Medical & Biological Engineering & Computing (2024) 62:357–369
https://doi.org/10.1007/s11517-023-02937-5

ORIGINAL ARTICLE

Adaptive enhancement of cataractous retinal images for contrast standardization

Bingyu Yang[1] · Lvchen Cao[2] · He Zhao[1] · Huiqi Li[1] · Hanruo Liu[3] · Ningli Wang[3]

Received: 8 May 2023 / Accepted: 9 September 2023 / Published online: 18 October 2023

Abstract

Cataract affects the quality of fundus images, especially the contrast, due to lens opacity. In this paper, we propose a scheme to enhance different cataractous retinal images to the same contrast as normal images, which can automatically choose the suitable enhancement model based on cataract grading. A multi-level cataract dataset is constructed via the degradation model with quantified contrast. Then, an adaptive enhancement strategy is introduced to choose among three enhancement networks based on a blurriness classifier. The blurriness grading loss is proposed in the enhancement models to further constrain the contrast of the enhanced images. During test, the well-trained blurriness classifier can assist in the selection of enhancement networks with specific enhancement ability. Our method performs the best on the synthetic paired data on PSNR, SSIM, and FSIM and has the best PIQE and FID on 406 clinical fundus images. There is a 7.78% improvement for our method compared with the second on the introduced P_h score without over-enhancement according to P_{oe}, which demonstrates that the proper enhancement by our method is close to the high-quality images. The visual evaluation on multiple clinical datasets also shows the applicability of our method for different blurriness. The proposed method can benefit clinical diagnosis and improve the performance of computer-aided algorithms such as vessel tracking and vessel segmentation.

Keywords Retinal image enhancement · Contrast standardization · Blurriness grading · Adaptive enhancement

1 Introduction

Cataract is one of the common eye diseases in clinical photography and diagnosis. The quality of cataractous fundus images is not always satisfactory due to the opacity of the refractive medium. Low-quality cataractous fundus images often have low contrast globally or locally, which affects the accuracy of clinical diagnosis as well as the performance of automatic retinal image processing [1]. Retinal image enhancement can be employed as a pre-processing step for cataractous retinal images to improve the visibility of image structure and lesion area [2]. Figure 1 shows the synthetic cataractous image and the enhancement example via the proposed method. The influence of blurriness on vessel segmentation is also displayed in Fig. 1.

The blurriness of cataractous images is divided into four classes [4], which is shown in Fig. 2. A non-cataract image is displayed in Fig. 2a, in which the retinal structure such as optic disc and blood vessels can be seen clearly. Therefore, it does not require enhancement. Figure 2b, c, and d are the fundus images of mild, moderate, and severe cataracts respectively. The fundus structure is gradually blurred and becomes invisible with the aggravation of cataract. The contrast of images decreases as the cataract worsens. Accordingly, they require different levels of enhancement to standardize the contrast to the normal fundus.

Many classical [5, 6] and deep learning fundus image enhancement methods [7, 8] have been investigated in recent years. It is difficult to obtain clear and blurry fundus image pairs clinically for the training of deep learning methods. Therefore, some methods degenerate high-quality fundus images to get corresponding blurry images [9]. Convolutional neural networks or transformer-based networks are used for the enhancement in these deep learning methods [10].

✉ Huiqi Li
huiqili@bit.edu.cn

1 Beijing Institute of Technology, Beijing 100081, China

2 School of Artificial Intelligence, Henan University, Zhengzhou 450046, China

3 Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China

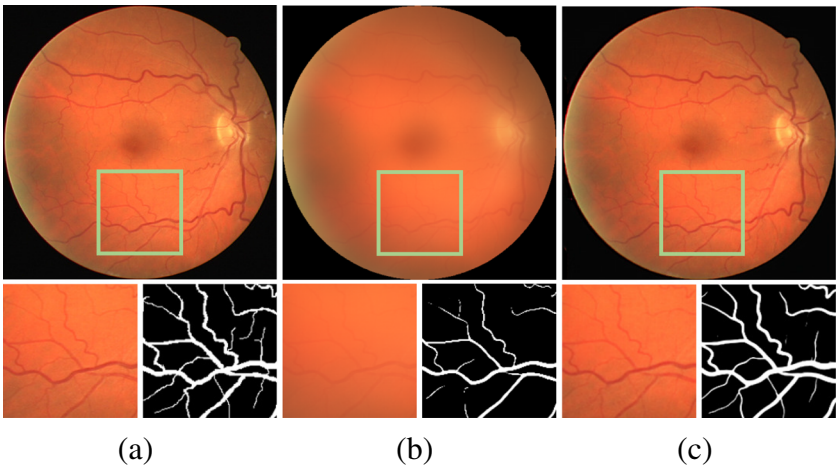

Fig. 1 Enhancement of the cataractous fundus image. The bottom left is the enlarged part, and the bottom right is the vessel segmentation. **a** High-quality fundus image and vessel labels from DRIVE [3]. **b** Cataractous fundus image synthesized through the introduced degeneration model. **c** Enhancement of the synthesized cataractous image through the proposed method. The image contrast is adjusted to the appropriate range, and blood vessels become clear after enhancement

However, these methods do not discriminate different blurriness in the enhancement and usually use a single enhancement network for different blurriness and degradation, which may limit the accuracy of the enhancement. As a single network is employed and cataractous images with different blurriness are put together as the low-quality image in the training, inappropriate enhancement such as over-enhancement or under-enhancement usually cannot be avoided. In addition, high-quality images will also be processed by these methods, which will lead to a decrease of the image quality.

In this paper, we propose a cataractous fundus image enhancement framework with a blurriness classifier and multiple enhancement networks to standardize the image contrast to a normal range. An adaptive enhancement strategy that combines cataract grading and fundus image generation algorithms is employed in this paper. During test, we use the blurriness classifier to estimate the blurriness of the fundus image and select the corresponding enhancement network. The image is no longer processed when it is classified as high quality. The main contribution of this paper is as follows:

- We are the first to incorporate cataract grading in deep learning-based retinal image enhancement methods. Our method can adaptively normalize the contrast of different cataractous images.
- A degradation model for cataractous image synthesis is introduced, where the blurriness is quantified with the ability to correspond to clinical cataractous image.
- Blurriness grading loss is employed to further constrain the contrast of the enhanced images, which can reduce over-enhancement as well as under-enhancement.
- We propose two metrics, P_h and P_{oe}, to evaluate the degree of enhancement. Statistical results indicate that image contrast after our enhancement is similar to high-quality fundus images, which means image contrast is standardized.

2 Related work

2.1 Classical retinal image enhancement

Classical retinal image enhancement algorithms can be categorized into three groups: histogram equalization (HE)-based methods [11, 12], retinex theory-based methods [6, 13], and image formation model (IFM)-based methods [14, 15]. HE-based methods like the contrast limited adaptive histogram equalization [5] improve the histogram equalization to reduce the over-enhancement. Retinex theory-based methods are also widely used in image enhancement, and estimating the illumination, reflection, and scale is the key factor [6]. IFM-based methods estimate the background light and transmission map to enhance fundus images, and dark channel prior [16] is always employed. Considering the complexity of degeneration, the above classical enhancement methods are sometimes used jointly or step by step [11, 13].

