Article

A Weakly Supervised Deep Learning Approach for Leakage Detection in Fluorescein Angiography Images

Wanyue Li[1,2], Wangyi Fang[3,4], Jing Wang[1,2], Yi He[1,2], Guohua Deng[5], Hong Ye[2], Zujun Hou[2], Yiwei Chen[2], Chunhui Jiang[3,4], and Guohua Shi[1,2,6]

[1] School of Biomedical Engineering (Suzhou), Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, People's Republic of China
[2] Jiangsu Key Laboratory of Medical Optics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, People's Republic of China
[3] Department of Ophthalmology and Vision Science, Eye and ENT Hospital, Fudan University, Shanghai, People's Republic of China
[4] Key Laboratory of Myopia of State Health Ministry, and Key Laboratory of Visual Impairment and Restoration of Shanghai, Shanghai, People's Republic of China
[5] Department of Ophthalmology, the Third People's Hospital of Changzhou, Changzhou, People's Republic of China
[6] Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, People's Republic of China

Correspondence: Yi He, Jiangsu Key Laboratory of Medical Optics, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, 88 Kelin Road, Suzhou 215163, People's Republic of China. e-mail: heyi_job@126.com

Purpose: The purpose of this study was to design an automated algorithm that can detect fluorescence leakage accurately and quickly without the use of a large amount of labeled data.

Methods: A weakly supervised learning-based method was proposed to detect fluorescein leakage without the need for manual annotation of leakage areas. To enhance the representation of the network, a residual attention module (RAM) was designed as the core component of the proposed generator. Moreover, class activation maps (CAMs) were used to define a novel anomaly mask loss to facilitate more accurate learning of leakage areas. In addition, sensitivity, specificity, accuracy, area under the curve (AUC), and dice coefficient (DC) were used to evaluate the performance of the methods.

Results: The proposed method reached a sensitivity of 0.73 ± 0.04, a specificity of 0.97 ± 0.03, an accuracy of 0.95 ± 0.05, an AUC of 0.86 ± 0.04, and a DC of 0.87 ± 0.01 on the HRA data set; a sensitivity of 0.91 ± 0.02, a specificity of 0.97 ± 0.02, an accuracy of 0.96 ± 0.03, an AUC of 0.94 ± 0.02, and a DC of 0.85 ± 0.03 on Zhao's publicly available data set; and a sensitivity of 0.71 ± 0.04, a specificity of 0.99 ± 0.06, an accuracy of 0.87 ± 0.06, an AUC of 0.85 ± 0.02, and a DC of 0.78 ± 0.04 on Rabbani's publicly available data set.

Conclusions: The experimental results showed that the proposed method achieves better performance on fluorescence leakage detection and can detect one image within 1 second and thus has great potential value for clinical diagnosis and treatment of retina-related diseases, such as diabetic retinopathy and malarial retinopathy.

Translational Relevance: The proposed weakly supervised learning-based method that automates the detection of fluorescence leakage can facilitate the assessment of retinal-related diseases.

Introduction

Fundus fluorescein angiography (FA) can reflect the damaged state of the retinal barrier in living human eyes and is the standard screening and diagnosis technique for retinal diseases.[1] Identifi-cation of high-intensity retinal leakage in FA images is a crucial step for clinicians to develop therapy planning and monitor treatment outcomes. However, current practical approaches for fluorescein leakage detection are usually labeled by trained graders[2] and require laborious and time-consuming work that is inevitably influenced by human factors. Thus, an

effective automated fluorescein leakage detection method is urgently needed.

Algorithms that tackle automated fluorescein leakage detection tasks can be mainly divided into two types: intensity-based methods and learning-based methods. The traditional intensity-based methods usually detect high-intensity leakages by analyzing the pixel intensity variation.[2–5] Although this kind of method has achieved relatively higher sensitivity and specificity in fluorescein leakage detection, it takes a long time to process an image (more than 20 seconds). Recently, with the development of machine learning and deep learning techniques, leaning-based methods are also applied to the fluorescein leakage detection task. Trucco et al.[6] and Tsai et al.[7] applied AdaBoost to detect the leakage regions of FA images, and Béouche-Hélias et al.[1] used random decision forests for leakage detection. However, these methods are all supervised and require a large amount of training data derived from manual annotation, which makes their performances inherently dependent on the quality of the annotations. To solve this problem, Li et al.[8] proposed an unsupervised learning-based fluorescence leakage detection method. However, this method cannot focus solely on the learning of leakage areas and leads to false leakage detection.

The main purpose of this work is to design an automated algorithm that can detect the fluorescence leakage more accurately and quickly than those in former reports and without the use of a large amount of labeled data.

Methods

Data Sets

The data set used in this study contains images with Spectralis HRA equipment (Heidelberg Engineering, Heidelberg, Germany) between March 2011 and September 2019 at the Third People's Hospital of Changzhou (Jiangsu, China); we call this data set the "HRA data set" in this work. The image types in our data set are normal FA images and abnormal FA images with three kinds of typical fluorescein leakage in retinal diseases, that is, optic disc leakage, large focal leakage, and punctate focal leakage (Fig. 1). These kinds of leakage have the same characteristic: the leakage of early angiography usually does not appear or is not obvious, but its size and brightness will increase in the late phase. The data sets initially contained 509 abnormal FA images and 343 normal FA images captured in the late phase (5–6 minutes) of angiography from 852 eyes of 462 patients (223 female, 239 male, ranging in age from 7 to 86 years) and one picture per eye. The resolution of each image is 768 × 768 pixels, and the field of views of these images includes 30°, 45°, and 60°. Twenty percent of the normal and abnormal FA images were randomly selected to comprise the testing set, and the remaining images were employed as the training set. Training data were augmented with random horizontal flips and rotations operations, leading to a final 1709 abnormal and 1149 normal FA images.

Thirty-two abnormal FA images with optic disc and large focal leakage in the test set were labeled by two specialists who demonstrated good intra- and interobserver consistency (Supplementary Table S1). The Visual Geometry Group Image Annotator,[9] an open-source tool, was used for annotations. Before commencing to grade the test data set, manual graders discussed and agreed upon the leakage definition and segmentation protocol. To define intraobserver reliability, one manual grader repeated his grading on the same images at least 5 weeks after the initial grading.

