



MUTE: A multilevel-stimulated denoising strategy for single cataractous retinal image dehazing

Shuhe Zhang^{*}, Ashwin Mohan, Carroll A.B. Webers, Tos T.J.M. Berendschot

University Eye Clinic Maastricht, Maastricht University Medical Center +, P.O. Box 5800, Maastricht 6202 AZ, The Netherlands



ARTICLE INFO

Keywords:
Ophthalmology
Retinal image
Cataract
Dehazing

ABSTRACT

In this research, we studied the duality between cataractous retinal image dehazing and image denoising and proposed that the dehazing task for cataractous retinal images can be achieved with the combination of image denoising and sigmoid function. To do so, we introduce the double-pass fundus reflection model in the YPbPr color space and developed a multilevel stimulated denoising strategy termed MUTE. The transmission matrix of the cataract layer is expressed as the superposition of denoised raw images of different levels weighted by pixel-wise sigmoid functions. We further designed an intensity-based cost function that can guide the updating of the model parameters. They are updated by gradient descent with adaptive momentum estimation, which gives us the final refined transmission matrix of the cataract layer. We tested our methods on cataract retinal images from both public and proprietary databases, and we compared the performance of our method with other state-of-the-art enhancement methods. Both visual assessments and objective assessments show the superiority of the proposed method. We further demonstrated three potential applications including blood vessel segmentation, retinal image registrations, and diagnosing with enhanced images that may largely benefit from our proposed methods.

1. Introduction

Retinal imaging provides a noninvasive measurement of fundus morphology and is widely used by ophthalmologists for early detection, diagnosis, and monitoring of ocular diseases and their progression (Ao et al., 2019; Doubal et al., 2010; Mary et al., 2016; Siegrist et al., 2014; Zafar et al., 2019; Zhang et al., 2018; Zhu et al., 2014). However, imaging of cataract patients' retina using fundus cameras is challenging since the image quality is severely degraded by light scattering of the turbid cataract layers. Dependent on the severity, many of the subtle retinal structures are hidden by the haze effect of cataracts which largely limits the diagnosing accuracy and reliability for ophthalmologists, as a result, diagnosing according to cataractous retinal images is error-prone.

Improving the cataractous retinal image quality by modifying the optical system of a fundus camera, for example, using a laser source or confocal imaging technique is expensive and also inconvenient for quick and universal retinal image checking. Therefore, image processing methods draw more attention to cataractous retinal image enhancement as they are economical and practical (Cao and Li, 2020; Luo et al., 2020; Mitra et al., 2018; Peli and Peli, 1989; Xiong et al., 2017). Various algorithms were investigated for cataractous retinal image enhancement,

within which the image formation model-based methods show their priority in experimental results.

Given the similarity between cataractous retinal images and hazy images for natural scenes, many studies have adopted natural scene dehazing methods to achieve cataractous retinal image dehazing (Cao and Li, 2021; Gaudio et al., 2020; Peli and Peli, 1989; Xiong et al., 2017), however, their results are limited. The main reason is that the spectral properties of retinal images are different from nature scenes due to the existence of intraocular lens, and absorption by ocular tissue (Berendschot et al., 2003). Therefore, the natural scene dehazing methods suffer from the problem where the transmission matrix of cataract layers is underestimated (Supplementary Note 1).

In our previous study (Zhang et al., 2022) we reviewed the historical image formation model of retinal imaging and developed the double-pass fundus reflection model (DPFR) dedicated to retinal image enhancement. It incorporates a specific double-pass fundus reflection feature (Artal et al., 1995a, 1995b; Berendschot et al., 2003) that had been neglected hitherto in modeling the light propagation of fundus imaging in all reports on retinal image enhancement. In this study, we continued this research and show that the dehazing of retinal images can be achieved in a framework of image denoising or smoothing. We

* Corresponding author.

E-mail address: shuhe.zhang@maastrichtuniversity.nl (S. Zhang).

developed a multilevel-stimulated denoising strategy, called MUTE. In MUTE, the transmission matrix of the cataract layer is expressed as the superposition of denoised raw retinal images of different levels weighted by pixel-wise sigmoid functions. We further designed an intensity-based cost function that guides and stimulates the updating of the transmission matrix. A demonstration of the experimental results of MUTE on cataractous retinal image dehazing is shown in Fig. 1.

The subsequent manuscript is organized as follows. Section 2 briefly reviews related work. Section 3 introduces the basic formulas for our retinal image formation model and the preprocessing procedures that are applied to the retinal image before the retinal image dehazing. Section 4 presents and analyzes the proposed multilevel-stimulated denoising (MUTE) strategy in detail. Section 5 shows the experimental results for both visual and objective assessment in comparison to state-of-the-art (SOTA) algorithms. Demonstrations of some potential clinical applications of MUTE are also included in Section 5. Section 6 contains a detailed ablation study of each module of MUTE. Section 7 provides a Hessian denoising strategy to suppress the noise signals during the MUTE dehazing, and Section 8 is discussions and the concluding remarks.

2. Related works

2.1. Retinal imaging formation/structuring model

The early-stage of the retinal image formation model was directly adopted from Koschmieder & McCartney's model (Koschmieder, 1938; McCartney, 1976) of hazy nature scenes, which is given by

$$\mathbf{S}_c(\mathbf{r}) = \mathbf{T}(\mathbf{r}) \circ \mathbf{O}_c(\mathbf{r}) + A_c[1 - \mathbf{T}(\mathbf{r})], \quad (1)$$

where \mathbf{O} is the haze-free image, \mathbf{T} is the transmission matrix of the haze medium describing the portion of the light that is not scattered and reaches the camera. A is the global atmospheric light, \mathbf{S} is the observed image. $\mathbf{r} = (x, y)$ is the vector of spatial coordinates and $c \in \{R, G, B\}$ is the index of the color channels. \circ is the pixel-wise multiplication.