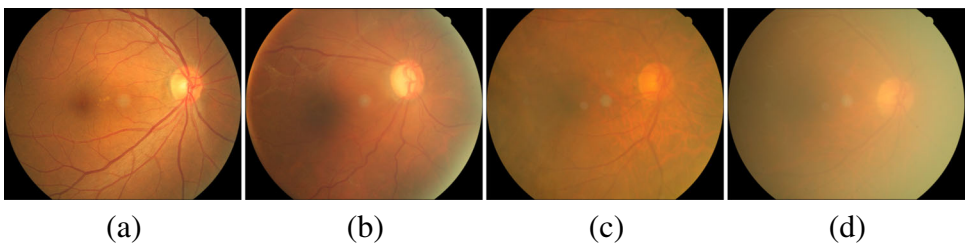

Fig. 2 Cataractous fundus images of different blurriness. **a** Non-cataract. **b** Mild cataract. **c** Moderate cataract. **d** Severe cataract. The blurriness increases as the cataract deteriorates

Table 1 Comparison of different methods

Method	Training data	Advantages and disadvantages
Classical method	None	A: No training process
		D: Require prior knowledge to design multiple steps
Enhancement via unpaired image translation	Unpaired data	A: Can leverage the abundant unpaired data
		D: Artifacts sometimes appear
Enhancement via image degradation	Synthetic paired data	A: Networks are straightforward and easy to converge
		D: Rely on degradation algorithms

A and *D* represent advantages and disadvantages

Different kinds of degeneration, such as insufficient brightness or low contrast, are processed separately in classical methods [14, 17]. The discriminative treatment is straightforward and effective for different degeneration, which can be utilized in deep learning methods for cataractous retinal images with different contrast.

2.2 Enhancement based on retinal image degradation

Acquiring pixel-to-pixel retinal image pairs of different clarities under the same imaging condition such as angle, focal length, and exposure is extremely difficult in retinal photography. Some fundus image enhancement methods employ the cycle consistency [18] through unpaired image translation to loosen the constraint of insufficient paired training data [19]. However, unreal artifacts sometimes appear without the pixel level constraint in these methods [20].

Therefore, researchers use the degradation of high-quality fundus images to obtain corresponding blurred fundus images. Then, the corresponding network is designed and trained by these pre-generated paired data. A fundus image degradation model is proposed in [9], where light transmission disturbance, image blurring, and retinal artifact are simulated on high-quality fundus images to obtain the degraded images. Then, a clinically oriented fundus enhancement network (cofe-Net) is proposed to correct the low-quality fundus images. Luo et al. [21] designed a two-step fundus image degradation method. The CataractDehazeNet is trained using pairs of synthesized cataract-like images through supervised learning. An additive cataract blur is designed in [22], which gives a higher intensity at the fundus border to simulate the disparity of illumination intensity. Then, the annotation-freely restoration network (ArcNet) for cataractous fundus images is proposed to boost the clinical practicability of restoration. These methods usually utilize a single enhancement network to deal with different image degeneration. Table 1 shows the comparison of different methods.

2.3 Cataract grading

Automatic cataract grading and diagnosis algorithms have been studied recently [23], involving data in different modalities [24] such as color fundus photographs and slit lamp photographs. DST-ResNet and EDST-ResNet [25] are proposed for cataract classification that can prevent overfitting and reduce storage memory. In [26], a global–local attention network is proposed to handle the cataract classification task. Keenan et al. [27] developed the DeepLensNet for the classification of age-related cataract and the performance approaches that outperforms ophthalmologists and medical students. An automated deep learning-based artificial intelligence platform is proposed in [28], where three convolutional neural networks are ensembled in their algorithms. The grading can provide auxiliary information such as the blurriness level for the adaptive enhancement, which is not exploited by current enhancement methods yet.

3 Method

3.1 Overview

Our approach first degenerates high-quality fundus images and then trains enhancement networks, which belongs to enhancement based on retinal image degradation. The flowchart is shown in Fig. 3. Denote the high-quality fundus image as $\boldsymbol{x_h} \in \mathbb{R}^{W\times H\times 3}$, where W and H represent the width and height of the image. The cataract-like fundus images $\{\boldsymbol{x_b^k}\}_{k=1}^3 \in \mathbb{R}^{W\times H\times 3}$ are obtained through the degradation model M, where $k \in \{1, 2, 3\}$ represents the blurriness grades of the image. Image contrast varies among different blurriness. After getting the multi-level cataract dataset, three sets of generative adversarial networks $\{G_k, D_k\}_{k=1}^3$ with the same structure are trained using the corresponding synthetic cataractous image and high-quality image pairs $\{\{\boldsymbol{x_b^k}\}_{k=1}^3, \boldsymbol{x_h}\}$ to enhance different blurriness. The classifier

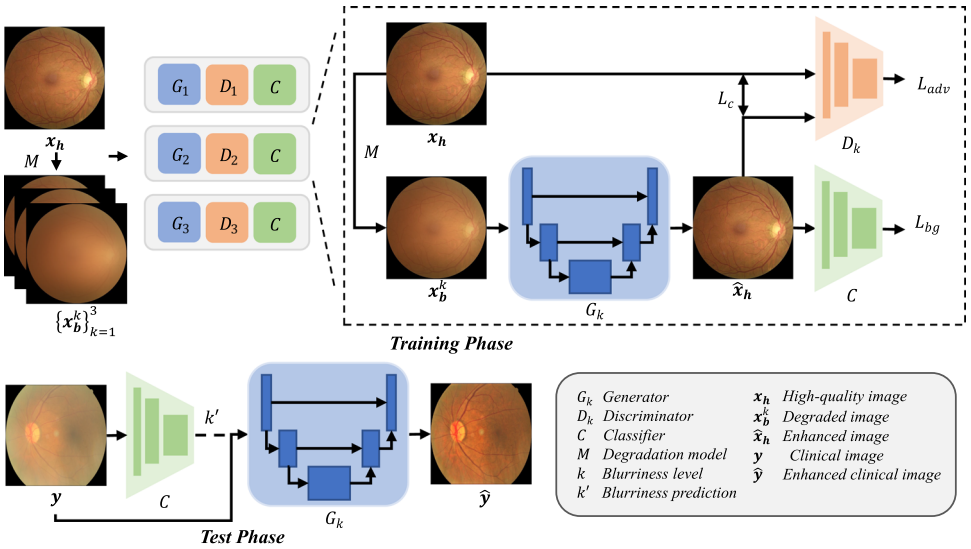

Fig. 3 Flowchart of our approach. First, synthetic fundus images $\{\boldsymbol{x}_{\boldsymbol{b}}^k\}_{k=1}^3$ with different blurriness are obtained through the degradation model M. During training, the blurriness classifier C is trained in advance. Then, multiple networks $\{G_k, D_k\}_{k=1}^3$ with the same structure are trained independently to enhance different blurriness. During test, the classifier C predicts the blurriness k' of the clinical fundus image $\boldsymbol{y}$. It will be sent to the corresponding generator G_k when it is classified as blur, and it will not be processed when it is classified as high quality

C is trained in advance. Then, the blurriness grading loss $\mathcal{L}_{bg}$ through the classifier C can constrain the enhancement level to avoid over-enhancement and under-enhancement. During the test phase, we use the classifier to predict the blurriness level k' of the clinically collected cataractous fundus image $\boldsymbol{y}$. If the image is classified as blurred, it will be sent to the corresponding enhancement network G_k for enhancement, and it will not be processed if the image is classified as high quality.