Anomaly Mask Calculation

Class activation maps (CAMs) have been widely used in many tasks, such as object localization,[10,11]
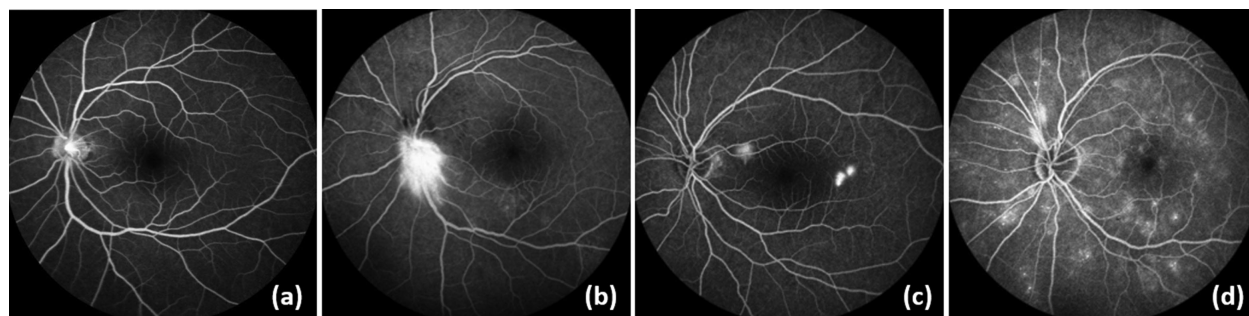

Figure 1. Examples of the types of FA images. (a) Normal FA image. FA image with (b) optic disc leakage, (c) large focal leakage, and (d) punctate focal leakage.
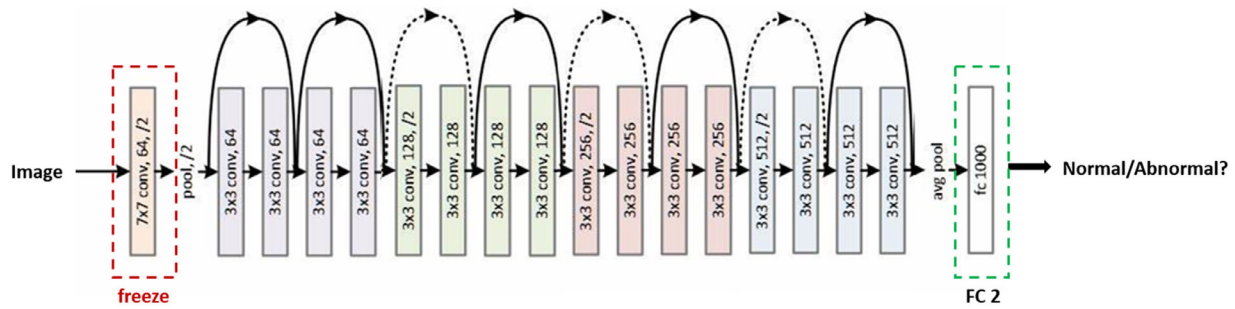
Figure 2. Architecture of the normal and abnormal FA image classification network. conv, convolutional layer; FC, fully connected layer.

image segmentation,[12,13] and unpaired image translation.[14] In this work, we mainly used CAMs to roughly localize the leakage areas and generate a binary anomaly mask of abnormal FA images. The anomaly mask can provide weak constraints for the network training process.

The calculating process of the anomaly mask can be described as three steps. First, a classification network is trained to classify the normal and abnormal FA images. We adopt ResNet18[15] pretrained on the ImageNet data set[16] as the backbone network (Fig. 2). Since the FA images and the images in the ImageNet data set are different in nature, we freeze the low-level layers (i.e., the first layer of ResNet18, highlighted by the red dotted box in Fig. 2) and redesign the last layer of the pretrained network for our binary classification task (highlighted by the green dotted box in Fig. 2), and the remaining architectures are all retrained based on the weights pretrained on "ImageNet." This network achieves an accuracy of 0.993 on the normal and abnormal FA image classification. Second, the gradient-weighted CAM method[17] is applied to generate the CAM of abnormal FA images. Finally, Otsu's binarization method[18] is used to generate the binary mask of abnormal FA images. To ensure that the leakage areas can be included in the white areas of the mask as much as possible, we set the threshold to 0.6 times Otsu's original threshold for all images. Some examples of the generated CAM and mask of abnormal FA images are shown in Figure 3.

Normal-Looking FA Image Generation and Leakage Detection

The main idea of the proposed method is similar to the study proposed by Li et al.,[8] that is, to train a model
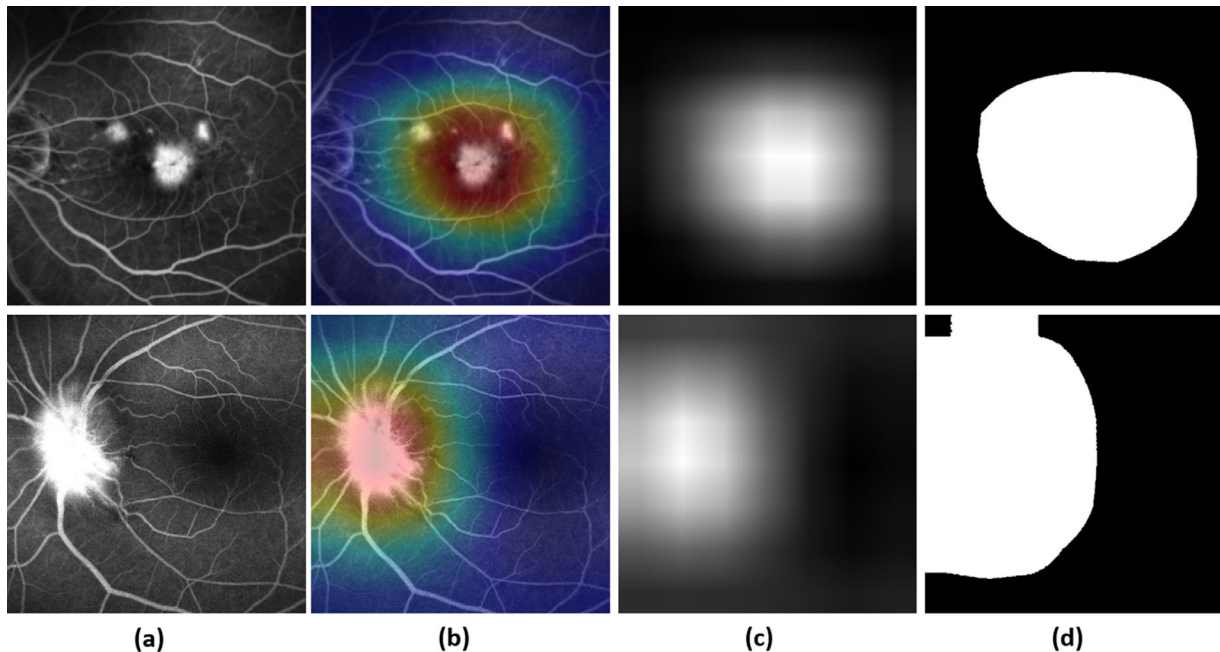

