DOMAIN GENERALIZATION IN RESTORATION OF CATARACT FUNDUS IMAGES VIA HIGH-FREQUENCY COMPONENTS

Haofeng Liu[1], *Heng Li*[1,∗], *Mingyang Ou*[1], *Yitian Zhao*[2], *Hong Qi*[3], *Yan Hu*[1,∗], *Jiang Liu*[1,2,4]

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

ABSTRACT

Cataracts are the most common blinding disease, and also impact the observation of the fundus. To boost the fundus examination of cataract patients, restoration algorithms have been proposed to address the degradation of fundus images caused by cataracts. However, it is impractical in clinics to collect paired or annotated fundus images for developing restoration models. In this paper, a restoration algorithm is designed for cataractous images without paired or annotated data. Domain generalization (DG) is applied to learn domain-invariant features (DIFs) from synthesized data, and the high-frequency components (HFCs) are extracted to conduct domain alignment. The proposed algorithm is used on unseen target data in the experiments. The effectiveness of the algorithm is demonstrated in the ablation study and compared with state-of-the-art methods. The code of this paper will be released at https://github.com/HeverLaw/Restoration-of-Cataract-Images-via-Domain-Generalization.

Index Terms— Cataract, image restoration, domain generalization, domain-invariant features

1. INTRODUCTION

Cataracts are one of the leading causes of blindness and vision loss in the world, and surgery is the most effective treatment for cataracts in the clinic [1]. Unfortunately, the surgical treatment is of severe risk and uncertainty for cataract patients with ocular comorbidities [2], where more than 30% of cataract patients also suffer from fundus diseases [3]. The turbidity in the cataractous lens attenuates and scatters the light for fundus observation [4], resulting in challenges for preoperative fundus examination of cataracts patients.

The degraded quality of cataract fundus images increases the risk of misdiagnosis and the uncertainty of surgery. Thus

∗ Correspondence: lih3, huy3@sustech.edu.cn

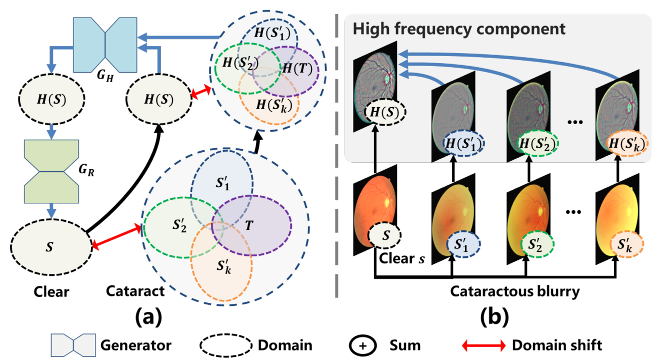

Fig. 1. Overview of the proposed model. The bottom of (a) and (b) exhibit that k cataract-like fundus images s'_i are randomly synthesized from an identical clear image s to cover the potential target domain T . On the top part, HFCs $H(\cdot)$ are extracted from the images to reduce the domain shift and then achieve domain alignment. Finally, the clear image is reconstructed from the aligned HFCs.

image restoration had been proposed to clearly visualize the fundus and promote the preoperative fundus examination [5, 6]. Extensive image denoising and enhancement algorithms, such as CLAHE [7] and SGRIF [8], were modified to improve the quality of fundus images collected from cataract patients. With the advances of CNN-based algorithms, novel networks [9, 10, 11] were designed to automatically suppress the random noise in images. And segmentation masks of vessels and optic disc are imported to the networks for preserving fundus structures in the restoration [12, 13]. However, the above algorithms require high-low quality image pairs to supervise the model training and segmentation annotations to enhance the structures, which are extremely costly in medical scenarios. To mitigate the requirements, a restoration network is proposed by Li et al. [14] to correct cataract fundus images using domain adaptation (DA) between the synthesized and real domains. Nevertheless, clinic images are still necessary in the training stage of networks with DA. This condition limits their applicability for new clinical centers.