3.2 Multi-level cataractous fundus image synthesis

3.2.1 Degradation model

It is difficult to obtain cataractous and high-quality image pairs clinically. Therefore, we propose a degradation model for cataractous fundus image synthesis. The degradation process is based on the image formation model [16], which is widely used in natural image dehazing and has shown good performance in the synthesis of hazy images [29]. The degradation model M can be expressed as

$$\boldsymbol{x_b} = M(\boldsymbol{x_h}) = \boldsymbol{x_h} \cdot t + A(\boldsymbol{x_h}) \cdot (1 - t), \tag{1}$$

where $\boldsymbol{x_h}$ is the input high-quality image, $\boldsymbol{x_b}$ is the degenerated cataract-like image, and $A(\cdot)$ is used to estimate the local average brightness of the high-quality images, where a two-dimensional Gaussian convolution is employed. The radius of the Gaussian kernel is set as $\frac{1}{6}$ of the image size to cover the basic retinal structure like optic disc, and the standard deviation is set as the radius divided by π. The transmission parameter $t \in (0, 1]$ is adjustable to control the blurriness, which describes the transmissivity of the refractive medium in the fundus. In fundus photography, the reflection light from the retina passes through the vitreum, lens, and cornea before it reaches the camera. If the eye is healthy, the refractive medium is often with good transmissivity and the acquired image is of good clarity with $t \to 1$. When the fundus has cataract, the refractive medium is turbid. Therefore, the transmission parameter t is reduced, which ranges from 0 to 1. The contrast of the input image is stretched if t is greater than 1.

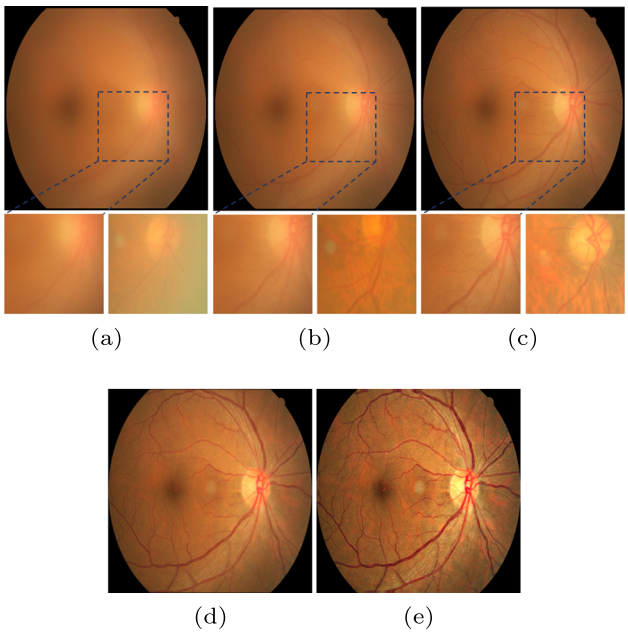

Fig. 4 Synthesized multi-level cataractous fundus images through the degeneration model. **a** Severe cataract, $t = 0.15$. **b** Moderate cataract, $t = 0.25$. **c** Mild cataract, $t = 0.4$. **d** The reference high-quality image. **e** Over-enhanced image, $t = 2$. Synthetic images (bottom left) are compared to another clinical cataractous image with the same blurriness (bottom right) in **a**–**c**

3.2.2 Image synthesis

Initially, we generate 20 degraded retinal images for each high-quality image by adjusting the transmissivity of t, and the step of t is 0.05 and the range is from 0 to 1. These images are then assessed and compared visually with the real cataractous retinal images. Based on the suggestions of two ophthalmologists, we determine to generate the retinal images when the transmissivity t is 0.15, 0.25, and 0.4. The synthesized cataractous images are shown in Fig. 4. Figure 4a–c are synthetic cataractous fundus images of different blurriness. The synthetic image has similar blurriness to the real cataractous image, where the contrast and visibility of their blood vessels are similar. When $t = 0.15$, the severe cataractous fundus image is synthesized. When $t = 0.25$ and $t = 0.4$, moderate cataractous fundus image and mild cataractous fundus image are synthesized respectively. Figure 4e is an over-enhanced fundus image $\boldsymbol{x_{oe}}$, which is synthesized through the model with $t = 2$, to train the classifier C to recognize the over-enhanced images.

3.3 Adaptive enhancement for cataractous retinal images

3.3.1 Adaptive enhancement strategy

The enhancement model can be trained after getting the paired training data $\{\{\boldsymbol{x_b^k}\}_{k=1}^3, \boldsymbol{x_h}\}$. We propose an adaptive enhancement strategy for different cataractous images. Specifically, a classifier for cataract grading and three sets of generative adversarial networks with the same structure are employed in the adaptive enhancement strategy. The enhancement is adjusted according to cataract grading. During training, the blurriness classifier is trained in advance using multi-level cataractous fundus images synthesized through the degradation model to determine the blurriness of input images. Then, three enhancement models of generative adversarial networks are trained independently with synthetic data of different cataract grades to enhance different blurriness. The blurriness in the training set is quantified through the degradation model, so the enhancement ability for each network is guaranteed. In addition, we use the well-trained classifier to calculate the blurriness grading loss $\mathcal{L}_{bg}$ to further constrain the enhancement ability of the model. During test, the blurriness of the test image is predicted, and the enhancement model is selected automatically to achieve adaptive enhancement. Inappropriate enhancement in Fig. 5 cannot be avoided without the assistance of the blurriness classifier C. Our enhancement can standardize the image contrast to a suitable range as high-quality images, as well as avoid performance decrease caused by training jointly with different blurriness, which is discussed in Sect. 4.3.

3.3.2 Objective functions

As shown in Fig. 3, our method has a blurriness classifier C and three sets of generative adversarial networks $\{G_k, D_k\}_{k=1}^3$. The network structure and the training process of the three sets of generative adversarial networks are the same, and only the training data is different. Detailed model structures are shown in Table 2, and objective functions are as follows.

Classifier C: Our blurriness classifier adopts the structure of ResNet-18 [30]. Training the classifier requires cataractous fundus images of three different blurriness $\{\boldsymbol{x_b^k}\}_{k=1}^3$, high-quality fundus images $\boldsymbol{x_h}$, and the over-enhanced fundus image $\boldsymbol{x_{oe}}$. Denote the input image as $\boldsymbol{x_c}$ and their one-hot label $\boldsymbol{o}$, where $\boldsymbol{x_c} \in \{\{\boldsymbol{x_b^k}\}_{k=1}^3, \boldsymbol{x_h}, \boldsymbol{x_{oe}}\}$. The multiclass cross entropy is used as the classification loss:

$$\mathcal{L}_{cls} = -\sum \boldsymbol{o} \cdot \log(C(\boldsymbol{x_c})). \tag{2}$$

After training, the classifier can predict the blurriness $k' \in \{1, 2, ..., 5\}$ of the fundus image and assist in the selection of the enhancement model.

Generator G: Our generator adopts the U-net [31] structure. Skip connections are utilized between the downsampling layers and the corresponding upsampling layers. The loss of the generator consists of three parts: the adversarial loss $\mathcal{L}_{adv}^G$, the content $\mathcal{L}_c$, and the blurriness grading loss $\mathcal{L}_{bg}$.

The adversarial loss of the generator $\mathcal{L}_{adv}^G$ is as follows:

$$\mathcal{L}_{adv}^G = (D_k(\boldsymbol{x_b^k}, G_k(\boldsymbol{x_b^k})) - 1)^2. \tag{3}$$

It facilitates the generator to generate the fundus image as realistic as possible.

The content loss $\mathcal{L}_c$ uses the $\mathcal{L}_1$ loss between the enhanced image $\hat{\boldsymbol{x}}_{\boldsymbol{h}}$ and the high-quality image $\boldsymbol{x_h}$ to constrain the enhanced image as similar as the ground truth. It can be expressed as follows:

$$\mathcal{L}_c = \|\boldsymbol{x_h} - G_k(\boldsymbol{x_b^k})\|_1, \tag{4}$$

where $\|\cdot\|_1$ is the 1-norm.