Based on Eq. (1), many studies have been proposed to enhance the visual quality of hazy and underwater images, achieving promising results (Mujbaile and Rojatkar, 2020; Saxena and Bhaduria, 2020; Senthilkumar and Sivakumar, 2019; Singh and Kumar, 2018). Some studies also introduced Eq. (1) for retinal image enhancement. In these applications, \mathbf{T} was regarded as the transmission matrix of the cataract layer, as cataract also scatters light and imposes a haze-effect on captured retinal images. Xiong et al. proposed to use intensity correction and histogram adjustment to preprocess the image, generating a transmission map according to the intensity of the preprocessed image in each color channel (Xiong et al., 2017). Haze effects were then suppressed through dehazing. Gaudio et al. demonstrated a pixel color amplification method for retinal imaging enhancement that showed good performance in enhancing the detailed structure of retinal images (Gaudio et al., 2020). Cao et al. proposed the detail-richest-channel to estimate the transmission matrix \mathbf{T} (Cao and Li, 2021). However, Eq. (1) was developed for natural scenes and may not be the optimal choice for fundus imaging, ignoring its double-pass property.

Peli et al. (Peli and Peli, 1989) developed an optical model for imaging the retina through cataracts which is

$$\mathbf{S}_c(\mathbf{r}) = \alpha L \circ \mathbf{T}(\mathbf{r}) \circ \mathbf{O}_c(\mathbf{r}) + L \circ [1 - \mathbf{T}(\mathbf{r})], \quad (2)$$

L is the flash illumination of the fundus camera and α is the attenuation of retinal illumination due to the cataract. Both L and α are considered to be constant. Following Eq. (2), many methods have been proposed to enhance the cataractous retinal image. Mitra et al. proposed to use the Retinex theory to remove the haze effect (Mitra et al., 2018), which is analogous to Peli's work. Luo et al. developed an unpaired generative adversarial network to achieve haze removal (Luo et al., 2020).

Different from Eq. (1), Eq. (2) reveals that the illumination pattern also impacts the quality of retinal imaging. However, as L is constant, Eq. (2) loses the ability to correct the uneven (spatially varying) illumination of retinal imaging. In addition, the existing parameter α shows the basic idea of the double-pass property where the illumination light

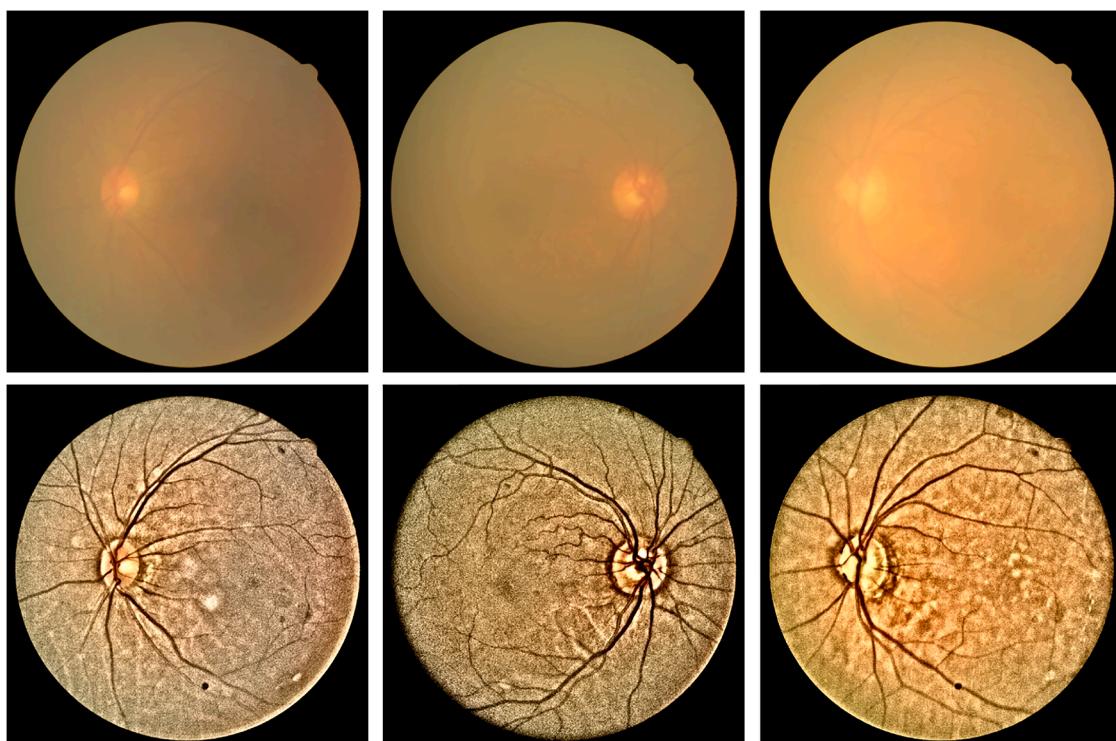


Fig. 1. Demonstrations of MUTE on cataractous retinal image dehazing. Images in the first row are cataractous retinal images. Images in the second row are outputs of MUTE.

interacts twice with the cataract layer. (When the light goes inside the eye, and when it is reflected out from the fundus).

In our previous study (Zhang et al., 2022), we proposed the double-pass fundus reflection (DPFR) that deals with image formation in retinal imaging. This DPFR model is given by

$$\mathbf{S}_c(\mathbf{r}) = \mathbf{L}(\mathbf{r}) \circ \mathbf{T}_{\text{lens}}^2(\mathbf{r}) \circ [\mathbf{T}_{sc}^2(\mathbf{r}) \circ \mathbf{O}_c(\mathbf{r}) + 1 - \mathbf{T}_{sc}(\mathbf{r})], \quad (3)$$

where $\mathbf{L}(\mathbf{r})$ is the illumination from the outside of the eye and is delivered by the illumination system of the fundus camera. We consider that the retina is illuminated by white light (identical value in R, G and B channels) which may have an uneven and insufficient illumination pattern. $\mathbf{T}_{\text{lens}}(\mathbf{r})$ is the transmission matrix of the combination of cornea and lens. $\mathbf{T}_{sc}(\mathbf{r})$ is the transmission matrix of intraocular scatter including cataractous. Eq. (3) reveals that the degeneration of the retinal image is mainly due to three parts: (1) an uneven illumination condition, (2) filtering by the human lens, and (3) intraocular scattering. Based on Eq. (3), we now can correct the illumination problem and suppress the haze effect.