Figure 3. Examples of the generated CAM and mask of abnormal FA images. (a) Original FA images. (b) Corresponding CAMs (represented as heatmaps on the FA images). (c) Corresponding CAMs (grayscale). (d) Corresponding anomaly masks.
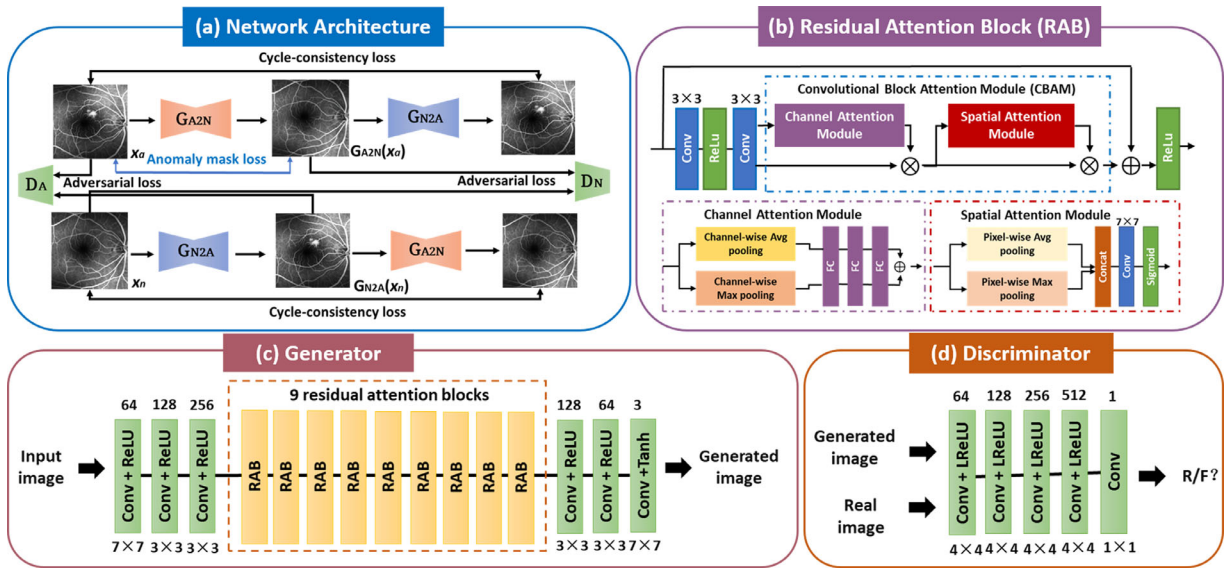
Figure 4. Architecture of the proposed method. (a) The proposed CycleGAN-based network consists of two generators, G_{A2N} and G_{N2A}, and two discriminators, D_A and D_N. Architecture of (b) RAB, (c) generator G_{A2N} and G_{N2A}, and (d) discriminator D_A and D_N. LReLU, leaky rectified linear unit; ReLU, rectified linear unit; Tanh, TanHyperbolic function.

that can generate a normal-looking FA image from the input abnormal image and then detect the leakage by calculating the difference between the abnormal and generated normal images.

To transfer the abnormal FA image into the normal domain without the use of paired images, a cyclic generative adversarial network (CycleGAN)–based network is designed. As shown in Figure 4a, this network also consists of two generators and two discriminators, which makes the network trained in an adversarial way. Generator G_{A2N} takes an image x_a as input and generates a normal-looking image $G_{A2N}(x_a)$; the main goal of generator G_{N2A} is to translate an image from a normal domain to an abnormal domain. Two discriminators, D_A and D_N, aim to discriminate between real and generated images. Figure 4c shows the architecture of the proposed generators, which contain three key components: an encoder (composed of a 7×7 convolutional layer and two 3×3 convolutional layers), nine residual attention blocks (RABs), and a decoder. The discriminators are the same as CycleGAN's (i.e., a 70×70 PatchGAN), as shown in Figure 4d.

Residual Attention Block

The main aim of the designed RAB was to increase representation ability by using an attention mechanism: focusing on important features and suppressing unnecessary ones. Details of the RAB are illustrated in Figure 4b. The RAB is mainly composed of two 3 $\times$ 3 convolutional layers and a skip connection, and

a convolutional block attention module (CBAM)[19] is concatenated following the second convolutional layer. CBAM is a plug-and-play module to learn "what" and "where" to focus on the channel and spatial dimensions, respectively. As shown in Figure 4b, this module consists of two submodules: a channel attention module and a spatial attention module. Thus, CBAM can enhance meaningful features along both channel- and spatial-wise dimensions.

All experiments were implemented on an Ubuntu 16.04 + Python 3.6 + PyTorch 1.7.0 environment. The proposed model was trained for 200 epochs using linear decay with a batch size of 1 and the Adam optimizer. It took nearly 72 hours on one GeForce GTX 1080Ti GPU to train the model.

Loss Function

To ensure that the network generates more realistic normal domain images, we formulate the loss function as a combination of adversarial loss L_{GAN}, cycle consistency loss L_{CC}, and anomaly mask loss L_{AM}. The full loss function of this network can be written as follows:

$$\text{L} = \lambda_{GAN} L_{GAN} + \lambda_{CC} L_{CC} + \lambda_{AM} L_{AM}, \quad (1)$$

where λ_{GAN}, λ_{CC}, and λ_{AM} are the experimentally determined hyperparameters that control the effect of adversarial loss, cycle consistency loss, and anomaly mask loss, respectively. Pursuing balance among the three losses is not a trivial task. After multiple experiments, we set $\lambda_{GAN} = 1$, $\lambda_{CC} = 10$, and $\lambda_{AM} = 10$ in this task.

Adversarial Loss

The proposed network adopts a bidirectional transform model with two generators, G_{A2N} and G_{N2A}, trained simultaneously. This strategy can help stabilize the model training. Since we have two generators and discriminators, the GAN loss can be defined as

$$L_{GAN} = \mathrm{E}_{p_a}[\log D_A(x_a)] + \mathrm{E}_{p_a}[\log(1 - D_N(G_{A2N}(x_a)))]$$
$$+ \mathrm{E}_{p_n}[\log D_N(x_n)] + \mathrm{E}_{p_n}[\log(1 - D_A(G_{N2A}(x_n)))]. \quad (2)$$

Cycle Consistency Loss

The cycle consistency loss is adopted to transform normal and abnormal FA into one another and aid the learning of G_{A2N} and G_{N2A}:

$$L_{CC} = \mathrm{E}_{p_a}[\|G_{N2A}(G_{A2N}(x_a)) - x_a\|_1]$$
$$+ \mathrm{E}_{p_n}[\|G_{A2N}(G_{N2A}(x_n)) - x_n\|_1]. \quad (3)$$

Anomaly Mask Loss

The anomaly mask (AM) loss is proposed to assume that each abnormal FA image has a corresponding binary mask that indicates where the leakage areas are within the image. The anomaly mask M_x of the abnormal FA image can be calculated using the method described earlier, and this mask is not needed during testing. Since we want generator G_{A2N} to automatically isolate and modify the leakage areas within the image without changing nonleakage areas, the loss function is defined as

$$L_{AM} = E_{p_a}\left[\|(\mathbf{1} - M_x) \odot (G_{A2N}(x_a) - x_a)\|_2^2\right], \quad (4)$$

where $\odot$ represents element-wise multiplication and $\mathbf{1}$ is an all-ones matrix of the same size of the anomaly mask. That is to say, if the generator modifies the pixels in an abnormal image x_a that does not correspond to the leakage areas, an L2 cost is paid.