The blurriness grading loss $\mathcal{L}_{bg}$ aims to constrain the blurriness of the enhanced image. It utilizes the classifier to identify the blurriness of the generated image and

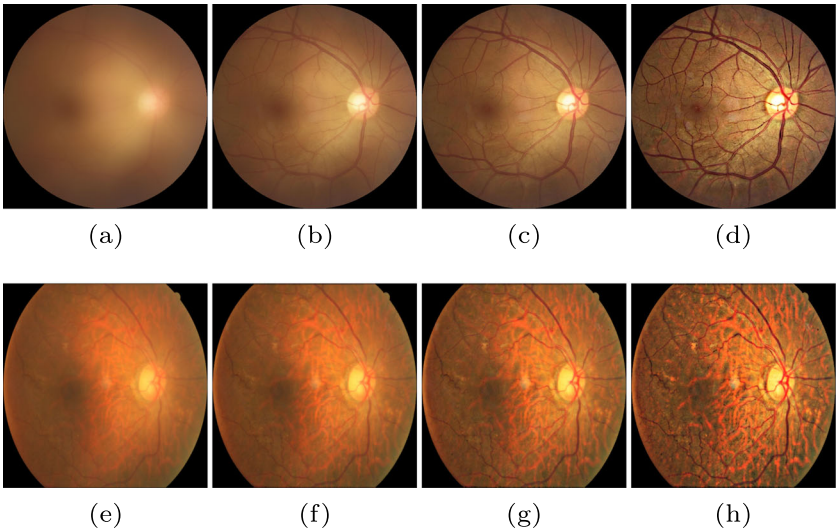

Fig. 5 Inappropriate enhancement. **a** Synthetic cataractous image. **b** Under-enhancement. **c** Enhancement with the matched model. **d** Over-enhancement. **e** Real cataractous image. **f** Under-enhancement. **g** Enhancement with the matched model. **h** Over-enhancement

constrains the blurriness to be close to high-quality images to avoid over-enhancement or under-enhancement.

$$\mathcal{L}_{bg} = -\sum \boldsymbol{o}_h \cdot \log(C(G_k(\boldsymbol{x}_{\boldsymbol{b}}^k))), \tag{5}$$

where $\boldsymbol{o}_h$ is the one-hot label of the high-quality image $\boldsymbol{x}_{\boldsymbol{h}}$.

The total loss of the generator can be expressed as follows:

$$\mathcal{L}_G = \mathcal{L}_{adv}^G + \lambda_c \mathcal{L}_c + \lambda_{bg} \mathcal{L}_{bg}, \tag{6}$$

where λ_c and λ_{bg} are the weights of the losses.

Discriminator D**:** The discriminator is a 5-layer convolutional network. The conditional adversarial network [32] is used, where the conditional image $\boldsymbol{x}_{\boldsymbol{b}}^k$ concatenates with the enhanced image $\hat{\boldsymbol{x}}_{\boldsymbol{h}}$ or the high-quality image $\boldsymbol{x}_{\boldsymbol{h}}$. We express the adversarial loss $\mathcal{L}_{adv}$ separately, where $\mathcal{L}_{adv}^D$ is for the discriminator and $\mathcal{L}_{adv}^G$ is for the generator. The least-square loss is employed for the adversarial loss of the discriminator $\mathcal{L}_{adv}^D$, which can be expressed as follows:

$$\mathcal{L}_{adv}^D = (D_k(\boldsymbol{x}_{\boldsymbol{b}}^k, \hat{\boldsymbol{x}}_{\boldsymbol{h}}))^2 + (D_k(\boldsymbol{x}_{\boldsymbol{b}}^k, \boldsymbol{x}_{\boldsymbol{h}}) - 1)^2. \tag{7}$$

The discriminator is used to distinguish whether the input is from the real fundus or the fundus enhanced by the generator.

Table 2 Model structures

	Operator	Channels
Classifier	ResNet-18 [30]	
	Parameters: 11.179 M	
Generator	Conv4×4 → TransConv4×4, stride 2	64
	Conv4×4 → TransConv4×4, stride 2	128
	Conv4×4 → TransConv4×4, stride 2	256
	Conv4×4 → TransConv4×4, stride 2	512
	Conv4×4 → TransConv4×4, stride 2	512
	Conv4×4 → TransConv4×4, stride 2	512
	Conv4×4 → TransConv4×4, stride 2	512
	Parameters: 41.826 M	
	InstanceNorm, Down: LeakyReLU, Up: ReLU	
Discriminator	Conv4×4, stride 2	64
	Conv4×4, stride 2	128
	Conv4×4, stride 2	256
	Conv4×4, stride 1	512
	Conv4×4, stride 1	1
	Parameters: 2.768 M	
	InstanceNorm, LeakyReLU	

4 Results and discussion

4.1 Implement details

We select 503 high-quality fundus images from the Ocular Disease Intelligent Recognition (ODIR) dataset [33], where images with the label of cataract in the dataset are excluded. In addition, the structure of the fundus should be as clear as possible, including main retinal structures such as optic disc, macula, main blood vessels, and small retinal structures such as microaneurysms, hard exudation, drusens, and small blood vessels after two bifurcations. The high-quality images are divided into *Synthetic Training Set* and *Synthetic Test Set* at a ratio of 2:1 (336:167), and then they are processed using the degeneration model M. *Clinical Test Set* contains 406 fundus images. Blurred fundus images are not pre-selected, so some high-quality fundus images are mixed in the test set. The in-house clinical data used in our experiments is from the Beijing Eye Study collected by Beijing Tongren Hospital. The Beijing Eye Study is a population-based cross-sectional investigation that was performed in an urban district and a rural region in Greater Beijing. The Ethics Committee of the Beijing Tongren Hospital approved the study protocol, and all participants gave their written informed consent. Images are resized to 512×512 in our experiments.

The batch size of the classifier C is set as 24 and 15 epochs are trained. The learning rate is $1e^{-3}$ and decays twice at the 5th and 10th epochs. The enhancement networks use a batch size of 1 and a learning rate of $2e^{-4}$. λ_c and λ_{bg} are set as 200 and 0.2 respectively. The networks are trained for 100 epochs. Random brightness augmentation following the settings in cofe-Net [9] is utilized during the training of both classifier and enhancement networks considering that the brightness may be variable for clinically acquired fundus images.