Besides the image formation model, there are also image structure models used for retinal image enhancement (Cao and Li, 2020; Cao et al., 2020; Dai et al., 2016; Wang et al., 2021), and the image structure models can be summarized as

$$\mathbf{S}_c(\mathbf{r}) = \mathbf{S}_{c, \text{background}}(\mathbf{r}) + \mathbf{S}_{c, \text{details}}(\mathbf{r}), \quad (4)$$

where $\mathbf{S}_{\text{background}}$ is the background information of the observed image which corresponds to the low-frequency components, while $\mathbf{S}_{\text{details}}$ denotes the detailed information implying the detailed structures and textures of the image. By giving a large weight to $\mathbf{S}_{\text{details}}$ and suppressing the $\mathbf{S}_{\text{background}}$ one can obtain a contrast-enhanced image. $\mathbf{S}_{\text{background}}$ can be obtained by low-pass filtering of \mathbf{S} (Cao and Li, 2020; Cao et al., 2020) and total variation regularization (Wang et al., 2021). While $\mathbf{S}_{\text{details}}$ can be obtained by high pass filtering of \mathbf{S} or subtracting $\mathbf{S}_{\text{background}}$ from \mathbf{S} . Note that Eq. (4) is not based on the optical process of how the image is formed and the physical insight is different from Eqs. (1) to (3).

2.2. Retinal image illumination correction

Since retinal images may suffer from uneven and insufficient illumination conditions, correction of the uneven illumination is necessary to enhance its visual quality. Methods for illumination correction are categorized into model-based and non-model-based.

For non-model-based methods, Zhou et al. proposed to use gamma correction to the V channel of retinal images in HSV color space (Zhou et al., 2018). This process provides a luminance gain matrix, which can later be applied to RGB channels to correct the uneven illumination. Gupta et al. modified Zhou's method by combining the adaptive gamma correction to refine the luminance gain matrix (Gupta and Tiwari, 2019).

For model-based methods, the Retinex theory and its variations are used in many reports about correcting the uneven illumination of retinal images. Based on the assumption where the illumination pattern of an image is the low-frequency component, an estimation of the illumination pattern can be obtained by low-pass filtering the given retinal images and removed in the logarithm domain (Cao and Li, 2020; Cao et al., 2020). Besides, based on the observation that the color reversed image which has insufficient/uneven illumination conditions is similar to the image with hazed effects, the bright channel prior (color inverted version of dark-channel prior) also provides an illumination correction method (Cao and Li, 2021; Savelli et al., 2017), and they can be combined with the Retinex theory (Zhang et al., 2022).

2.3. Cataractous retinal image dehazing

Dehazing of cataractous retinal image seems to be similar to a natural

scene. Dehazing of a single natural scene image is ill-posed, therefore, much prior knowledge were proposed to obtain the transmission matrix \mathbf{T}_{sc} from known only of raw haze image \mathbf{S} . For example, in the framework of the dark channel prior (He et al., 2009), \mathbf{T}_{sc} can be obtained by local minimum filtering of \mathbf{S}_Y^{IC} among RGB color channel. In the color attenuation prior (Zhu et al., 2015), \mathbf{T}_{sc} can be estimated from the difference between image brightness and saturation. In the gamma correction prior (Ju et al., 2021), \mathbf{T}_{sc} is given by the intensity difference before and after the gamma correction of the raw image.

However, as we show in **Supplementary Note 1**, prior knowledge of natural haze-images fails to obtain the appropriate \mathbf{T}_{sc} in the case of retinal image dehazing. For example, the dark channel prior fails because there are insufficient shadow areas in retinal image \mathbf{S} . The color attenuation prior fails because the spectral statistical property of retinal images is different from nature images due to the existence of the ocular lens. While the failure of gamma correction prior fails is due to that Eq. (3) cannot be simplified to meeting the derivation of gamma correction prior due to the squared term of \mathbf{T}_{sc} . In general, new dehazing algorithms that are dedicated to retinal image dehazing tasks need to be designed.

3. The DPFR model and image preprocessing

3.1. The DPFR in the YPbPr color space

As shown in Fig. 2(a), the DFPR model for retinal image formation in RGB color space is described by Eq. (3). Let's consider the linear transform matrix Ψ that converts RGB color space to YPbPr color space, Ψ is given by

$$\Psi = \begin{pmatrix} K_R & K_G & 1 - K_R - K_G \\ \frac{K_R}{2(K_R + K_G)} & \frac{K_G}{2(K_R + K_G)} & \frac{1}{2} \\ \frac{1}{2} & \frac{K_G}{2(1 - K_R)} & \frac{1 - K_R - K_G}{2(1 - K_R)} \end{pmatrix}, \quad (5)$$

and $\Psi(1, 1, 1)^T = (1, 0, 0)$. Applying Eq. (5) to both sides of Eq. (3), we obtained

$$\begin{pmatrix} \mathbf{S}_Y \\ \mathbf{S}_{P_B} \\ \mathbf{S}_{P_R} \end{pmatrix} = \mathbf{L} \cdot \mathbf{T}_{\text{lens}}^2 \cdot \left[\mathbf{T}_{sc}^2 \cdot \begin{pmatrix} \mathbf{O}_Y \\ \mathbf{O}_{P_B} \\ \mathbf{O}_{P_R} \end{pmatrix} + \begin{pmatrix} 1 - \mathbf{T}_{sc} \\ 0 \\ 0 \end{pmatrix} \right]. \quad (6)$$

Eq. (6) reveals the mathematical insight that the hazed effect only happens to the Y channel of the retinal image as the Y channel has the haze term $\mathbf{I}_{\text{ill}} \cdot \mathbf{T}_{\text{lens}}^2 \cdot (1 - \mathbf{T}_{sc})$ which denotes the backscattering of light from intraocular scatters as shown in Fig. 2(b). For P_B and P_R channels,

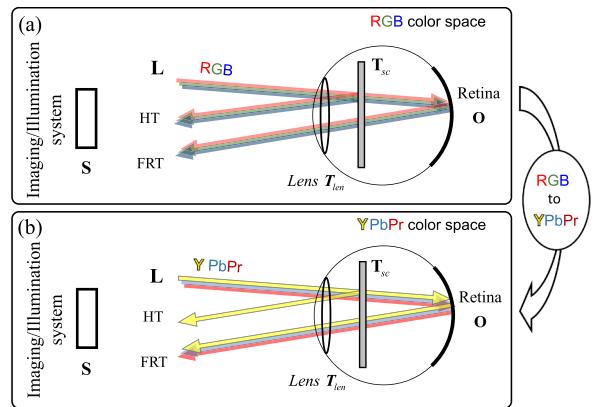


Fig. 2. The sketch of the double-pass fundus reflection model. \mathbf{S} is the final output image of the imaging system. (a) in RGB color space. (b) in YPbPr color space. HT: haze term; FRT: fundus reflection term.

the existence of \mathbf{T}_{sc} works similar to \mathbf{I}_{ll} which only decreases the intensity of corresponding color channels of the haze-free image. According to ITU-R BT.601 standard, $K_R = 0.299$ and $K_G = 0.587$.