Evaluation Metrics

For a fair comparison with other state-of-the-art methods, the testing results were all evaluated with the criteria of sensitivity (Sen), specificity (Spe), accuracy (Acc), area under the curve (AUC), and dice coefficient (DC), as in the existing leakage detection works.[2,4,8] These metrics can be calculated as

$$Sen = \frac{TP}{TP + FN} \quad (5)$$

$$Spe = \frac{TN}{TN + FP} \quad (6)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$AUC = \frac{Sen + Spe}{2} \quad (8)$$

$$DC = \frac{2(|A \cap B|)}{|A| + |B|}, \quad (9)$$

where *TP*, *TN*, *FP*, and *FN* represent the number of true positives (correctly identified leakage pixels or regions), true negatives (correctly identified background pixels or regions), false positives (incorrectly identified leakage pixels or regions), and false negatives (incorrectly identified background pixels or regions), respectively. *A* indicates the ground truth regions, *B* indicates the segmented regions, and |*A*∩*B*| denotes the number of pixels in the intersecting region between *A* and *B*. All pixels are treated equally in their counting without considering the severity of the symptoms they depict.

Results

Results on the HRA Data Set

In this section, we compared the proposed method with the one by Li et al.,[8] which is also the baseline model of the proposed method. Starting from this baseline model (CycleGAN), which consists of generator G_{A2N} without AM loss and CBAM, we also conducted an ablation study to validate the inclusion of AM loss and CBAM.

Results on Optic Disc and Large Focal Leakage Detection

Table 1 shows that the proposed method (Cycle-GAN + CBAM + L_{AM}) achieves the best overall performance when compared with the other three models and reaches the highest specificity, accuracy, AUC, and DC of 0.97 ± 0.03, 0.95 ± 0.05, 0.86 ± 0.04, and 0.87 ± 0.01, respectively. The qualitative comparison also illustrates the better performance of the proposed method on optic disc and large focal leakage detection (Figs. 5, 6).

Results on Optic Punctate Focal Leakage Detection

Since only the FA images with optic disc and large focal leakages were labeled by experts, only a qualitative analysis is illustrated for punctate leakage detection (Fig. 7). As seen in Figures 7a3–e3, the proposed

Table 1. Performances of Different Methods on Detecting Optic Disc and Large Focal Leakages on the HRA Data Set (32 Images) at the Pixel Level

Characteristic	CycleGAN	CycleGAN + L_{AM}	CycleGAN + CBAM	Proposed Method
Sen	**0.79 ± 0.05**	0.77 ± 0.04	0.74 ± 0.04	0.73 ± 0.04
Spe	0.93 ± 0.06	0.94 ± 0.05	0.96 ± 0.04	**0.97 ± 0.03**
Acc	0.93 ± 0.06	0.91 ± 0.05	0.93 ± 0.05	**0.95 ± 0.05**
AUC	0.85 ± 0.04	0.85 ± 0.05	0.84 ± 0.06	**0.86 ± 0.04**
DC	0.66 ± 0.01	0.66 ± 0.01	0.73 ± 0.01	**0.87 ± 0.01**

Data are expressed as mean ± standard deviation. The bold values represent the best results.
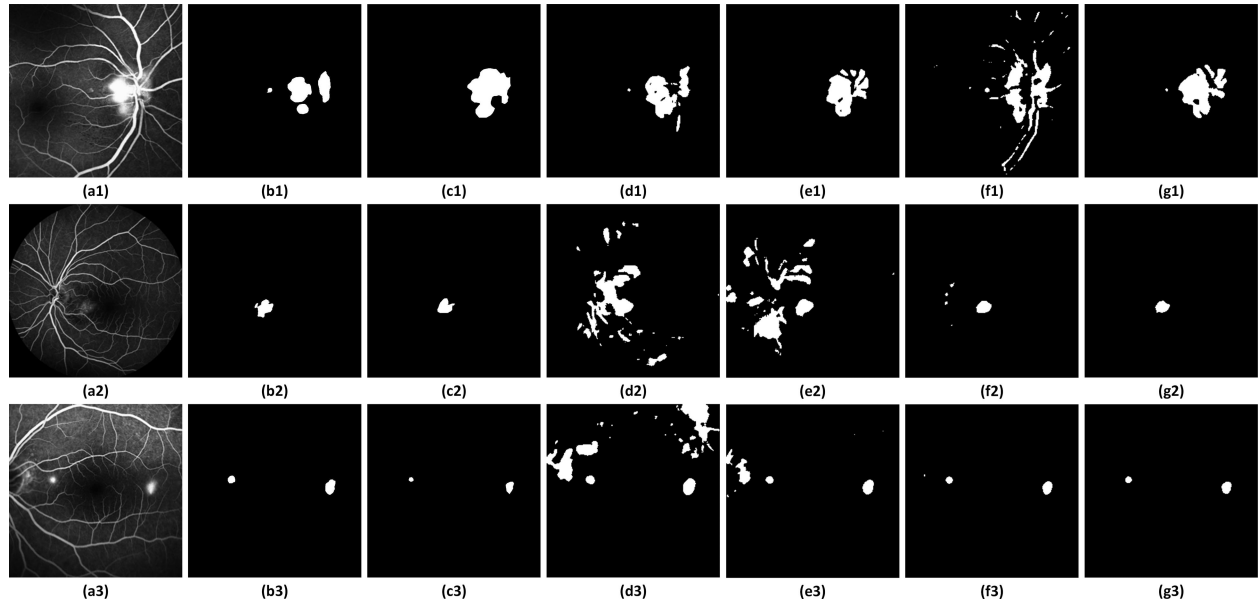