Quantitative evaluation metrics are selected based on whether the test set has reference images. PSNR [34], SSIM [35], and FSIM [36] are used for *Synthetic Test Set* as full-reference evaluation metrics. For *Clinical Test Set*, Perception-based Image Quality Evaluator (PIQE) [37], Frechet Inception Distance (FID) [38], and cataract grading results are leveraged. The standard deviation (STD) along with the mean value has been calculated for quantitative evaluation. We utilize the well-trained classifier C to evaluate the blurriness level of the enhanced images and observe whether they are under-enhanced or over-enhanced. P_h represents the probability that the enhanced image is recognized as high quality by the classifier C, and P_{oe} indicates the probability of over-enhancement. The calculation is as follows:

$$P_h = \frac{N_h}{N}, \tag{8}$$

$$P_{oe} = \frac{N_{oe}}{N}, \tag{9}$$

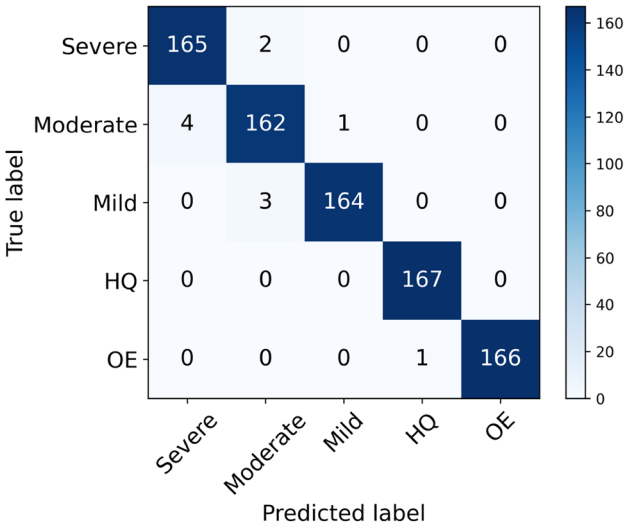

Fig. 6 Confusion matrix of the classification results on *Synthetic Test Set*. Severe, moderate, and mild represent the grades of cataract. HQ and OE are high quality and over-enhancement respectively

where N denotes the number of input image. N_h and N_{oe} are the number classified as high quality and over-enhancement respectively.

4.2 Ablation study

We use *Synthetic Test Set* in our ablation study. The classification results are shown in Fig. 6. The classification accuracy of the blurriness reached 98.68%. In the case of misclassification, they are only classified into adjacent categories, which has less impact than other wrong categories.

The ablation experiments of enhancement networks are shown in Table 3. We randomly select the enhancement networks when we test without the assistance of the classifier C. The inappropriate enhancement (over-enhancement, under-enhancement) leads to a large drop in results. The proposed adaptive enhancement strategy can grade and enhance cataractous fundus images with the help of the classifier C, which is shown in the second row in Table 3. It has the best SSIM score on moderate cataract. After adding $\mathcal{L}_{bg}$, the enhancement results are further improved. The average value of SSIM reaches 0.9297.

Table 3 Ablation experiments

C	$\mathcal{L}_{bg}$	Severe	Moderate	Mild	Average
×	×	0.9007	0.9124	0.8849	0.8994
√	×	0.9039	**0.9315**	0.9403	0.9252
√	√	**0.9116**	0.9295	**0.9480**	**0.9297**

SSIM on *Synthetic Test Set* is shown in the table
Bold numbers in the table show the best scores

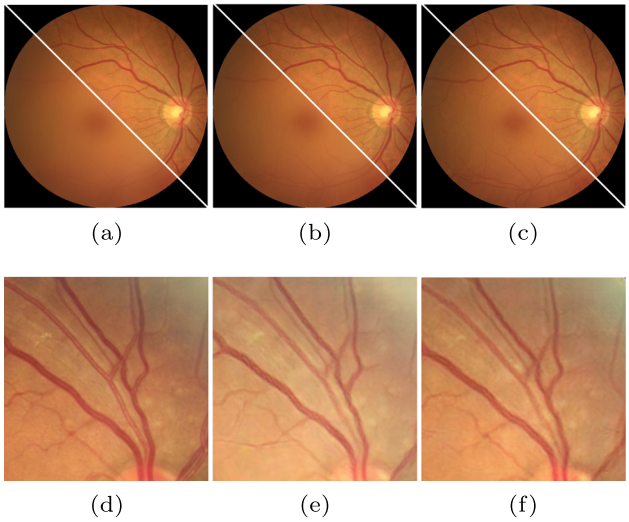

Fig. 7 Enhancement results of ablation study. **a** Synthetic severe cataract and the enhancement. **b** Moderate cataract. **c** Mild cataract. **d** High-quality image. **e** Enhancement without $\mathcal{L}_{bg}$. **f** Enhancement with $\mathcal{L}_{bg}$

Visual comparison of enhancement results of different cataract is shown in Fig. 7a–c. Different grades of cataract all achieve good enhancement results, and there is no obvious difference between the enhanced results. Our method can unify cataractous images with large differences of blurriness into a suitable and observable contrast as high-quality images. Figure 7d–f display the impact of $\mathcal{L}_{bg}$. Figure 7e is a little whiter than the other two images without the help of $\mathcal{L}_{bg}$.

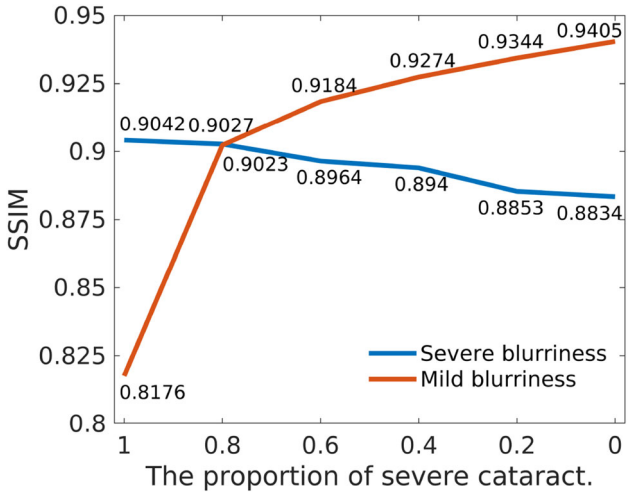

Fig. 8 Enhancement results for the joint training strategy. We jointly train an enhancement network with images of severe cataract and mild cataract. The horizontal axis represents the sampling proportion of severe blur in the training data without changing the sampling range. With the proportion of severe cataract data decreasing during training, the enhancement ability for severe cataractous images decreases in the test phase

4.3 Influence of training jointly

We also investigate the influence of training jointly. The baseline generative adversarial networks without the classifier are leveraged in the experiments. By adjusting the proportion of mild cataract and severe cataract in *Synthetic Training Set*, multiple enhanced networks are trained. The test results on *Synthetic Test Set* are shown in Fig. 8. The best result for severe cataract is to train with only the severe cataractous images. As the proportion of mild cataract increases in the training data, the enhancement accuracy reduces for the severe cataract, although the network has the ability to enhance different blurriness. The experiments show that different blurriness affects each other when they are trained together, which will lead to a decrease in the accuracy of the results. Our adaptive enhancement can avoid the interaction of training jointly.

4.4 Comparison with other methods

4.4.1 Comparison on synthetic data

We first perform the comparison on *Synthetic Test Set*. We compare classical retinal image enhancement methods: Zhou et al. [11], Xiong et al. [14], Gupta et al. [12], Cao et al. [6], Zhang et al. [13], unpaired image translation methods: You et al. [7], Zhao et al. [39], Yang et al. [20], and methods with data degradation: cofe-Net [9], ArcNet [22]. The deep learning methods use the well-trained weights provided by the authors or they are retrained with refined parameters if the weights are not available.

The results of the quantitative evaluation are shown in Table 4. Our method performs the best on *Synthetic Test Set* with 27.92 PSNR, 0.9297 SSIM, and 0.9858 FSIM with the minimum standard deviation value $2.5e^{-4}$. The STD values of PSNR and SSIM are 4.9 and $3.6e^{-2}$, which does not perform as well as other methods, such as the classical method Cao et al. [6] with STD of 1.4 and $2.8e^{-2}$. Our method uses data augmentation on brightness during training following the settings in cofe-Net [9] to adjust the inappropriate intensity, while Cao et al. [6] design specific steps to retain the original color. The adjusted brightness is easier to observe for images with imperfect illumination, but it will increase the standard deviation for PSNR and SSIM. If we remove the data augmentation, the STD of PSNR and SSIM will drop to 1.6 and $8.5e^{-3}$. But we retain the data augmentation on brightness considering the improvement on visual assessment and the variable brightness in clinical photography.