According to Eq. (6), our proposed algorithm converts the color retinal image in $Y\mathbf{P}_B\mathbf{P}_R$ color space, and three channels are treated differently. Both intensity correction and dehazing are applied to the Y channel, while only intensity correction is applied to the \mathbf{P}_B and \mathbf{P}_R channels. After then, the retinal image is converted back to RGB color space. In our proposed method, the grayscale is normalized between [0, 1] by dividing the loaded images (uint8) by 255.

3.2. Image preprocessing

Before entering the dehazing process, preprocessing including background padding and illumination correction is applied to the raw images. A retinal image is composed of two major regions - a circular retinal region of interest (ROI) and a black background. Let Γ be all pixels in the circular ROI. Since the image will be filtered by Gaussian kernels during the following process, unexpected enhancement results often occur when the convolution kernel slides through the retinal image due to the sudden change of pixel values. Padding the black areas before enhancement is an effective way to reduce the boundary effect during the convolution. Here we use mirror reflection padding to fill the black background.

Hereafter, we apply our previously reported adaptive illumination correction method to the Y channel of the retinal image to correct its uneven and insufficient illumination pattern. The illumination correction can unify the intensity distribution of input images, and benefit following dehazing process. Since this research focuses on the dehazing process, we refer to Ref (Zhang et al., 2022) or Supplementary Note 2 for detailed information on illumination correction. Moreover, the ablation study for illumination correction will be discussed in Sections

6.2 and 6.3. Fig. 3 shows the raw cataractous retinal images and corresponding Y channels before and after illumination correction treatment.

4. Dehazing using pixel-wise stimulated multilevel denoising

After retinal image illumination correction, the Y channel becomes

$$\mathbf{S}_Y^{IC} = \mathbf{T}_{sc}^2 \circ \mathbf{O}_Y + \mathbf{1} - \mathbf{T}_{sc}, \quad (7)$$

\mathbf{S}_Y^{IC} denotes the Y channel of the source image after illumination correction. The first term $\mathbf{T}_{sc}^2 \circ \mathbf{O}_Y$ of Eq. (7) is the scene term denoting the light reflected from the fundus and attenuated by the cataract layer. The second term $\mathbf{1} - \mathbf{T}_{sc}$ is the haze term denoting the backscattering light from the cataract layer. Once \mathbf{T}_{sc} is known, \mathbf{O}_Y can be retrieved by

$$\mathbf{O}_Y = \frac{\mathbf{S}_Y^{IC} - \mathbf{1}}{\mathbf{T}_{sc}^2} + \min\left(\frac{\mathbf{1}}{\mathbf{T}_{sc}}, 1\right). \quad (8)$$

Considering that (I) the retinal image \mathbf{O}_Y contains blood vessels, and many speckle-like textures due to the scattering of biological tissues, and (II) the haze term is spatially slow-varying since the cataract lays on the pupil plane which is severely out-of-focus from the retinal plane (Arias and Artal, 2020; Peli and Peli, 1989). (III) Eq. (8) is analog to Eq. (4) as $\mathbf{T}_{sc}^2 \cdot \mathbf{O}_Y$ can be regarded as the structure term and $\mathbf{1} - \mathbf{T}_{sc}$ can be regarded as the background term. We, therefore, are able to regard $\mathbf{T}_{sc}^2 \cdot \mathbf{O}_Y$ as the “noise” imposed on $\mathbf{1} - \mathbf{T}_{sc}$ and can be removed by image denoising, yielding

$$\text{Denoising}(\mathbf{S}_Y^{IC}) \approx \mathbf{1} - \mathbf{T}_{sc} \Rightarrow \mathbf{T}_{sc} \approx \mathbf{1} - \text{Denoising}(\mathbf{S}_Y^{IC}), \quad (9)$$

Eq. (9) implies that the transmission matrix \mathbf{T}_{sc} can be obtained by denoising (or smoothing) the input hazy image \mathbf{S}_Y^{IC} .

