaptation is a subtask of transfer learning, which maps features of different domains to the same feature space, and utilizes the labels of the source domain to enhance the training of the target domain. The mainstream approach is to learn the domain invariant representation using adversarial training. DANN [23] pioneers this field by training a domain discriminator to distinguish the source domain from the target domain, and training a feature extractor to cheat the discriminator to align the features of the two domains. CDAN [24] utilizes the discrimination information predicted by the classifier to condition the adversarial model. GVB [25] improves adversarial training by building bridging layers between the generator and the discriminator. MetaAlign [26] treats domain alignment tasks and classification tasks as meta-training and meta-testing tasks in a meta-learning scheme for domain adaptation.

## References

1. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Conference on Neural Information Processing Systems (NeurIPS 2015), Montreal, QC, Canada, 7–10 December 2015; Volume 28.
2. Zhou, Y.; Wang, B.; Huang, L.; Cui, S.; Shao, L. A benchmark for studying diabetic retinopathy: Segmentation, grading, and transferability. *IEEE Trans. Med. Imaging* 2020, 40, 818–828.
3. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.

4. Zhou, Y.; He, X.; Huang, L.; Liu, L.; Zhu, F.; Cui, S.; Shao, L. Collaborative learning of semi-supervised segmentation and classification for medical images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2079–2088.
5. Wang, Z.; Yin, Y.; Shi, J.; Fang, W.; Li, H.; Wang, X. Zoom-in-net: Deep mining lesions for diabetic retinopathy detection. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Quebec City, QC, Canada, 11–13 September 2017; pp. 267–275.
6. Sun, R.; Li, Y.; Zhang, T.; Mao, Z.; Wu, F.; Zhang, Y. Lesion-aware transformers for diabetic retinopathy grading. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 10938–10947.
7. Wang, X.; Gu, Y.; Pan, J.; Jia, L. Diabetic Retinopathy Detection Based on Weakly Supervised Object Localization and Knowledge Driven Attribute Mining. In Proceedings of the International Workshop on Ophthalmic Medical Image Analysis, Strasbourg, France, 27 September 2021; pp. 32–41.
8. Yang, Y.; Li, T.; Li, W.; Wu, H.; Fan, W.; Zhang, W. Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Quebec City, QC, Canada, 11–13 September 2017; pp. 533–540.
9. Lin, Z.; Guo, R.; Wang, Y.; Wu, B.; Chen, T.; Wang, W.; Chen, D.Z.; Wu, J. A framework for identifying diabetic retinopathy based on anti-noise detection and attention-based fusion. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018; pp. 74–82.
10. Foo, A.; Hsu, W.; Lee, M.L.; Lim, G.; Wong, T.Y. Multi-task learning for diabetic retinopathy grading and lesion segmentation. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 13267–13272.
11. Sudharshan, P.; Petitjean, C.; Spanhol, F.; Oliveira, L.E.; Heutte, L.; Honeine, P. Multiple instance learning for histopathological breast cancer image classification. *Expert Syst. Appl.* **2019**, *117*, 103–111.
12. Lerousseau, M.; Vakalopoulou, M.; Classe, M.; Adam, J.; Battistella, E.; Carré, A.; Estienne, T.; Henry, T.; Deutsch, E.; Paragios, N. Weakly supervised multiple instance learning histopathological tumor segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, 4–8 October 2020; pp. 470–479.
13. Chikontwe, P.; Kim, M.; Nam, S.J.; Go, H.; Park, S.H. Multiple instance learning with center embeddings for histopathology classification. In Proceedings of the Medical Image Computing

- and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, 4–8 October 2020; pp. 519–528.
14. Li, H.; Yang, F.; Zhao, Y.; Xing, X.; Zhang, J.; Gao, M.; Huang, J.; Wang, L.; Yao, J. DT-MIL: Deformable transformer for multi-instance learning on histopathological image. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, 27 September–1 October 2021; pp. 206–216.
  15. Zhang, H.; Meng, Y.; Zhao, Y.; Qiao, Y.; Yang, X.; Coupland, S.E.; Zheng, Y. DTFD-MIL: Double-tier feature distillation multiple instance learning for histopathology whole slide image classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 18802–18812.
  16. Qian, Z.; Li, K.; Lai, M.; Chang, E.I.C.; Wei, B.; Fan, Y.; Xu, Y. Transformer based multiple instance learning for weakly supervised histopathology image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Singapore, 18–22 September 2022; pp. 160–170.
  17. Li, B.; Li, Y.; Eliceiri, K.W. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 14318–14328.
  18. Zhou, Y.; Sun, X.; Liu, D.; Zha, Z.; Zeng, W. Adaptive pooling in multi-instance learning for web video annotation. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 318–327.
  19. Wang, Y.; Li, J.; Metze, F. A comparison of five multiple instance learning pooling functions for sound event detection with weak labeling. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 31–35.
  20. Ilse, M.; Tomczak, J.; Welling, M. Attention-based deep multiple instance learning. In Proceedings of the International Conference on Machine Learning, PMLR, Stockholm, Sweden, 10–15 July 2018; pp. 2127–2136.
  21. Morfi, V.; Stowell, D. Data-efficient weakly supervised learning for low-resource audio event detection using deep learning. arXiv 2018, arXiv:1807.06972.
  22. Seibold, C.; Kleesiek, J.; Schlemmer, H.P.; Stiefelhagen, R. Self-Guided Multiple Instance Learning for Weakly Supervised Thoracic Disease Classification and Localization in Chest Radiographs. In Proceedings of the ACCV, Kyoto, Japan, 30 November–4 December 2020.
  23. Ganin, Y.; Lempitsky, V. Unsupervised domain adaptation by backpropagation. In Proceedings of the International Conference on Machine Learning, PMLR, Lille, France, 6–11 July 2015; pp.

1180–1189.

24. Long, M.; Cao, Z.; Wang, J.; Jordan, M.I. Conditional adversarial domain adaptation. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montreal, QC, Canada, 3–8 December 2018; Volume 31.
25. Cui, S.; Wang, S.; Zhuo, J.; Su, C.; Huang, Q.; Tian, Q. Gradually vanishing bridge for adversarial domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 12455–12464.
26. Wei, G.; Lan, C.; Zeng, W.; Chen, Z. Metaalign: Coordinating domain alignment and classification for unsupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 16643–16653.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/113282>