The inappropriate enhancement may lead to the decline of the enhancement visually and quantitatively. As shown in Fig. 9, our method can enlarge the contrast of the image to a suitable range on the synthetic dataset. After the enhancement, the structure of blood vessels and optic discs in the

Table 4 Quantitative comparison with other methods on synthetic and clinical datasets

	Synthetic Test Set			*Clinical Test Set A*		*Clinical Test Set B*	
	PSNR ↑	SSIM ↑	FSIM ↑	PIQE ↓ [37]	P_h ↑	FID ↓ [38]	P_{oe} ↓
Zhou et al. [11]	$20.13_{1.7}$	$0.9015_{3.1e^{-2}}$	$0.9237_{9.1e^{-4}}$	$14.96_{4.2}$	0.6333	125.0	0.6838
Xiong et al. [14]	$26.32_{1.5}$	$0.7872_{5.4e^{-2}}$	$0.8601_{6.9e^{-4}}$	$16.44_{5.3}$	0.3370	72.04	1
Gupta et al. [12]	$21.38_{1.6}$	$0.8767_{2.8e^{-2}}$	$0.8441_{8.6e^{-4}}$	$14.88_{4.3}$	0.6556	61.43	0.8456
Cao et al. [6]	$25.72_{1.4}$	$0.9172_{2.8e^{-2}}$	$0.9206_{3.2e^{-3}}$	$14.73_{4.3}$	0.1815	103.7	1
Zhang et al. [13]	$19.52_{1.7}$	$0.8577_{3.4e^{-2}}$	$0.8383_{7.7e^{-4}}$	$16.10_{6.1}$	0.6296	110.8	0.7941
You et al. [7]	$25.54_{3.9}$	$0.8226_{1.1e^{-1}}$	$0.9259_{4.3e^{-4}}$	$16.92_{6.1}$	0.4889	37.29	0.1103
Zhao et al. [39]	$27.48_{2.9}$	$0.8884_{3.1e^{-2}}$	$0.9276_{7.0e^{-4}}$	$14.34_{4.1}$	0.6852	39.65	0
Yang et al. [20]	$24.84_{2.1}$	$0.8972_{2.7e^{-2}}$	$0.9353_{1.9e^{-3}}$	$16.06_{3.0}$	0.7148	36.18	0
cofe-Net [9]	$22.96_{3.0}$	$0.8979_{4.4e^{-2}}$	$0.9173_{1.4e^{-3}}$	$22.33_{4.8}$	0.4000	27.61	0
ArcNet [22]	$20.34_{1.3}$	$0.8966_{3.3e^{-2}}$	$0.9089_{6.2e^{-3}}$	$19.39_{3.8}$	0.6185	41.33	0.3456
Ours	$\mathbf{27.92}_{5.5}$	$\mathbf{0.9297}_{3.6e^{-2}}$	$\mathbf{0.9858}_{2.5e^{-4}}$	$\mathbf{13.48}_{3.5}$	**0.7926**	**18.74**	**0**

The standard deviation is displayed after the mean value
Bold numbers in the table show the best scores

image becomes clearer, which is very close to the ground truth high-quality image.

4.4.2 Comparison on clinical data

Clinical Test Set is clinically collected data, where images of different quality are collected at the same time during clinical photography. We use the classifier to identify high-quality images apart from images that need to be enhanced. The classification results are shown in Table 5. *Clinical Test Set A* and *Clinical Test Set B* are tested and discussed separately.

Clinical Test Set A contains 270 images that need to be enhanced. The comparison results are displayed in Table 4. Our method performs the best on the PIQE score of 13.48 with the second small standard deviation value of 3.5. Furthermore, we evaluate the enhanced results with the blurriness classifier C. The large value of P_h represents the contrast after enhancement is close to the high-quality image. Our enhancement has the best P_h score of 0.7926 and improves by 7.78% compared with the second. It means that the proposed method has the appropriate enhancement ability, which can help to avoid over-enhancement and under-enhancement.

The visual comparison on *Clinical Test Set A* of different methods is shown in the second row of Fig. 9. Figure 9g looks very blurry that needs to be enhanced. Structures such

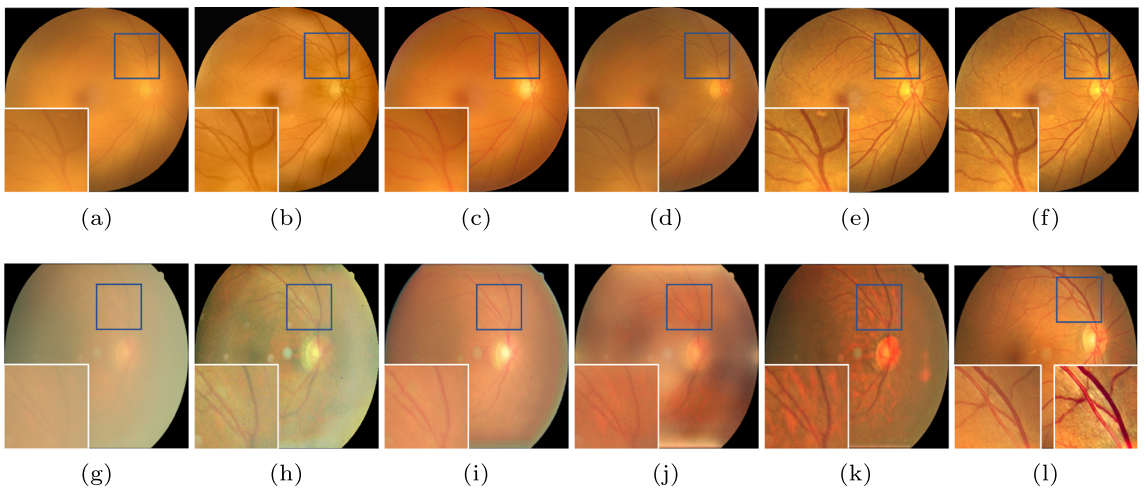

Fig. 9 Visual comparison with other methods on synthetic and clinical datasets. **a** Synthetic moderate cataract. **b** Zhou et al. [11]. **c** Zhao et al. [39]. **d** cofe-Net [9]. **e** Ours. **f** Ground truth high-quality image. **g** Clinical blurry image classified by C. **h** Cao et al. [6]. **i** Yang et al. [20]. **j** ArcNet [22]. **k** Ours. **l** High-quality image from *Clinical Test Set B* classified by C, where the bottom right is the over-enhanced image of Xiong et al. [14]. **b** and **h** are classical retinal image enhancement methods. **c** and **i** are unpaired image translation methods. **d** and **j** are methods with data degradation. The enlarged part of the image is displayed in the left bottom of the image

Table 5 Classification on *Clinical Test Set* using the well-trained classifier

Clinical Test Set A			*Clinical Test Set B*	
Severe	Moderate	Mild	HQ	OE
9	41	220	136	0

as blood vessels are almost invisible, which is challenging for fundus enhancement algorithms. Other methods are somewhat under-enhanced for such severe blurriness. The blurriness of this image is evaluated as severe by our blurriness classifier. The proposed method has an enhancement network for severe blurriness. The fundus structure is well restored after enhancement. Our method can enhance the image contrast to a suitable range similar to high-quality images according to the blurriness.

