2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)
April 13-16, 2021, Nice, France

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) | 978-1-6654-1246-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISBI48211.2021.9433795

RESTORATION OF CATARACT FUNDUS IMAGES VIA UNSUPERVISED DOMAIN ADAPTATION

Heng Li[1], *Haofeng Liu*[1], *Yan Hu*[1,*], *Risa Higashita*[2], *Yitian Zhao*[3], *Hong Qi*[4], *Jiang Liu*[1,3,5,*]

[1] School of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
[2] Tomey Corporation, Nagoya 451-0051, Japan
[3] Cixi Institute of Biomedical Engineering, Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, Ningbo 315201, China
[4] Department of Ophthalmology, Peking University Third Hospital, Beijing 100191, China
[5] Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation, Southern University of Science and Technology, Shenzhen 518055, China

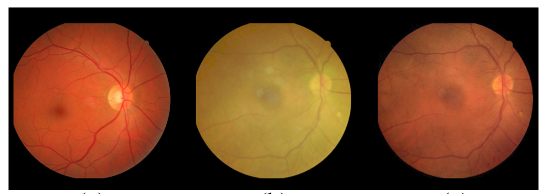
ABSTRACT

Cataract presents the leading cause of preventable blindness in the world. The degraded image quality of cataract fundus increases the risk of misdiagnosis and the uncertainty in preoperative planning. Unfortunately, the absence of annotated data, which should consist of cataract images and the corresponding clear ones from the same patients after surgery, limits the development of restoration algorithms for cataract images. In this paper, we propose an end-to-end unsupervised restoration method of cataract images to enhance the clinical observation of cataract fundus. The proposed method begins with constructing an annotated source domain through simulating cataract-like images. Then a restoration model for cataract images is designed based on pix2pix framework and trained via unsupervised domain adaptation to generalize the restoration mapping from simulated data to real one. In the experiment, the proposed method is validated in an ablation study and a comparison with previous methods. A favorable performance is presented by the proposed method against the previous methods. The code of of this paper will be released at https://github.com/liamheng/Restoration-of-Cataract-Images-via-Domain-Adaptation.

***Index Terms*—** Cataract, image restoration, unsupervised domain adaptation

1. INTRODUCTION

Cataract presents the most prevalent cause of vision loss and blindness in the world, which will lead to 40 million blindness in 2025 estimated by the World Health Organization [1]. A cataract is an opacification of the normally clear lens of the eye, and it develops with aging, general disease, congenital disorder or injury. As cataracts grow, they can lead to a decrease in the vision, including cloudy or blurred vision, faded colors, glare, poor night vision and double vision. Moreover, cataract not only decreases the vision, but also affects the diagnosis and treatment of other eye diseases.



(a) (b) (c)

Fig. 1. Retinal fundus images. (a) from eyes without cataract; (b) from eyes with cataract; (c) restored image from (b).

Retinal fundus photographs have been routinely used in the clinic to diagnose and monitor many ocular diseases. However, it is challenging to diagnose based on hazy cataract fundus images, either by human experts or computer-aided diagnosis systems [2]. Since fundus images are captured through the lens, the light will attenuate and scatter when traveling through the turbulent lens of cataract patients. The image quality of the cataract fundus photograph is consequently severely degraded, which results in the the risk of misdiagnosis and the uncertainty in preoperative planning. As shown in Fig. 1 (a) and (b), the cataract degrades the fundus images and prevents the vessel and disc region from being clearly observed.