Since off-the-shelf denoising algorithms are not designed for cata-

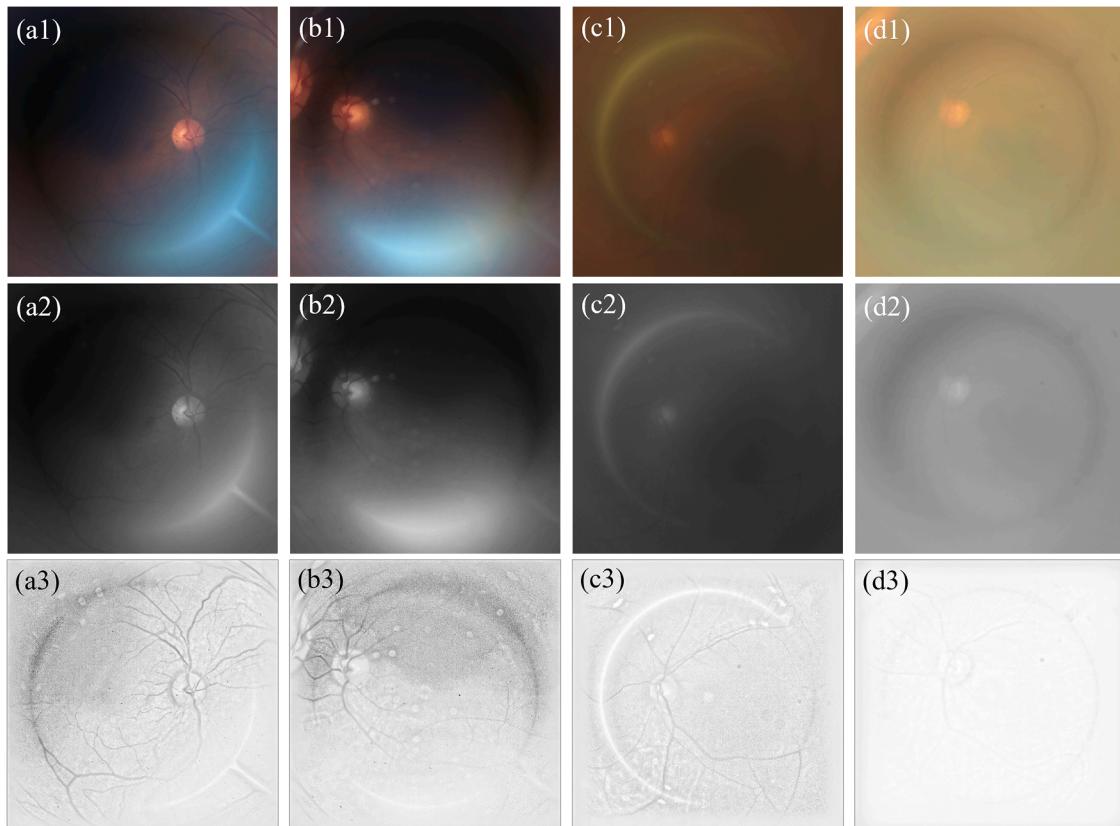


Fig. 3. Image preprocessing. (a1) to (d1) are background-padded retinal images. (a2) to (d2) are their Y channels, respectively. (a3) to (d3) show the illumination correction results.

ractious retinal image dehazing tasks, the denoised image needs to be further refined to form a good estimation of \mathbf{T}_{sc} . Inspired by multi-scale treatment, we consider the \mathbf{T}_{sc} as a linear combination of multiple denoised images of different denoising strengths. Images with strong denoising can be regarded as coarse estimations of \mathbf{T}_{sc} since they contain large-scale features of \mathbf{T}_{sc} . While images with weak denoising can be regarded as fine estimation of \mathbf{T}_{sc} but corrupted by the reflected components $\mathbf{T}_{sc}^2 \cdot \mathbf{O}_Y$. We are further inspired by stimulating functions, and we assign stimulated functions to different levels of denoised images so that we can design a cost function that guides the model on which pixel should contain more details and which pixel shouldn't. (For the detailed relationship between image dehazing and image denoise, please refer to **Supplementary Note 3**.)

Based on this idea, here we proposed the adaptively multilevel stimulated denoising strategy (MUTE) to obtain a delicate estimation of \mathbf{T}_{sc} . The structure of MUTE is shown in Fig. 4. Eq. (9) is rewritten as

$$\mathbf{T}_{sc} = 1 - \mathcal{A} \cdot \left\{ \mathcal{S}_0[\mathbf{S}_Y^{IC}(\mathbf{r})] + \sum_{l=1}^{\mathcal{L}} \theta_l(\mathbf{r}) \circ \mathcal{S}_l[\mathbf{S}_Y^{IC}(\mathbf{r})] \right\}, \quad (10)$$

where \mathcal{A} is a global amplitude term that controls the magnitude of all denoising terms. \mathcal{S}_l , ($l = 0, 1, 2 \dots \mathcal{L} - 1$) is the denoising of $\mathbf{S}_Y^{IC}(\mathbf{r})$ at l -th level of denoising strength (a total of \mathcal{L} levels, $\mathcal{L} \geq 2$). $\theta_l(\mathbf{r})$ is the pixel-wise stimulated (sigmoid) function which is given by

$$\theta_l(\mathbf{r}) = \frac{1}{\mathcal{L} - 1} \left\{ \frac{1}{1 + \exp[-\varphi_l(\mathbf{r})]} \right\} \quad (11)$$

that assigns different weights at point \mathbf{r} for l -th denoising level. $\varphi_l(\mathbf{r})$ is the parameter that controls the stimulated weight for $\theta_l(\mathbf{r})$. We choose the sigmoid function due to its simple expression of its first-order derivative and the sigmoid function gets saturated even if the value of $\varphi_l(\mathbf{r})$ gets very large, which makes the gradient descent won't go too far from the optimal points in each iteration. The remaining tasks are to design a cost function that guides the automatic adjustment of $\varphi_l(\mathbf{r})$ and the amplitude \mathcal{A} .

Recall that an underestimation of \mathbf{T}_{sc} darkens the intensity of \mathbf{O}_Y but enriches its texture information (Galdran et al., 2018), while the overestimation of \mathbf{T}_{sc} brightens the intensity of \mathbf{O}_Y but decreases the texture information due to the haze effect. To optimize the parameters $\varphi_l(\mathbf{r})$ and \mathcal{A} , we design an intensity-based cost function

$$E = \sum_{\mathbf{r}} (\mathbf{O}_Y - \bar{I})^2 - \lambda \sum_{\mathbf{r}} |\nabla \mathbf{O}_Y|^2, \quad (12)$$

where $\bar{I} = I_0 \times \text{mean}(\mathbf{S}_Y^{IC})$, and I_0 is the intensity ratio that controls the

dehazing strength. A small value of I_0 leads to a strong dehazed effect while also darkening the image. The first term ensures that the output \mathbf{O}_Y has a proper intensity distribution, and the second term ensures the image has rich texture information. λ ($\lambda < 1$, or $\lambda = 0$) is a small value to ensure the first term contributes most to the gradient. The MUTE can be regarded as a one-layer network, with only sigmoid functions. The derivative of Eq (12) with respect to parameters $\varphi_l(\mathbf{r})$ and \mathcal{A} can be either calculated by automatic differentiation or calculated in closed-form as written in **Supplementary Note 4**.

In MUTE, we use the Adam optimizer to update model parameters (Kingma and Ba, 2015), $\eta = 0.01$ (step size), $\alpha = 0.9$ (first-order decay rate), and $\beta = 0.99$ (second-order decay rate) in all following experiments. Moreover, Let $R_n = \underset{\mathbf{r} \in \Gamma}{\text{mean}}(\mathbf{O}_Y - \bar{I})^2$ be part of the residual of the cost function in the n -th iteration and we stop the iteration process when $R_n < 0.005$. Fig. 5 shows the flowchart of the proposed dehazing algorithm.