Images in *Clinical Test Set B* are classified as high quality. Our method does not additionally process these fundus images. The comparison results are shown in Table 4. We use another dataset containing 200 high-quality fundus images as the reference dataset for FID. In the experiments, the FID of the classical enhancement methods is generally worse than that of the deep learning methods, which may be caused by over-enhancement. As shown by P_{oe} in the table, the over-enhancement ratio of the classical methods is also higher. Four retinal image enhancement methods including our method do not have over-enhanced images, and their FID values are also smaller. Figure 9l is the high-quality image of *Clinical Test Set B* identified by the blurriness classifier C. The image contrast and brightness do not need to be adjusted again. The redundancy enhancement may lead to over-enhancement as shown in the bottom right of the image.

We also test clinical image pairs before and after cataract surgery from Beijing Tongren Hospital for visual assessment. The visibility of the fundus structure is improved after cataract surgery, which can be regarded as the ground truth of the low-quality image before surgery. As shown in Fig. 10, the visibility of the fundus structure before surgery decreases as blurriness increases. After enhancement, the fundus contrast is corrected to a normal range as high-quality images and easy to observe. Some enhanced vessels are even clearer than after surgery, such as the case of severe blurriness in the third row. The enhanced images are all classified as high quality by our classifier. The experiments on the cataract surgery dataset demonstrate that our method can suitably enhance the fundus with different blurriness.

4.4.3 Statistical test

We conduct a statistical analysis for the experiments. First, the Shapiro-Wilk test [40] is employed for the SSIM of different enhancement results on the *Synthetic Test Set*. The

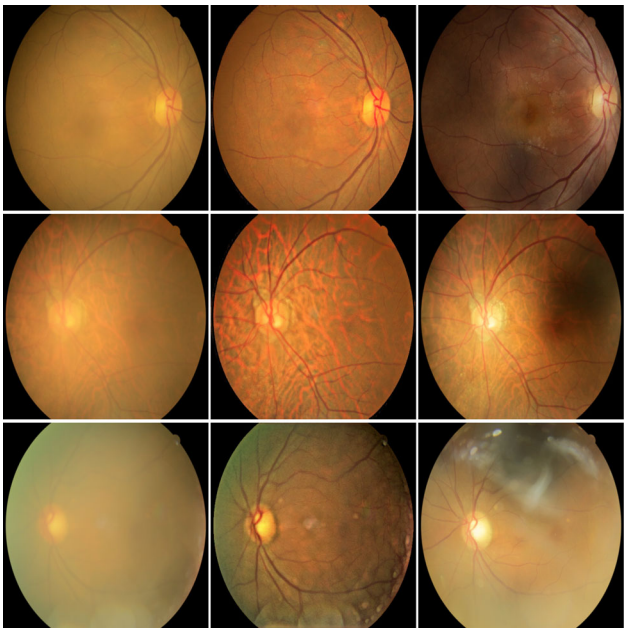

Fig. 10 Visual comparison of clinical image pairs before and after cataract surgery. **a** Cataractous images before surgery. **b** Enhancement via our method. **c** Ground truth after surgery. The first, second, and third rows are mild, moderate, and severe blurriness respectively recognized by our classifier

results of the p-value are all less than 0.05, indicating the SSIM results are not Gaussian distribution. Then, we carry out the Wilcoxon signed-rank test [41], a nonparametric statistical hypothesis test of matched-pair data, for the statistical analysis of different methods. The enhancement of cofe-Net [9] and ArcNet [22] are tested along with the ablation experiments because they are based on image degradation, which has a similar process to ours. In Table 6, the p-values of the statistical test are less than 0.05, which demonstrates our method is statistically different from others.

4.5 Applications

This method can be applied as preprocessing for clinical diagnosis or related algorithms. As shown in Fig. 11a–c, low contrast of the image affects the accuracy of the blood vessel tracking algorithm on clinically fundus images. The tracking results are corrected after enhancement by our method.

Table 6 Wilcoxon signed-rank test for SSIM on the *Synthetic Test Set*

	p-value
Baseline	$2.54e^{-53}$
Baseline + C	$1.76e^{-17}$
cofe-Net [9]	$9.12e^{-66}$
ArcNet [22]	$1.56e^{-65}$

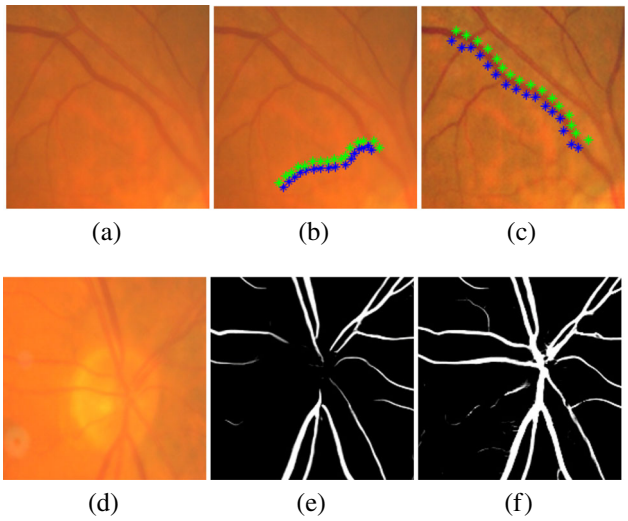

Fig. 11 Vessel tracking and segmentation of the enhancement. **a** Blurry image. **b** Vessel tracking before enhancement. **c** Vessel tracking after enhancement. **d** Blurry image. **e** Vessel segmentation before enhancement. **f** Vessel segmentation after enhancement

Our method can also improve vessel segmentation, which is shown in Fig. 11d–f.

We propose an adaptive retinal image enhancement method in this paper. Image quality of cataractous images is improved, and inappropriate enhancement is avoided after the enhancement. Our approach benefits clinical diagnosis via enhancing cataractous retinal images to the same quality as normal images and improves the performance of computer-aided algorithms such as vessel tracking and vessel segmentation.

5 Conclusion

We propose an adaptive cataractous retinal image enhancement method in this paper. Our method can automatically enhance the fundus to the appropriate contrast with the help of cataract grading. A degradation model is proposed to degenerate high-quality fundus images into different grades of cataract with quantified contrast. An adaptive enhancement strategy is developed for contrast standardization. Our method keeps the contrast of enhanced images similar to normal retinal images so that the adverse impact caused by cataract can be removed for clinical diagnosis.

The contrast of some images is extremely small, and all retinal structures cannot be seen, where the blurriness exceeds the severe cataract investigated in this paper. We cannot enhance these images well at present, which is one of the limitations of the proposed work. In addition, the low-quality fundus images in clinics have different forms of deterioration, such as uneven brightness and out-of-focus. Our future work will concentrate on other image deteriorations besides low contrast caused by cataract and incorporate more efficient network structures.

Funding This work was supported by the National Natural Science Foundation of China (No. 82072007), the Key Scientific Research Projects of Colleges and Universities in Henan Province (No. 23A520011), and Key R&D and Promotion Projects of Henan Province (No. 232102211089).