To boost the clinical application of restoration for cataract fundus images, the following challenges should be addressed: a) training the restoration model without high-low quality image pairs; b) preserving fundus structures during the restoration; c) agnostic to the variance across clinical images. Although DA allows the restoration models to properly adapt to real images, the requirement of real images for model training is unpractical in applications. Through introducing variances to source domains, DG learns DIFs, which can be straightforwardly generalized to unseen target domains [15].

Motivated by DG, this paper proposes a restoration network for cataractous images without annotation or target data. As shown in Fig. 1, to construct source domains for DG, cataract-like images are firstly randomly simulated. Then, DIFs are acquired by the domain alignment using HFCs. Finally, restored images are reconstructed from the aligned HFCs. The proposed algorithm is independent of annotation or target data in the training stage, and achieves decent performance in unseen target domains. The main contributions of this paper are summarized as follows:

- We propose a DG-based restoration model to effectively restore cataractous images unseen in training without any paired or annotated data.
- Source domains are constructed by randomly synthesizing cataract-like images, and then DIFs are captured using the domain alignment of HFCs.
- Adequate experiments show that our model trained without any annotation or target domain data can achieve state-of-the-art performance steadily.

2. METHODOLOGY

An overview of the proposed restoration algorithm is shown in Fig. 2. Domain randomization (DR) is first implemented by synthesizing multiple cataract-like images. Then, HFCs are adopted to reduce the domain shift and conduct domain alignment. Finally, the aligned HFCs are used as the DIFs to reconstruct the clear fundus image.

2.1. DR by Cataract-like Image Synthesis

DR importing variances to source domains is fundamental for implementing DG [16]. Motivated by the fundus imaging model for cataract patients in [4], cataract fundus images can be synthesized to conduct DR. According to [4], the cataract fundus image I is reformulated as:

$$I = \alpha \cdot L \cdot \gamma + (1-\tau)(L - \alpha \cdot L \cdot \gamma). \quad (1)$$

where α is the attenuation constant of retinal illumination. L is the flash illumination of the fundus camera, γ and τ represent the retinal reflectance function and the transmission function of the lens.

To simulate the uneven transmission caused by cataracts, a panel J is adopted to model $(1-\tau)$. As the illumination intensity of fundus changed under different spectral illuminations, Eq. 1 is decomposed into RGB channels as:

$$C(s_c) = \alpha \cdot s_c * g_B(r_B, \sigma_B) + \beta \cdot J * g_L(r_L, \sigma_L) \cdot (L_c - s_c), \quad (2)$$

where s_c is the clear image, subscript $c \in r, g, b$ denotes the image channels of red, green, and blue. α and β denote the weight of the clear image and the noise from cataract. The $*$ is the convolution operator and $g(r, \sigma)$ is the Gaussian filter with a radius of r and spatial constant σ. In the experiment, r was set to 11 and σ was set to 10.

As the figure illustrated in Fig. 1 (a), multiple cataract-like images are simulated using Eq.2 with random parameters to construct source domains with variances. Specifically, K cataract-like images $[s'_1, s'_2, ...s'_k, ...s'_K]$ are randomly simulated from a clear image s, where K was set as 16.

2.2. Domain Alignment via HFCs

In the restoration of cataractous images, fundus structures need to be preserved and enhanced. Therefore, the fundus structures are used as the DIFs in the restoration task. Because the retinex theory [17] pointed out that the blur illumination bias is considered to be a low-frequency noise in the image, and the fundus structures should be contained in the HFCs of images. The HFCs are thus extracted and aligned to capture the DIFs of fundus images.

The HFCs $H(I)$ are effectively extracted by removing the low frequency components (LFCs) from the fundus image:

$$H(I) = I - I * g_P(r_P, \sigma_P) \quad (3)$$

where r_P and σ_P represent the radius and spatial constant of low-pass Gaussian filter g_P. Due to removing the most domain-variant features, the shift between domains is reduced by the HFCs.