Consequently, image restoration takes a significant role in the promotion of the diagnosis and treatment for cataract fundus. Studies on image restoration have been conducted for years, and lots of methods have been proposed to solve the dehaze problems in computer vision. Features extracted from the images, such as dark channel prior and guided image filtering (GIF) [3], were employed as driven knowledge for the image correction. Thanks to the advance in conditional generative adversarial networks (cGAN), image-to-image translation was constructed between degraded images and high-quality ones [4, 5]. Nevertheless, these algorithms often could not properly preserve the retinal anatomical structures, which might be essential in clinical diagnosis or analysis. Inspired by GIF, Cheng et al. [2] proposed a structure-preserving guided retinal image filtering (SGRIF) to achieve the fundus image enhancement. Luo et al. [6] reported a two-stage dehaze algorithm for catatact image using 400 retinal images without cataracts and 400 hazy images from cataract patients. However, it is extremely daunting to acquire the cataract images paired with the corresponding clear ones after surgery. The deficiency of training data hence impacts the restoration of cataract images. This motivated us to develop an unsupervised algorithm to remove the cataractous effect and improve the readability of cataract fundus images.

In this paper, inspired by the technique of domain adaptation [7], an end-to-end unsupervised restoration method is proposed to remove artifacts in cataract fundus images. An instance of the restoration is provided in Fig. 1 (c). The restoration model is derived from an image-to-image translation model by appending structure guidance. And an annotated source domain was constructed by simulating cataract-like images from clear ones. Then unsupervised domain adaptation was introduced to generalize the restoration model from the source domain to the unannotated target domain of real cataract images. The main contributions of this paper are summarized as follows:

- An end-to-end restoration model for cataract fundus images is proposed to enhance the degraded images for clinical diagnosis and treatment.
- By adapting the model from simulated cataract-like domain to real cataract domain, the scarcity of annotated data are circumvent.
- The ablation study and comparison in the experiment demonstrate the capacity of the proposed method that it performed favorably against previous approaches.

2. METHODOLOGY

The proposed algorithm in this paper consists of three components: a simulation model for cataract-like images, an image-to-image translation model and a domain adaptation model. As presented in Fig. 2, the proposed algorithm begins with the simulation of cataract-like images to address the lack of clear and cataract image pairs. Subsequently, unsupervised domain adaptation is integrated with image-to-image translation to construct an end-to-end network, which adapts the restoration model from the simulated cataract-like domain to the real cataract one.

2.1. Cataract-like image simulation

Following the fundus degradation model proposed in [8], cataract-like images are simulated in this study to construct the annotated source domain. Given a clear fundus image s, the simulated cataract-like image s' is defined as:

$$s' = \alpha(s * f_B(r_B, \sigma_B)) + n, \tag{1}$$

where f_B is a Gaussian filter for image blurry with a radius of r_B and spatial constant σ_B, and α denotes the weight for the image. The additive noise n is given by $\beta(J * f_L(r_L, \sigma_L))$, where J is defined as an illumination bias smoothed by the Gaussian filter f_L and β denotes the weight for the noise.

Specifically, the publicly available retinal image dataset DRIVE [1] was employed as the clear images, and the cataract-like ones were consequently simulated. Through the simulation of cataract-like images, the clear and cataract-like images were paired to compose the annotated source domain.

In addition, to augment the training data, ten cataract-like images are generated from each case in DRIVE with random parameters of Eq. 1. Therefore, 400 pairs of source data are acquired to training the restoration model.

2.2. Image restoration by image-to-image translation

The image-to-image translation model of pix2pix [4] is employed as the foundation of our restoration model. And it is trained under the supervision of the source domain, where the simulated cataract-like images and the corresponding clear images were respectively employed as the input and the ground truth. The architecture of pix2pix contains a generator G and a discriminator D_p, while the training loss $\mathcal{L}_p$ is given by:

$$\mathcal{L}_p = \mathcal{L}_{pGAN} + \lambda \mathcal{L}_{L_1}, \tag{2}$$

in which

$$\begin{aligned}\mathcal{L}_{pGAN} =& \mathbb{E}\left[\log D_p\left(s', s\right)\right] + \\ & \mathbb{E}\left[\log\left(1 - D_p\left(s', G\left(s'\right)\right)\right)\right],\end{aligned} \tag{3}$$

$$\mathcal{L}_{L_1} = \mathbb{E}[\|s - G(s')\|_1]. \tag{4}$$

Retinal fundus images are the foundations for clinical diagnosis. Therefore, the incorrectly generated components during the restoration may heavily skew the pathological features. Considering it is fundamental to preserve the realistic and convincing contents in the restoration, the retinal structure is extracted from the cataract images to guide the restoration procedure.