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Zhang J, Li H, Nie Q, Cheng L (2014) A retinal vessel boundary tracking method based on Bayesian theory and multi-scale line detection. Comput Med Imaging Graph 38(6):517–525
2. Cao L, Li H, Zhang Y, Zhang L, Xu L (2020) Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. Information Fusion 53:196–208
3. Staal J, Abrámoff MD, Niemeijer M, Viergever MA, Van Ginneken B (2004) Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging 23(4):501–509
4. Yang J-J, Li J, Shen R, Zeng Y, He J, Bi J, Li Y, Zhang Q, Peng L, Wang Q (2016) Exploiting ensemble learning for automatic cataract detection and grading. Comput Methods Prog Biomed 124:45–57
5. Setiawan, AW, Mengko, TR, Santoso, OS, Suksmono, AB (2013) Color retinal image enhancement using CLAHE. In: International conference on ICT for Smart Society, pp 1–3. IEEE
6. Cao L, Li H (2020) Enhancement of blurry retinal image based on non-uniform contrast stretching and intensity transfer. Medical & Biological Engineering & Computing 58(3):483–496
7. You, Q, Wan, C, Sun, J, Shen, J, Ye, H, Yu, Q (2019) Fundus image enhancement method based on CycleGAN. In: 2019 41st Annual international conference of the ieee engineering in medicine and biology society (EMBC):pp 4500–4503. IEEE
8. Cheng, P, Lin, L, Huang, Y, He, H, Luo, W, Tang, X (2023) Learning enhancement from degradation: a diffusion model for fundus image enhancement. arXiv:2303.04603
9. Shen Z, Fu H, Shen J, Shao L (2020) Modeling and enhancing low-quality retinal fundus images. IEEE Trans Med Imaging 40(3):996–1006
10. Deng Z, Cai Y, Chen L, Gong Z, Bao Q, Yao X, Fang D, Yang W, Zhang S, Ma L (2022) RFormer: transformer-based generative adversarial network for real fundus image restoration on a new clinical benchmark. IEEE Journal of Biomedical and Health Informatics 26(9):4645–4655
11. Zhou M, Jin K, Wang S, Ye J, Qian D (2017) Color retinal image enhancement based on luminosity and contrast adjustment. IEEE Trans Biomed Eng 65(3):521–527
12. Gupta B, Tiwari M (2019) Color retinal image enhancement using luminosity and quantile based contrast enhancement. Multidim Syst Sign Process 30(4):1829–1837
13. Zhang S, Webers CA, Berendschot TT (2022) A double-pass fundus reflection model for efficient single retinal image enhancement. Signal Process 192:108400
14. Xiong L, Li H, Xu L (2017) An enhancement method for color retinal images based on image formation model. Comput Methods Prog Biomed 143:137–150
15. Gaudio, A, Smailagic, A, Campilho, A (2020) Enhancement of retinal fundus images via pixel color amplification. In: Image anal-

ysis and recognition: 17th international conference, ICIAR 2020, Póvoa de Varzim, Portugal, June 24–26, 2020, Proceedings, Part II 17, pp 299–312. Springer
16. He K, Sun J, Tang X (2010) Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(12):2341–2353
17. Cao L, Li H, Zhang Y (2020) Retinal image enhancement using low-pass filtering and α-rooting. Signal Process 170:107445
18. Zhu, J-Y, Park, T, Isola, P, Efros, AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pp 2223–2232
19. Wan C, Zhou X, You Q, Sun J, Shen J, Zhu S, Jiang Q, Yang W (2022) Retinal image enhancement using cycle-constraint adversarial network. Frontiers in Medicine 8:2881
20. Yang B, Zhao H, Cao L, Liu H, Wang N, Li H (2023) Retinal image enhancement with artifact reduction and structure retention. Pattern Recogn 133:108968
21. Luo Y, Chen K, Liu L, Liu J, Mao J, Ke G, Sun M (2020) Dehaze of cataractous retinal images using an unpaired generative adversarial network. IEEE Journal of Biomedical and Health Informatics 24(12):3374–3383
22. Li, H, Liu, H, Hu, Y, Fu, H, Zhao, Y, Miao, H, Liu, J (2022) An annotation-free restoration network for cataractous fundus images. IEEE Transactions on Medical Imaging
23. Masruroh, SU, Rahman, DA, Putri, RA (2022) Systematic literature review: detecting cataract with deep learning. In: 2022 10th International conference on Cyber and IT service management (CITSM):pp 01–05. IEEE
24. Süleyman E (2020) Medical data analysis for different data types. International Journal of Computational and Experimental Science and Engineering 6(3):138–144
25. Zhou Y, Li G, Li H (2019) Automatic cataract classification using deep neural network with discrete state transition. IEEE Trans Med Imaging 39(2):436–446
26. Xu X, Li J, Guan Y, Zhao L, Zhao Q, Zhang L, Li L (2021) GLA-Net: a global-local attention network for automatic cataract classification. J Biomed Inform 124
27. Keenan TD, Chen Q, Agrón E, Tham Y-C, Goh JHL, Lei X, Ng YP, Liu Y, Xu X, Cheng C-Y et al (2022) DeepLensNet: deep learning automated diagnosis and quantitative classification of cataract type and severity. Ophthalmology 129(5):571–584
28. Son KY, Ko J, Kim E, Lee SY, Kim M-J, Han J, Shin E, Chung T-Y, Lim DH (2022) Deep learning-based cataract detection and grading from slit-lamp and retro-illumination photographs: model development and validation study. Ophthalmology Science 2(2)
29. Zhang, Y, Ding, L, Sharma, G (2017) HazeRD: an outdoor scene dataset and benchmark for single image dehazing. In: 2017 IEEE international conference on image processing (ICIP). IEEE, pp 3205–3209
30. He, K, Zhang, X, Ren, S, Sun, J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
31. Ronneberger, O, Fischer, P, Brox, T (2015) U-Net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, pp 234–241. Springer
32. Isola, P, Zhu, J-Y, Zhou, T, Efros, AA (2017) Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1125–1134
33. Ocular disease intelligent recognition ODIR-5K
34. Hore, A, Ziou, D (2010) Image quality metrics: PSNR vs. SSIM. In: 2010 20th International conference on pattern recognition, pp 2366–2369. IEEE
35. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
36. Zhang L, Zhang L, Mou X, Zhang D (2011) FSIM: a feature similarity index for image quality assessment. IEEE Trans Image Process 20(8):2378–2386
37. Venkatanath, N, Praneeth, D, Bh, MC, Channappayya, SS, Medasani, SS (2015) Blind image quality evaluation using perception based features. In: 2015 Twenty first national conference on communications (NCC):pp 1–6. IEEE
38. Heusel, M, Ramsauer, H, Unterthiner, T, Nessler, B, Hochreiter, S (2017) GANs trained by a two time-scale update rule converge to a local Nash equilibrium. Advances in Neural Information Processing SSystems **30**
39. Zhao, H, Yang, B, Cao, L, Li, H (2019) Data-driven enhancement of blurry retinal images via generative adversarial networks. In: International conference on medical image computing and computer-assisted intervention, pp 75–83. Springer
40. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3/4):591–611
41. Woolson, RF (2007) Wilcoxon signed-rank test. Wiley encyclopedia of clinical trials, pp 1–3

Bingyu Yang is a Ph.D. candidate at the School of Information and Electronics, Beijing Institute of Technology, Beijing, China.

Lvchen Cao received the Ph.D. degree from Beijing Institute of Technology, Beijing, China, in 2021. His research interests are classification and enhancement of low-quality medical images.

He Zhao received the Ph.D. degree from Beijing Institute of Technology, Beijing, China, in 2020. His research interests are medical image synthesis and disease detection.

Huiqi Li received Ph.D. degree from Nanyang Technological University, Singapore. She is currently a professor at Beijing Institute of Technology. Her research interests are medical image processing and computer-aided diagnosis.

Hanruo Liu is an associate professor of ophthalmology at Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Capital Medical University. Her research interests are digital eye health, diagnosis and treatment of glaucoma.

Ningli Wang is the Director of Beijing Tongren Eye Center, Dean of School of Ophthalmology, Capital Medical University, Head of National Committee for the Prevention of Blindness. His main areas of interests are the pathogenesis, diagnosis and treatment of glaucoma.