Many off-the-shelf image denoising algorithms can be used for computing \mathcal{S}_l in Eq. (10), here we choose the Gaussian filter for simplicity which is given by

$$\mathcal{S}_l(\mathbf{r}) = \mathbf{S}_Y^{IC}(\mathbf{r}) \otimes \left[\frac{1}{2\pi w_l^2} \exp\left(-\frac{\mathbf{r}^2}{2w_l^2}\right) \right], \quad (13)$$

and the size of the gaussian kernel at l -th level is

$$w_l = \left(\frac{\sqrt{M \cdot N}}{20} \right) \times \left(\frac{1}{2} \right)^{\mathcal{L}-l}. \quad (14)$$

Fig. 6 shows \mathbf{O}_Y and corresponding \mathbf{T}_{sc} when $\lambda = 1$, $\mathcal{L} = 4$ and $I_0 = 0.50$ for images in Figs. 3(a3) to 3(d3), respectively. As shown in Fig. 6 (a1) to (d1), the patterns of \mathbf{T}_{sc} have low transmission maps due to the presence of cataracts. Since the intensity of the pattern is low for haze retinal images, in order to have a clear observation of the transmission pattern, we normalized the intensity in the area Γ as shown in the second row of Fig. 6, where the areas near the blood vessels are regarded as having higher transmission than other places.

Comparing images in Fig. 3 (a3-d3) to Fig. 6. (a3-d3), the haze effect is significantly suppressed as the retinal structures including optical disks and blood vessels can be clearly observed. Video 1 and Video 2 demonstrate how the Y channel and the transmission matrix gradually evolve in each iteration step for Fig. 6(c2) and Fig. 6(b2). Since the iteration could be fast and converged in only a few steps, we insert additional frames in the videos for better demonstration.

In general, a flowchart for the proposed enhancement method is summarized in Fig. 7. The Hessian denoising is an optional process that will be discussed in Section 7 in detail.

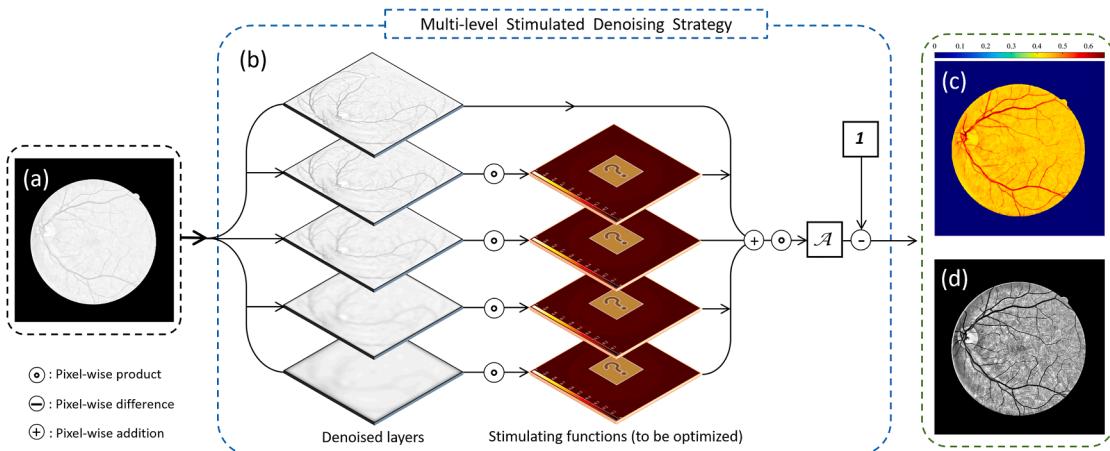


Fig. 4. Multilevel-stimulated denoising strategy for retinal image dehazing. (a) input Y channel. (b) Multilevel denoised image controlled by stimulating functions. (c) the final output of estimation of \mathbf{T}_{sc} . (d) final output Y channel of the restored image.

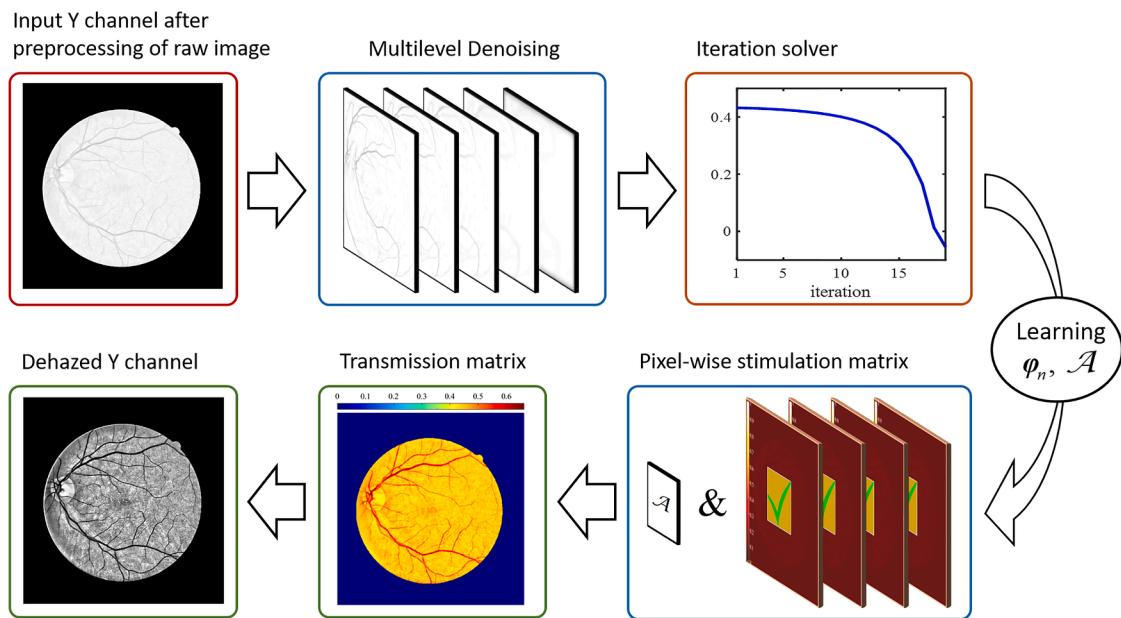


Fig. 5. flowchart of proposed Y channel dehazing method.