Figure 5. Leakage detection results on HRA data set. (a1–a3) Example of abnormal FA images. Leakage detected by (b1–b3) expert 1's annotation, (c1–c3) expert 2's annotation, (d1–d3) CycleGAN, (e1–e3) CycleGAN + L_{AM}, (f1–f3) CycleGAN + CBAM, and (g1–g3) the proposed method.
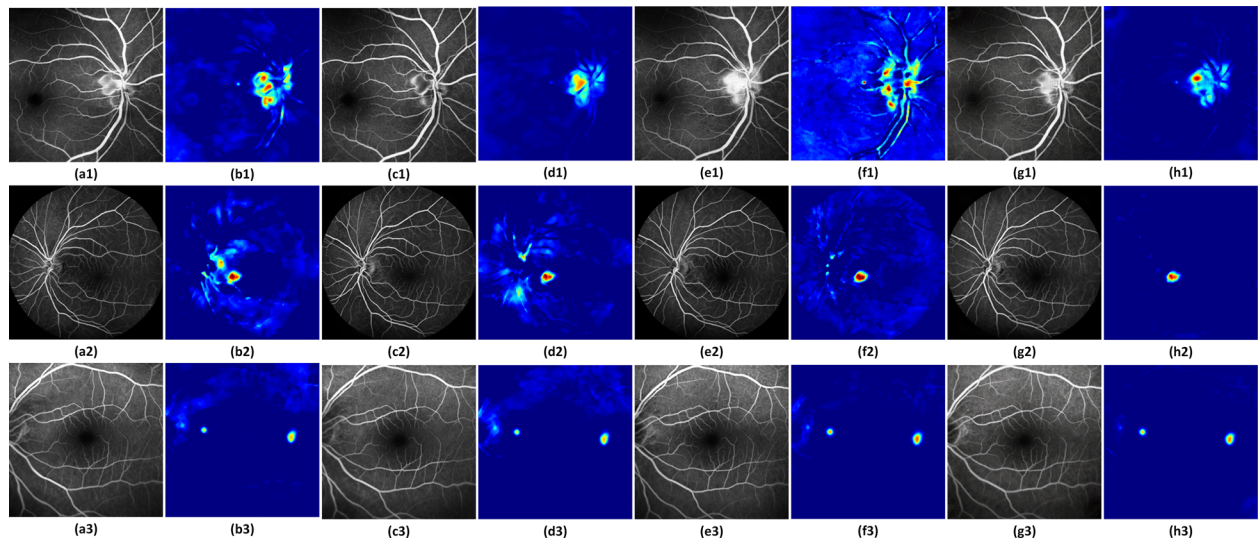

Figure 6. Generated normal-looking image corresponding to the abnormal FA images in Figure 4. Normal-looking FA image generated by (a1–a3) CycleGAN, (c1–c3) CycleGAN + L_{AM}, (e1–e3) CycleGAN + CBAM, and (g1–g3) the proposed method. Difference image between abnormal FA image and normal image generated by (b1–b3) CycleGAN, (d1–d3) CycleGAN + L_{AM}, (f1–f3) CycleGAN + CBAM, and (h1–h3) the proposed method.

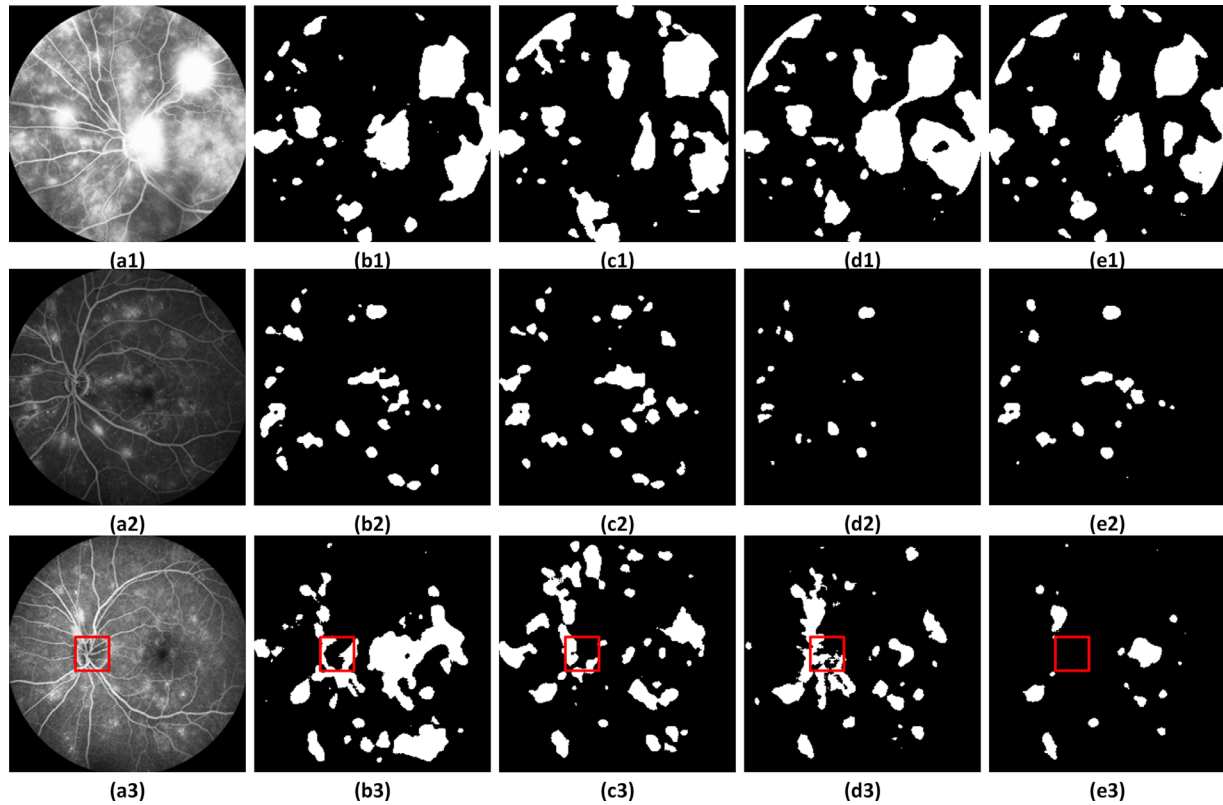translational vision science & technology

Figure 7. Punctate focal leakage detection results on HRA data set. (a1–a3) Example of abnormal FA images. Leakage detected by (b1–b3) CycleGAN, (c1–c3) CycleGAN $+ L_{AM}$, (d1–d3) CycleGAN $+$ CBAM, and (e1–e3) the proposed method.
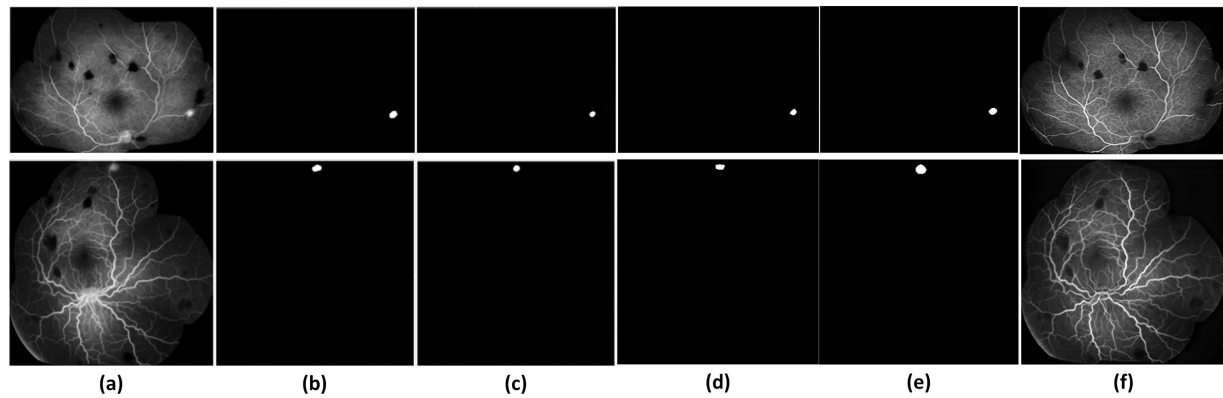

Figure 8. Leakage detection results on publicly available data set. (a) Example of abnormal FA images. Leakage identified by (b) an expert, (c) Zhao et al.,[2] (d) Li et al.[8] (CycleGAN), and (e) the proposed method. (f) The normal-looking FA image generated by the proposed method.

method focuses more on the identification of the real leakage areas, which illustrates its good specificity.