Then as shown in Fig. 2, a generator G_H is implemented to conduct domain alignment between the HFCs of cataract-like images and the clear one. L_1 distance is calculated to align the HFCs:

$$\mathcal{L}_H = \mathbb{E}\left[\sum_{k=1}^{K} \|H(s) - G_H(H(s_k{}'))\|_1\right]. \quad (4)$$

And an identity loss is also implemented to prevent changing clear fundus images:

$$\mathcal{L}_{idt} = \mathbb{E}[\|H(s) - G_H(H(s))\|_1]. \quad (5)$$

The final loss for domain alignment is given by:

$$\mathcal{L}_{G_H} = \mathcal{L}_H + \lambda_1 \mathcal{L}_{idt} \quad (6)$$

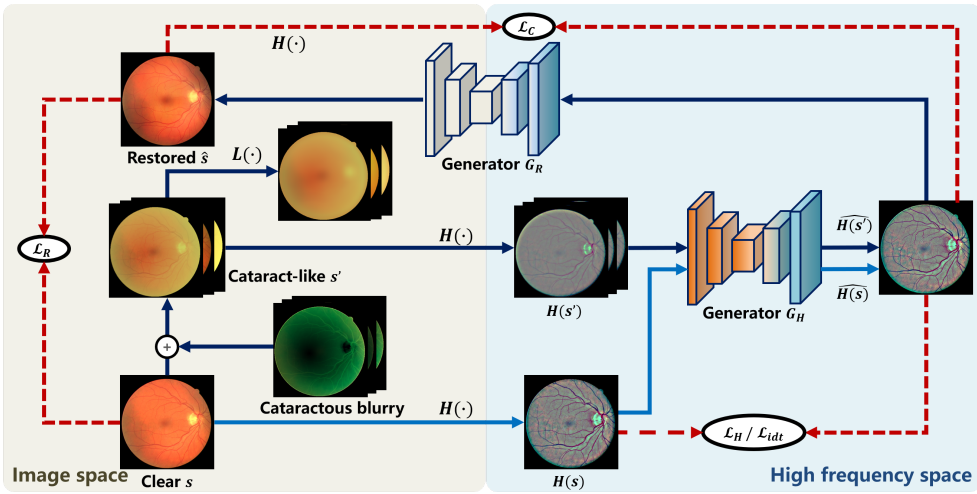

Fig. 2. Overview of the proposed model. Cataract-like images s' are synthesized from clear image s using DR to construct source domains. $H(\cdot)$ and $L(\cdot)$ are the extraction of HFCs and LFCs. Then, DIFs are acquired by domain alignment using HFCs and generator G_H. Finally, generator G_R reconstructs the clear fundus image from the aligned HFCs.

where λ_1 is the weight for balancing the losses which is equal to 0.01 in this model. Through domain alignment in HFCs, DIFs for the restoration task are acquired.

2.3. Clear Fundus Reconstruction

Since the color and style of the fundus images are removed with the LFCs, clear fundus images are reconstructed from the aligned HFCs, $\widehat{H(s')} = G_H(H(s'))$, for convenient visualization.

Another generator G_R is used to translate the aligned HFCs to clear fundus images. The loss function for reconstruction is defined as:

$$\mathcal{L}_R = \mathbb{E}[\left\| G_R(\widehat{H(s')}) - s \right\|_1]. \tag{7}$$

In addition, a consistency loss is also introduced using HFCs to preserve the structures in the reconstruction by:

$$\mathcal{L}_C = \mathbb{E}\left[\left\| H(G_R(\widehat{H(s')})) - \widehat{H(s')} \right\|_1 \right]. \tag{8}$$

The total loss function for G_R is defined by:

$$\mathcal{L}_{G_R} = \mathcal{L}_R + \lambda_2 \mathcal{L}_C \tag{9}$$

where λ_2 is for balancing the weights for reconstruction and structure preservation, which is set to 1.

As described above, a DG-based restoration algorithm for cataractous images is proposed without access to annotation or target data. It should be noted that the two generators for domain alignment and clear fundus reconstruction are trained independently.

3. EXPERIMENTS AND EVALUATION

3.1. Compliance with Ethical Standards

As summarized in Table 1, three fundus image datasets were applied in the experiments. Two public datasets, DRIVE[1] and Kaggle[2] were respectively imported to implement the proposed algorithm. The private dataset was applied as the target domain to evaluate the generalization performance of the model. The experiment followed the Declaration of Helsinki and obtained approval from the local ethics committee.

Table 1. Overview of the adopted datasets.