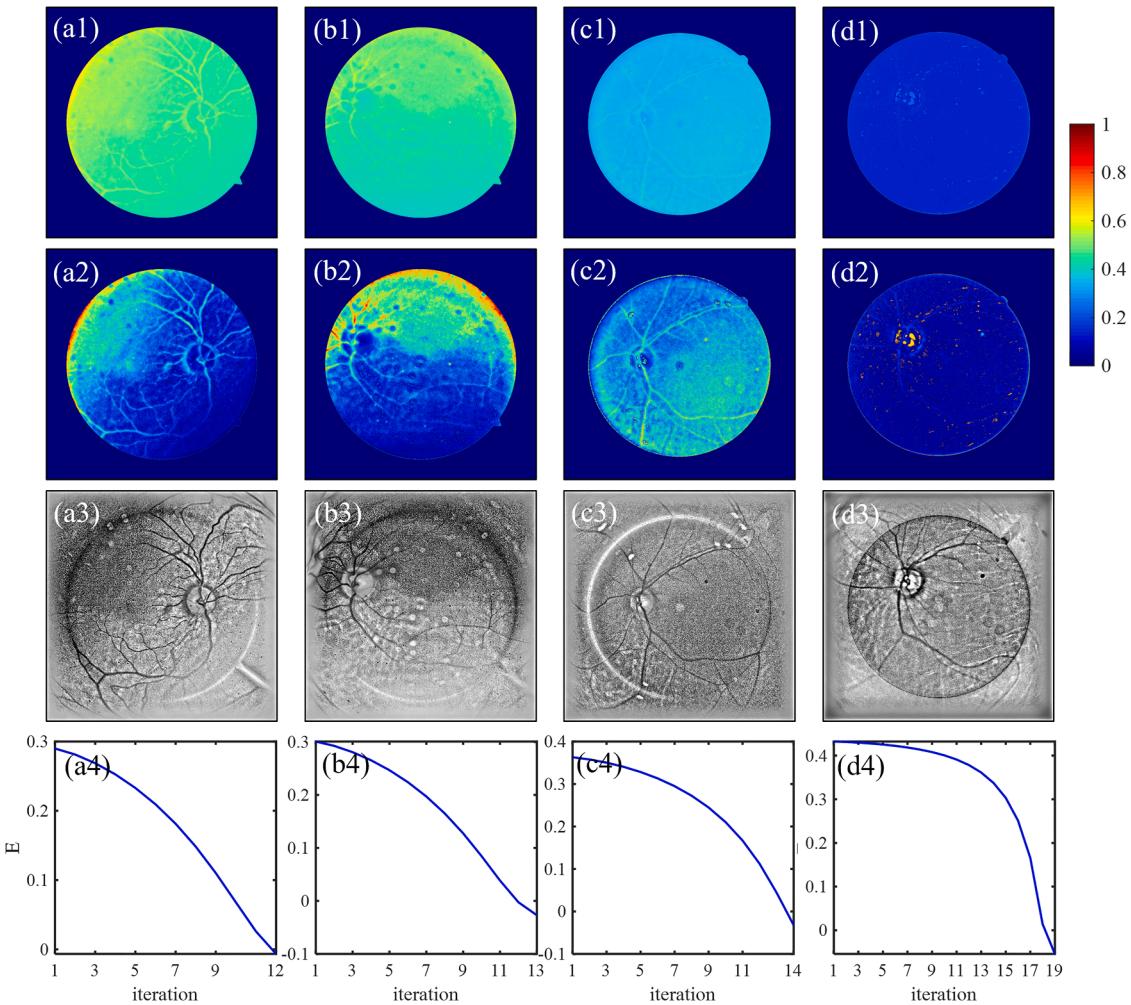


Fig. 6. Results for Y channel dehazing. (a1) to (d1) are obtained T_{sc} for images in Figs. 3(a3) to 3(d3). The patterns are truncated by the mask of the region of interest. (a2) to (d2) are intensity normalized pattern of (a1) to (d1). (a3) to (d3) are dehazed Y channel. (a4) to (d4) are changes of the cost function with respect to iterations.

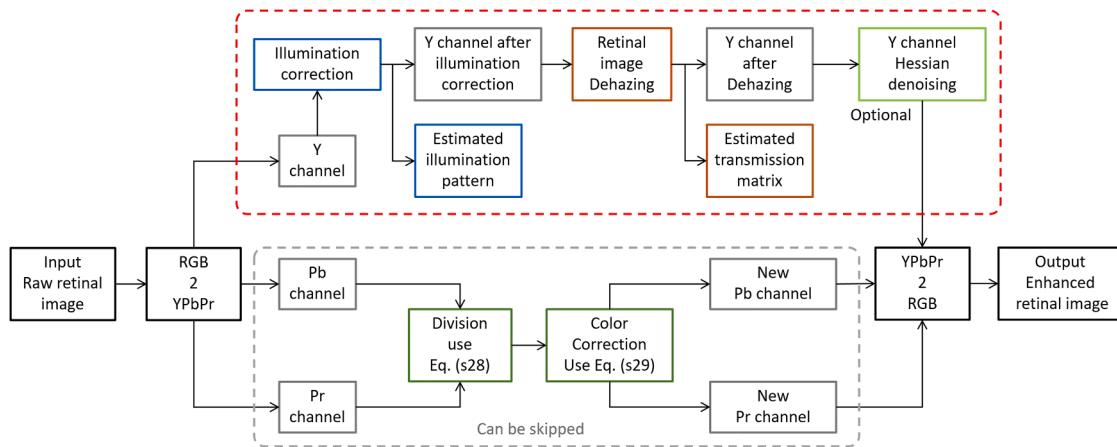


Fig. 7. A flowchart for MUTE. Eq. (s28) and Eq. (s29) are in the **Supplementary Note 5**.

5. Experimental results

To show its performance, this section compares, both qualitatively and quantitatively, MUTE with that of three recently published state-of-art methods, namely the decomposition and visual adaptation (DVA) (Wang et al., 2021), the detail-richest-channel method (DRC) (Cao and Li, 2021), and the low-pass filtering and α -rooting (LPAR) (Cao et al., 2020). Note that in (Cao and Li, 2021) and (Cao et al., 2020) it was proven that the enhancement performances of DRC and LPAR are better than methods presented in (Xiong et al., 2017) and (Mitra et al., 2018). Image black background padding was applied to all raw retinal images before retinal image enhancement. The treatment in Pb/Pr channel is

not applied.