Results on the Publicly Available Data Sets

The proposed model was also tested on two publicly available data sets (data sets from Zhao et al.[2] and Rabbani et al.[4]) and compared with the methods by Zhao et al.,[2] Rabbani et al.,[4] and Li et al.[8] Zhao et al.'s data set contains 30 abnormal FA images (20 large

focal and 10 punctate focal) with signs of malarial retinopathy (MR) on admission. Figure 8 shows that the leakage detection results of the proposed method are closer to the expert's annotation, and Figure 9 illustrates the better performance of the proposed method on leakage detection when compared with Li et al.'s method. Table 2 shows the good quantitative results of the proposed method, with the highest accuracy and DC of 0.96 ± 0.03 and 0.85 ± 0.03, respectively. Rabbani et al.'s data sets contain 24 images (10 predom-
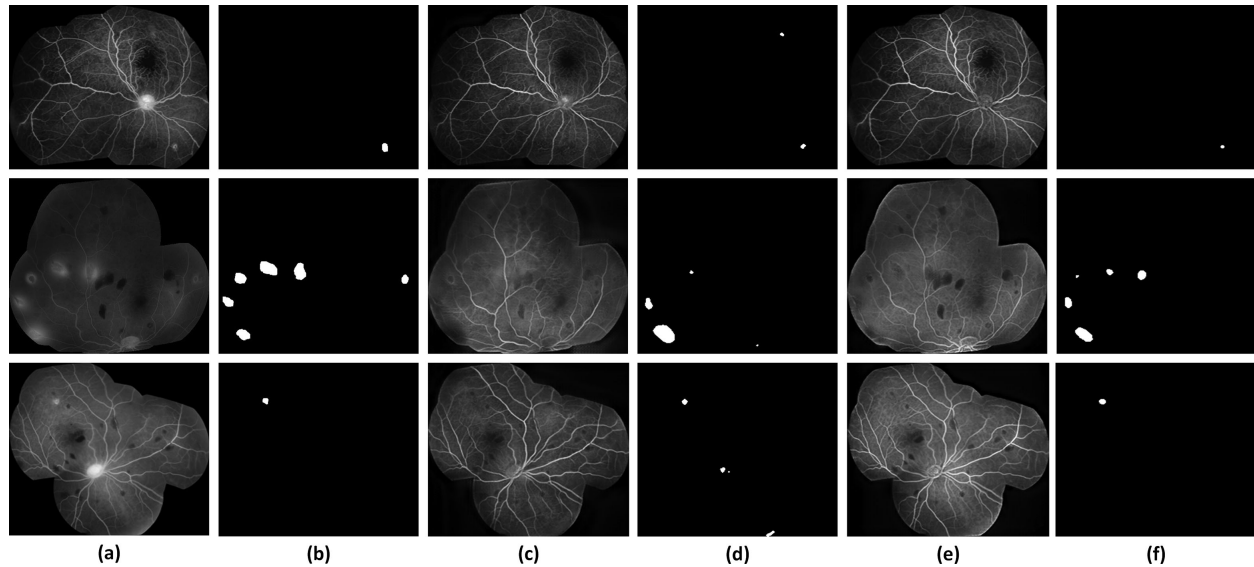
Figure 9. Leakage detection results on the publicly avaliable data set. (a) Example FA image. Leakage identified by (b) an expert, (d) Li et al.[8] (CycleGAN), and (f) the proposed method. The normal-looking FA image generated by (c) Li et al.[8] (CycleGAN) and (e) the proposed method.

Table 2. Performances of Different Methods on Detecting Focal Leakages over the Data Set by Zhao et al.[2] (20 images) at the Pixel Level

Characteristic	Intra obs. (Intraobserver)	Inter obs. (Interobserver)	Method by Rabbani et al.[4]	Method by Zhao et al.[2]	Method by Li et al.[8]	Proposed Method
Sen	0.96 ± 0.02	0.91 ± 0.04	0.81 ± 0.08	**0.93 ± 0.03**	0.92 ± 0.04	0.91 ± 0.02
Spe	0.97 ± 0.03	0.94 ± 0.05	0.87 ± 0.08	0.96 ± 0.02	**0.97 ± 0.02**	**0.97 ± 0.02**
Acc	0.96 ± 0.03	0.89 ± 0.04	0.83 ± 0.10	0.91 ± 0.03	0.92 ± 0.02	**0.96 ± 0.03**
AUC	0.96 ± 0.02	0.92 ± 0.04	0.84 ± 0.08	**0.94 ± 0.02**	0.94 ± 0.03	**0.94 ± 0.02**
DC	0.92 ± 0.04	0.80 ± 0.05	0.74 ± 0.05	0.82 ± 0.03	0.84 ± 0.03	**0.85 ± 0.03**

Data are expressed as mean ± standard deviation. The bold values represent the best results.

inantly focal, 7 predominantly diffuse, 7 mixed pattern leakage) captured from 24 patients who had signs of diabetic retinopathy (DR) on admission. As described in Rabbani et al.,[4] quantitative analysis of a circular region centered at the fovea with a radius of 1500 μm is of the greatest significance for clinical diagnosis and treatment. To make a fair comparison, we also limited the proposed method of detecting the leakages in this area. The quantitative and qualitative results of the proposed method are compared in Figure 10 and Table 3. The proposed method reached the highest specificity of 0.99 ± 0.06.

Discussion

Fundus FA is a valuable imaging technique that provides a map of retinal vascular structure and function by highlighting blockage of, and leakage from, retinal vessels. Detecting and evaluating the high-intensity leakages is a crucial step for disease recognition and treatment.

Current automated leakage identification methods can be classified as intensity based and learning based. The pixel intensity–based methods[2–5] usually require a long time to detect leakages in an image (more than 20 seconds). Also, the supervised-learning based methods[1,6,7] require a large amount of annotated data to train the model. In addition, the existing supervised learning–based methods are all a machine learning–based method that needs to extract the features manually or use some methods before segmentation. This makes the extracted features not various enough, which leads to unsatisfying detection results.