Clear fundus image datasets	1) DRIVE: 40 clear fundus images with segmentation masks. 2) Normal subset of Kaggle: 300 clear images.
Target domain datasets	3) Private dataset: 26 image pairs of cataract patients collected before and after surgery.

To evaluate the performance, we applied Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) to quantify the restored image quality with the reference of the postoperative images. The Unet-like network [18] was used as the backbone of G_R and G_H. The networks were trained for 200 epochs with an image size of 256×256 and a batch size of 8. The optimizer was Adam with the momentum of 0.9 and 0.99, and the learning rate was set to 2×10^{-4} in 150 epochs and 5×10^{-5} in the next 50 epochs.

3.2. Ablation Study

An ablation study was conducted to demonstrate the effectiveness of the proposed modules. As summarized in Table 2, the

[1]http://www.isi.uu.nl/Research/Databases/DRIVE/
[2]https://www.kaggle.com/jr2ngb/cataractdataset

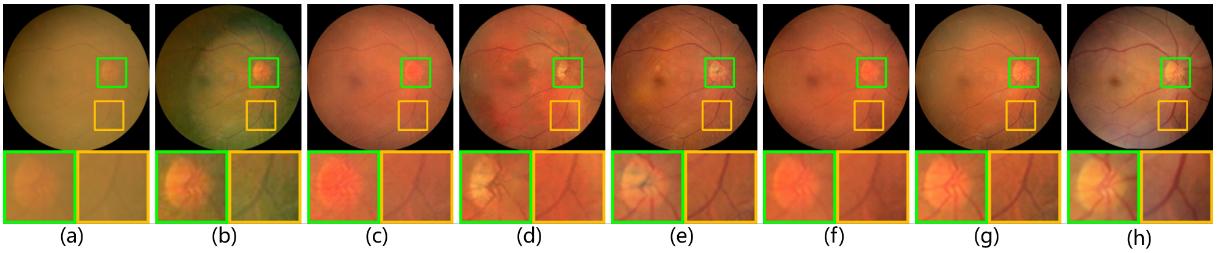

Fig. 3. Comparison between the cataract restoration algorithms. (a) cataract fundus image. (b) SGRIF [8]. (c) pix2pix [18]. (d) Luo et al. [12]. (e) CofeNet [13]. (f) Li et al. [14]. (g) The proposed method. (h) clear image after surgery.

source domains were synthesized with DRIVE and our annotated private dataset was used to test the algorithm. DR refers to randomly simulating multiple cataract-like images, $H(\cdot)$ represents the extraction of HFCs, and the domain alignment is implemented by G_H.

Table 2. Ablation study of the proposed method.

DR	HFC	G_H	SSIM	PSNR
			0.729	17.47
√			0.731	17.80
√	√		0.755	18.02
√	√	√	**0.769**	**18.32**

With the deployment of the proposed modules, the restoration performance is gradually improved. DR endows the models to learn DIFs from source domains efficiently. HFCs preserve not only the fundus structures but also reduce the shift between domains. DIFs are explicitly acquired by the domain alignment using G_H. Consequently, the model is desirably generalized to the unseen target domain.

3.3. Comparison with state-of-the-art methods

To verify the performance of the proposed method, comparisons were conducted with state-of-the-art methods. The algorithms for comparison consist of a hand-crafted algorithm SGRIF [8], and CNN-based algorithms, including pix2pix [18], Luo et al. [12], CofeNet [13], and Li et al. [14]. Furthermore, the datasets of DRIVE and Kaggle were respectively used to synthesize cataract-like images for understanding how the source domains affect the performance.

Fig. 3 exhibits the restoration results with the models trained on the cataract-like images synthesized from DRIVE. As the image to be restored suffers from severe cataracts, it is challenging to enhance the quality. CofeNet [13] enhanced the vessels but led to artifacts in the optic disk. The image restored by Li et al. [14] had limited contrast. The proposed algorithm presented a superior performance in structure preservation and intensity contrast.