The edges detected by the Sobel operator is adopt to guide the generation of the clear fundus images. As illustrated in Fig. 2, the cataract-like image and the edges detected from it are forwarded to the restoration model. Thus the definition of

[1] http://www.isi.uu.nl/Research/Databases/DRIVE/

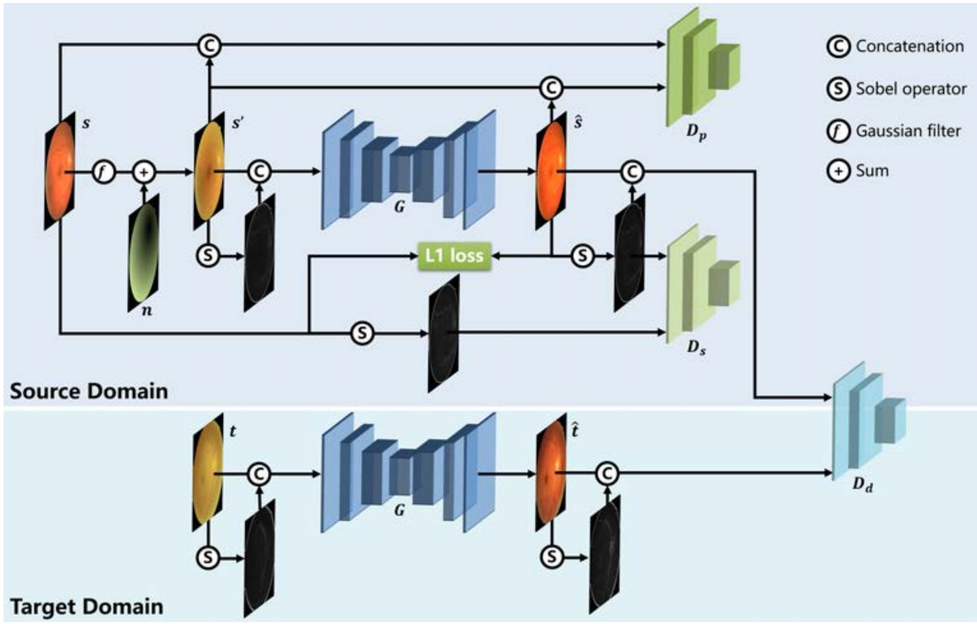


Fig. 2. Architecture of the proposed algorithm. Cataract-like images were simulated to construct the annotated date for the training of G. The Sobel operator and D_s preserved the retinal structure. And D_d generalized the restoration model from the source domain to the unannotated target domain.

the restored image $\hat{s}$ is modified from $G(s')$ to $G(s', \mathrm{S}(s'))$ in Eq. 3 and 4, where $\mathrm{S}(\cdot)$ denotes the Sobel operator.

Additionally, to guarantee the realistic and convincing content has been properly preserved, a discriminator D_s is implemented to distinguish whether the edge is detected from the ground truth s or the restored image $\hat{s}$. And the corresponding structure adversarial loss $\mathcal{L}_s$ is given by:

$$\mathcal{L}_s = \mathbb{E}\left[\log D_s\left(\mathrm{S}(s)\right)\right] + \mathbb{E}\left[\log\left(1 - D_s\left(\mathrm{S}(\hat{s})\right)\right)\right]. \quad (5)$$