Recently, with the development of generative adversarial networks (GANs) and unsupervised image translation techniques, unsupervised learning–based lesion
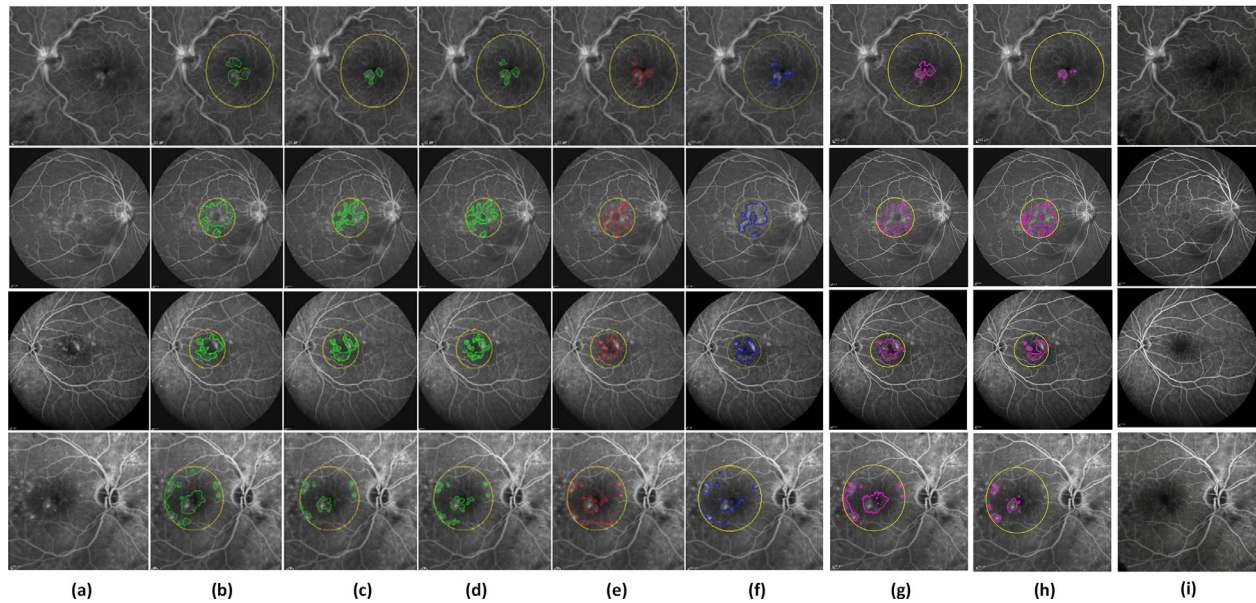
Figure 10. Leakage detection results on the data set by Rabbani et al.[4] (a) Example FA image. Leakage detected by (b) expert 1's annotation, (c) expert 2's annotation, (d) expert 2's annotation after 4 weeks, (e) Rabbani et al.,[4] (f) Zhao et al.,[2] (g) Li et al.[8] (CycleGAN), and (h) the proposed method. (i) The normal-looking FA image generated by the proposed method.

Table 3. Performances of Different Methods on Detecting Focal Leakages over the Data Set by Rabbani et al.[4] at the Pixel Level

Characteristic	Intra obs. (Intraobserver)	Inter obs. (Interobserver)	Method by Rabbani et al.[4]	Method by Zhao et al.[2]	Method by Li et al.[8]	Proposed Method
Sen	0.95 ± 0.05	0.78 ± 0.09	0.69 ± 0.16	0.78 ± 0.06	**0.80 ± 0.06**	0.71 ± 0.04
Spe	0.73 ± 0.27	0.94 ± 0.08	0.91 ± 0.09	0.94 ± 0.02	0.94 ± 0.04	**0.99 ± 0.06**
Acc	0.83 ± 0.16	0.90 ± 0.08	0.86 ± 0.08	0.89 ± 0.06	**0.90 ± 0.05**	0.87 ± 0.06
AUC	0.84 ± 0.16	0.91 ± 0.08	0.80 ± 0.12	0.86 ± 0.04	**0.87 ± 0.06**	0.85 ± 0.02
DC	0.80 ± 0.08	0.82 ± 0.03	0.75 ± 0.05	**0.81 ± 0.02**	0.81 ± 0.04	0.78 ± 0.04

Data are expressed as mean $\pm$ standard deviation. The bold values represent the best results.

detection methods have also emerged. Sun et al.[20] proposed an abnormal-to-normal translation GAN to generate a normal-looking medical image and using the difference between the abnormal and the generated normal image to guide the detection or segmentation of lesions. Schlegl et al.[21] proposed a model to detect the anomaly regions of optical coherence tomography retinal images. As for fluorescence leakage detection, Li et al.[8] proposed an unsupervised learning–based fluorescence leakage detection model based on CycleGAN,[22] which illustrated the potential of the unsupervised learning method. However, this method only applied the CycleGAN to generate the normal-looking image, and this model cannot focus solely on modifying the leakages areas and leaving the no-leakage areas unchanged, as this leads to false leakage detection.

The present study builds upon that by Li et al.,[8] with the purpose of optimizing the CycleGAN-based method and making it concentrate more on the leakage areas and keeping the pixel intensity of the no-leakage areas nearly unchanged. To achieve this objective, we made the following two improvements: (1) a CBAM was introduced and combined with the deep residual blocks. CBAM can extract main features in both spatial- and channel-wise dimensions, which enhances the representations of the important regions (leakage areas). (2) The CAM and its derivatives enable discriminative regions of images to be located with basic classification networks.[10] Inspired by this characteristic of CAM, we designed an anomaly mask loss to make the network focus on the generation of leakage areas. The proposed method is called a weakly supervised method since we need to leverage the normal or abnormal labels

of FA images to train a classification network to generate the anomaly mask of abnormal FA images.

In this work, we evaluated the proposed model on our HRA data set and two publicly available data sets to validate its effectiveness and universality. The presented HRA data set contained FA images with three common types of fluorescein leakage (large focal, punctate focal, and optic disc leakage). As shown in Figure 5, the proposed model obtained the best performance on the optic disc and large focal leakage detection when compared with the other three models. The reason is that the proposed model consists of both AM loss and CBAM, which can make the model focus more on the leakage areas and obtain less background redundant information in the difference image (Figs. 6h1–h3) and then gets better detection results. The proposed method achieved the highest specificity, accuracy, AUC, and DC of 0.97 ± 0.03, 0.95 ± 0.05, 0.86 ± 0.04, and 0.87 ± 0.01, respectively (Table 1). However, the sensitivity of the proposed method is lower than that of the other methods (Table 1) because the difference images of the other methods (Fig. 6) contain lots of redundant background information. The high-intensity redundant information can be identified as leakage, which allows the leakages with similar intensity of background to be detected and leads to high sensitivity, but this also makes the background detected as leakages and results in lower specificity.