Table 3 summarizes the quantitative results, where the proposed method achieves superior performance against the state-of-the-art methods. CofeNet [13] employing segmentation masks to boost the enhancement of vessels, achieved decent performance. Unsupervised DA was used by Li et al. [14] to adapt the restoration model to the target domain, and presented outstanding performance. However, these methods requiring annotations or target data to train the restoration model are not practical enough in clinics. Based on the DG technique, the proposed algorithm is independent of the annotations or target data in training and outperformed the state-of-the-art methods.

Moreover, comparable performances were presented by the models trained with DRIVE and Kaggle, where the proposed algorithm outperformed other methods using both datasets. This illuminates that the proposed algorithm is robust to the selection of training data.

Table 3. Comparisons on DRIVE and Kaggle.

Algorithms	DRIVE		Kaggle	
	SSIM	PSNR	SSIM	PSNR
SGRIF [8]	0.609	15.07	0.609	15.07
pix2pix [18]	0.731	17.68	0.739	17.71
Luo et al. [12]	0.704	17.29	0.735	17.74
CofeNet [13]	0.754	18.03	0.759	18.54
Li et al. [14]	0.755	18.07	0.748	18.24
The proposed method	**0.769**	**18.32**	**0.765**	**18.64**

The experiments of ablation study and comparison demonstrate the effectiveness of the proposed algorithm. The designed modules implement DG such that the model learned without annotations and target data performs outstandingly on the unseen target domain. The proposed algorithm provides superior performance and clinical applicability compared with the state-of-the-art methods.

4. CONCLUSIONS

Cataracts impact the quality of fundus imaging, leading to challenges in fundus examination. To boost the clinical fundus examination, this article proposes a DG-based restoration algorithm to enhance cataract fundus images. The proposed algorithm is independent of annotations or target data, and provides a superior restoration model to unseen target data. In the experiments, the proposed algorithm outperforms state-of-the-art algorithms, and the effectiveness of the designed modules is verified.

Acknowledgment

This work was supported in part by Guangdong Provincial Department of Education (2020ZDZX3043), Guangdong Provincial Key Laboratory (2020B121201001), Guangdong Basic and Applied Fundamental Research Fund Committee (2020A1515110286), and Shenzhen Natural Science Fund (JCYJ20200109140820699 and 20200925174052004).

5. REFERENCES

[1] Cameron M Lee et al., "The global state of cataract blindness," *Current opinion in ophthalmology*, vol. 28, no. 1, pp. 98–103, 2017.

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

[3] Thuan Quoc Pham et al., "Systemic and ocular comorbidity of cataract surgical patients in a western sydney public hospital," *Clinical & experimental ophthalmology*, vol. 32, no. 4, pp. 383–387, 2004.

[4] Eli Peli et al., "Restoration of retinal images obtained through cataracts," *IEEE TMI*, vol. 8, no. 4, pp. 401–406, 1989.

[5] Tao Li et al., "Applications of deep learning in fundus images: A review," *Medical Image Analysis*, p. 101971, 2021.

[6] Heng Li et al., "An annotation-free restoration network for cataractous fundus images," *IEEE TMI*, pp. 1–1, 2022.

[7] Mei Zhou et al., "Color retinal image enhancement based on luminosity and contrast adjustment," *IEEE TBME*, vol. 65, no. 3, pp. 521–527, 2017.

[8] 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.

[9] Tao Huang et al., "Neighbor2neighbor: Self-supervised denoising from single noisy images," in *CVPR*, 2021, pp. 14781–14790.

[10] Jaakko Lehtinen et al., "Noise2noise," in *ICML*. PMLR, 2018.

[11] Heng Li et al., "Ct scan synthesis for promoting computer-aided diagnosis capacity of covid-19," in *ICIC*. Springer, 2020, pp. 413–422.

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

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

[14] Heng Li et al., "Restoration of cataract fundus images via unsupervised domain adaptation," in *ISBI*. IEEE, 2021, pp. 516–520.

[15] Kaiyang Zhou et al., "Domain generalization: A survey," *arXiv preprint arXiv:2103.02503*, 2021.

[16] Myeongjin Kim et al., "Learning texture invariant representation for domain adaptation of semantic segmentation," in *CVPR*, 2020, pp. 12975–12984.

[17] Edwin H Land, "The retinex theory of color vision," *Scientific american*, vol. 237, no. 6, pp. 108–129, 1977.

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