2.3. Unsupervised domain adaptation

As the purpose of this study is to restore clear fundus images from the cataract ones, the restored images and the clear ones should belong to the same domain. Consequently, there is a hypothesis that no matter the input images are from the source or target domain, the output images should share strong similarities once they are convincingly restored. We utilize this property to adapt the restoration model from the source domain to the target one via adversarial learning. Specifically, the restored images $\hat{s}$ from the source domain and $\hat{t}$ from the target one, as well as their structure, were forwarded to a domain discriminator D_d using following loss:

$$\mathcal{L}_d = \mathbb{E}\left[\log D_d\left(\hat{s}, \mathrm{S}(\hat{s})\right)\right] + \mathbb{E}\left[\log\left(1 - D_d\left(\hat{t}, \mathrm{S}(\hat{t})\right)\right)\right], \quad (6)$$

where $\hat{t} = G(t, \mathrm{S}(t))$, and t represents real cataract fundus images. $\mathcal{L}_d$ would adapt the restored image $\hat{t}$ to the distribution of $\hat{s}$, through leading G to generate images that should fool D_d.

The final loss function for optimizing G of the proposed method is:

$$\mathcal{L}_G = \mathcal{L}_{pGAN} + \lambda_1 \mathcal{L}_{L_1} + \lambda_2 \mathcal{L}_s + \lambda_3 \mathcal{L}_d, \quad (7)$$

where λ_1, λ_2 and λ_3 are set to 100, 50 and 1 in this paper, respectively.

3. EXPERIMENTS AND EVALUATION

3.1. Compliance with Ethical Standards

The publicly available retinal image dataset DRIVE, and the fundus images of 26 cataract cases collected from Peking University Third Hospital, were used in this study. The experimental protocol is in accordance with the Declaration of Helsinki and has been approved by the local Ethics Committee. Cataract-like images were simulated from DRIVE dataset with the degradation model described in section 2.1 to construct the source domain. And the 26 cataract cases, which have images paired with the clear images from the same patients after surgery, were adopted as the target domain, as well as to evaluate the algorithms. Following [6, 8], the evaluation metric of Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) were utilized to quantify the restored image quality.

The proposed model was implemented in PyTorch. The images were resized to 256×256 with a batch size of 8 for each iteration. The learning rate was initialized to 1×10^{-4}

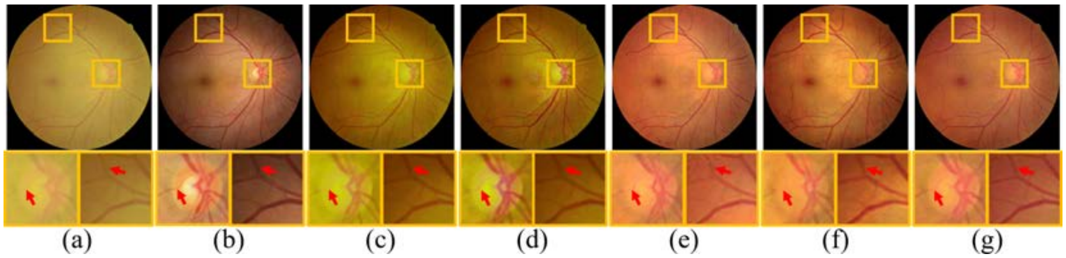


Fig. 3. Comparison of the restored fundus images. The proposed method correctly preserves and enhances the clinical features. (a) cataract image. (b) clear fundus image after surgery. (c) dark channel prior. (d) SGRIF. (e) pix2pix. (f) CycleGAN. (g) the proposed method.

for the first 150 epochs and and then gradually decayed to zero over the next 50 epochs.

3.2. Ablation Study

With the purpose of demonstrating the effectiveness of the proposed method, an ablation study is presented. Once removing the Sobel operator, D_s and D_d from the proposed method, the model becomes a vanilla pix2pix network. The contribution of each module is evaluated by separately adding them back to the network. Table 1 summarizes the results.

Table 1. Ablation study of the proposed method.