As shown in Figures 5 and 7, the proposed method performed well on the detection of large focal and optic disc leakage, but it cannot detect all leakage areas in the FA images with punctate focal leakage. This is because the introduction of the attention module and AM loss make the network focus more on the obvious areas and ignore the less obvious ones. Therefore, improving punctate leakage detection will be the main task in our future study.

The proposed method also compared with three state-of-the-art methods on two publicly available data sets from Zhao et al.[2] and Rabbani et al.[4]

Figure 8 and Table 2 show the qualitative and quantitative results on the data set by Zhao et al.,[2] and it can be seen that the overall performance of the proposed method is slightly better than the existing best intensity-based method by Zhao et al.[2]

The qualitative and quantitative comparable results of the proposed method over the data set by Rabbani et al.[4] are illustrated in Figure 9 and Table 3. As shown in Table 3, the proposed method reached the highest specificity (0.99 ± 0.06), which indicates that the proposed method focuses more on the detection of the real and obvious leakage. For this reason, the proposed method will eliminate the less obvious ones and lead to a low sensitivity (0.71 ± 0.04). Thus,

improving the detection sensitivity of this kind of leakage will also be the main task in our future study.

It should be noted that the HRA data set in our study does not have the disease labels. Since we cannot get the disease labels, the purpose of this study is to propose a method that can automatically and effectively detect the three types of fluorescein leakage: large focal, punctate focal, and optic disc leakage (these types of leakage are common in DR[4,23] and MR[5,24]), so as to achieve the auxiliary diagnosis of the retinal diseases with such fluorescein leakages. Also, the experimental results of the proposed method on two publicly available data sets indicate the effectiveness of the proposed method on DR and MR. This can demonstrate the potential value for application in clinical diagnosis to some extent.

In general, the proposed method can achieve comparable or even better performance to the existing intensity-based method on the publicly available data sets, which can demonstrate the effectiveness and universality of the proposed method to some extent. Furthermore, the detection time of the proposed method (within 1 second) is much less than that of intensity-based methods (more than 20 seconds). This depends on the characteristics of the learning-based method (i.e., spending a long time to train a model but needing a short time to test).

In summary, here we introduce a novel weakly supervised deep learning method for leakage detection in FA images. The results show that the proposed method has better performance on fluorescein leakage detection when compared with other methods, especially the FA image with optic disc and large focal leakage. Moreover, this method can detect one image within 1 second, which is far superior to the intensity-based methods (more than 20 seconds). These indicate the potential value of the proposed method in clinical diagnosis and treatment for retinal diseases, such as DR and MR.

Acknowledgments

References

1. Béouche-Hélias B, Helbert D, de Malézieu C, Leveziel N, Fernandez-Maloigne C. Neovascularization detection in diabetic retinopathy from fluorescein angiograms. *J Med Imaging*. 2017; 4(4):1.

2. Zhao Y, Zheng Y, Liu Y, et al. Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy. *IEEE Trans Med Imaging*. 2016;36(1):51–63.

3. Martínez-Costa L, Marco P, Ayala G, De Ves E, Domingo J, Simó A. Macular edema computer-aided evaluation in ocular vein occlusions. *Comput Biomed Res*. 1998;31(5):374–384.

4. Rabbani H, Allingham MJ, Mettu PS, Cousins SW, Farsiu S. Fully automatic segmentation of fluorescein leakage in subjects with diabetic macular edema. *Invest Ophthalmol Vis Sci*. 2015;56(3):1482–1492.

5. Zhao Y, MacCormick IJC, Parry DG, et al. Automated detection of leakage in fluorescein angiography images with application to malarial retinopathy. *Sci Rep*. 2015;5(1):10425.

6. Trucco E, Buchanan CR, Aslam T, Dhillon B. Contextual detection of ischemic regions in ultra-wide-field-of-view retinal fluorescein angiograms. *Annu Int Conf IEEE Eng Med Biol Soc*. 2007;2007:6740–6743.

7. Tsai C, Ying Y, Lin W. Automatic characterization of classic choroidal neovascularization by using AdaBoost for supervised learning automatic characterization and segmentation of classic choroidal neovascularization using AdaBoost for supervised learning. *Invest Ophthalmol Vis Sci*. 2011;52:2767–2774.

8. Li W, He Y, Wang J, Kong W, Chen Y, Shi G. An unsupervised adversarial learning approach to fundus fluorescein angiography image synthesis for leakage detection. In: *International Workshop on Simulation and Synthesis in Medical Imaging*. Lima: Springer; 2020:142–152.

9. Dutta A, Zisserman A. The VIA annotation software for images, audio and video. In: *Proceedings of the 27th ACM International Conference on Multimedia*. Nice: Association for Computing Machinery (ACM); 2019:2276–2279.

10. Guo H, Xu M, Chi Y, Zhang L, Hua X-S. Weakly supervised organ localization with attention maps regularized by local area reconstruction. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Lima: Springer; 2020:243–252.

11. Zhang X, Wei Y, Feng J, Yang Y, Huang TS. Adversarial complementary learning for weakly supervised object localization. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE Xplore; 2018:1325–1334.

12. Zhou Y, Zhu Y, Ye Q, Qiu Q, Jiao J. Weakly supervised instance segmentation using class peak response. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE Xplore; 2018:3791–3800.

13. Ahn J, Cho S, Kwak S. Weakly supervised learning of instance segmentation with inter-pixel relations. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE Xplore; 2019:2209–2218.

14. Zeng X, Pan Y, Zhang H, Wang M, Tian G, Liu Y. Unpaired salient object translation via spatial attention prior. *Neurocomputing*. 2021;453:718–730.

15. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE Xplore; 2016:770–778.

16. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. *Int J Comput Vis*. 2015;115(3):211–252.

17. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision*. Honolulu: IEEE Xplore; 2017:618–626.

18. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern*. 1979;9(1):62–66.

19. Woo S, Park J, Lee J-Y, Kweon IS. Cbam: convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. Munich: Springer; 2018:3–19.

20. Sun L, Wang J, Huang Y, et al. An adversarial learning approach to medical image synthesis for lesion detection. *IEEE Journal of Biomedical and Health Informatics,* 2020;24(8):2303–2314.

21. Schlegl T, Seeböck P, Waldstein SM, Langs G, Schmidt-Erfurth U. f-AnoGAN: fast unsupervised

anomaly detection with generative adversarial networks. *Med Image Anal*. 2019;54:30–44.

22. Zhu JY, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*. Venice: IEEE Xplore; 2017:2223–2232.

23. Wen Y, Chen L, Qiao L. Let's find fluorescein: cross-modal dual attention learning for fluorescein leakage segmentation in fundus fluorescein angiography. In: *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*. Virtual: IEEE; 2021:1–6.

24. Zhao Y, Su P, Yang J, et al. A compactness based saliency approach for leakages detection in fluorescein angiogram. *Int J Mach Learn Cyber*. 2017;8(6):1971–1979.