Sobel	D_s	D_d	SSIM	PSNR
			0.729	17.47
		√	0.744	17.59
√		√	0.746	17.83
√	√	√	**0.755**	**18.22**

Through the domain adaptation of output space, D_d efficaciously improves the capacity of the vanilla pix2pix network and increases the value of SSIM from 0.729 to 0.744. The Sobel operator and D_s guide the model to preserve and enhance the clinical structure. The top performance achieved by the proposed method shows that the Sobel operator, D_s and D_d provided reasonable and obvious improvements.

3.3. Comparison with previous methods

The proposed method was quantitatively compared with previously published algorithms to demonstrate the advantages. The dehaze method using dark channel prior proposed by He et al. [3], and SGRIF for cataract image dehaze [2] were performed. And the image-to-image translation models, including pix2pix [4] and CycleGAN [5], were trained in the source domain and tested in the target domain. It should be noted that the model reported in [6] was not conducted in the experiment, due to the lack of large dataset of cataract images.

The visualized and quantitative comparisons with previous methods are summarized in Fig. 3 and Table 2. The methods based on image-to-image translation performs more favorable than the dehaze models, since they not only enhance the content but also translate the style of the fundus images. According to the evaluation metrics of SSIM and PSNR, the proposed method outperforms the previous methods and achieves the highest score in the restored image quality. From the visualized comparison illustrated in Fig. 3, it could be inferred that the Sobel operator and the structure discriminator endowed the proposed method with an eminent ability on preserving and enhancing the clinical structure.

Table 2. Quantitative Comparison with previous methods.

Algorithms	SSIM	PSNR
Dark channel prior [3]	0.687	16.11
SGRIF [2]	0.633	15.05
pix2pix [4]	0.729	17.47
CycleGAN [5]	0.731	17.51
The proposed method	**0.755**	**18.22**

Furthermore, as a result of training in the source domain, the appearance of the restored images are close to DRIVE instead of the fundus after surgery (as shown in Fig. 1 (a) and Fig. 3 (b)). In our following work, more public datasets of fundus images will be tested, and efforts will be put on minimizing the gap between the public datasets and the collected images.

4. CONCLUSIONS

The low quality in retinal fundus images due to cataract, increases the uncertainty in clinical observation and treatment. In this paper, an end-to-end unsupervised restoration method has been proposed to enhance the contrast within cataract fundus images. Cataract-like images are first simulated from clear ones to construct an annotated source domain. Subsequently, the proposed model learns the restoration mapping from the source domain and generalizes it to the target domain of real cataract images. The proposed method performs favorably against the previous methods in terms of fidelity and image quality.

Acknowledgment

This study is supported by Ningbo Public Welfare Science and Technology Project (2019C50049) and Ningbo Natural Science Foundation(2019A610354). The authors would like to thank the members of the Intelligent Medical Imaging Program (iMED) for helping.

5. REFERENCES

[1] Wei Wang et al., "Cataract surgical rate and socioeconomics: a global study," *IOVS*, vol. 57, no. 14, pp. 5872–5881, 2016.

[2] Jun Cheng et al., "Structure-preserving guided retinal image filtering and its application for optic disk analysis," *IEEE TMI*, vol. 37, no. 11, pp. 2536–2546, 2018.

[3] Kaiming He et al., "Single image haze removal using dark channel prior," *IEEE PAMI*, vol. 33, no. 12, pp. 2341–2353, 2010.

[4] Phillip Isola et al., "Image-to-image translation with conditional adversarial networks," in *CVPR*, 2017, pp. 1125–1134.

[5] Jun-Yan Zhu et al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *ICCV*, 2017, pp. 2223–2232.

[6] Yuhao Luo et al., "Dehaze of cataractous retinal images using an unpaired generative adversarial network," *JBHI*, 2020.

[7] Eric Tzeng et al., "Adversarial discriminative domain adaptation," in *CVPR*, 2017, pp. 7167–7176.

[8] Ziyi Shen et al., "Understanding and correcting low-quality retinal fundus images for clinical analysis," *arXiv preprint arXiv:2005.05594*, 2020.