In actual implementation, the kernel parameters for DRC and LPAR are determined adaptively according to the size of input images. Other parameters for DRC and LPAR models such as intensity correction strength can be automatically determined by the models. While DVA method has two parameters, λ to determine the frequency band, and α to control the local contrast for the enhancement. In our implementation, $\lambda = 0.3$, $\alpha = 600$ based on (Wang et al., 2021). For MUTE, we choose $I_0 = 0.52$ in following experiments. An explanation of this selection will be discussed in Section 6.2.

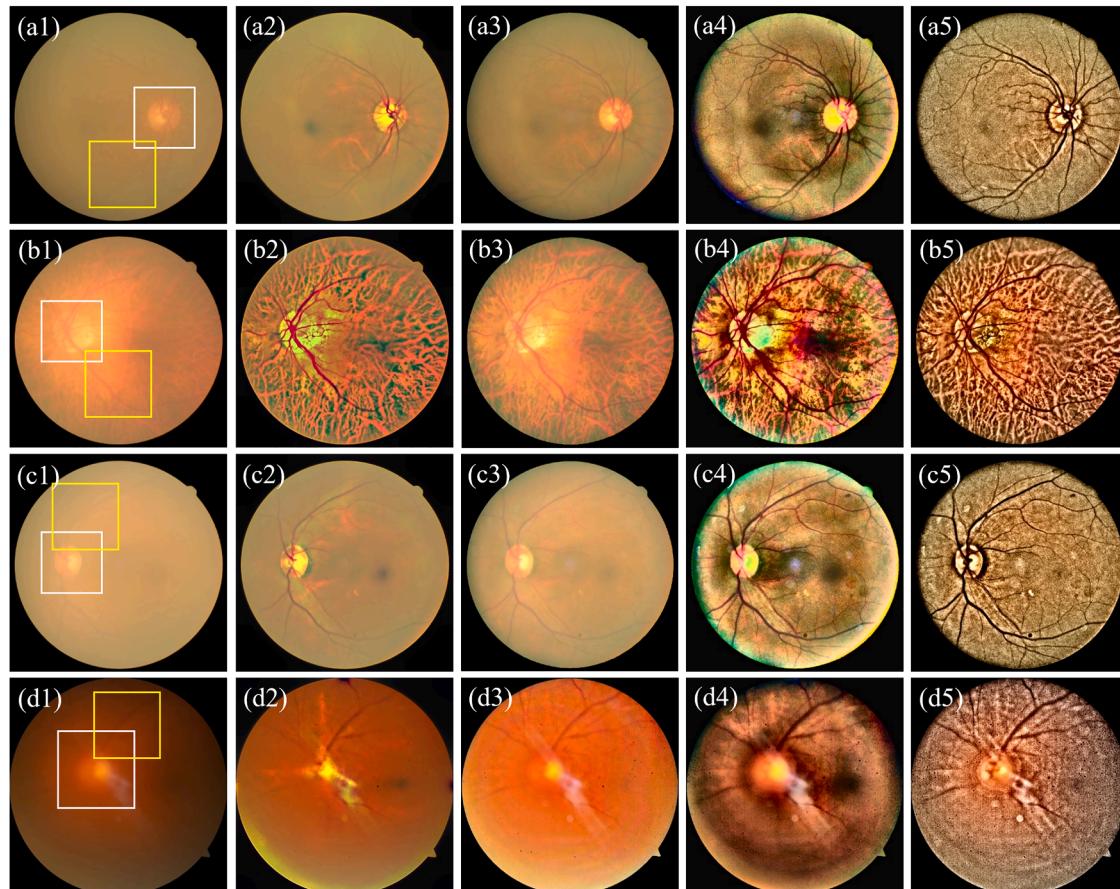


Fig. 8. Cataractous retinal image enhancement results. Columns from left to right are raw images, enhanced results of DVA, DRC, LPAR, and MUTE methods.

5.1. Comparing to SOTA non-learning methods

5.1.1. Visual assessments

First of all, we compared the results to assess the performance of the proposed method in terms of removing the haze effect and enhancing retinal structure visibility with images in public and our proprietary cataractous retinal image databases (**cataract**). Four groups of raw images and corresponding enhanced images are shown in Fig. 8. The first column shows raw cataract retinal images, where each image has a severe haze effect as most of the structures including blood vessels and optical disks are hidden behind the cataract layer.

From visual assessment, DVA corrects the haze effect for some images as shown in Fig. 8(b2) and 8(c2), as the blood vessels near the optical disks are enhanced. However, the performance of DVA is not stable as it failed to correct the haze effect for images in Fig. 8(a2) and Fig. 8(d2), especially for the region that is far from the optical disks.

The results of DRC are shown in Fig. 8(a3) to (d3), where the haze effects are still significant, although the detailed structures on the retinal image are enhanced. Results of LPAR are shown in Fig. 8(a4) to (d4), where the haze effect is removed as the blood vessels can be clearly observed. However, the LPAR causes an overexposure effect on the bright area of the optical disk as shown in Fig. 8(b4) and Fig. 8(c4).

Note that some DVA results are similar to that of LPAR method as shown in Fig. 8(b2) and Fig. 8(b4), the main reason being that both belong to variations of Eq. (4). DVA uses TV regularization while the LPAR uses the Gaussian filter to obtain the background components of the retinal image. Meanwhile, the LPAR uses an adaptive strategy to determine the parameters for retinal image enhancement, thus, the LPAR has better performance on retinal images of different conditions than that of DVA method whose parameters still need hand-tuning.

Results of MUTE are shown in the last column of Fig. 8. The haze effect is efficiently suppressed and blood vessels that are hidden behind the cataract can be now observed. Meanwhile, MUTE also uses an adaptive manner to determine model parameters. It does not lead to an overexposure effect, as it can restore a clear optical disk area as shown in Fig. 8(b5) and Fig. 8(c5).

Partially enlarged images for the areas in the white and yellow boxes are shown in Fig. 9 and Fig. 10, respectively. All four methods show their ability in contrast enhancement, while the DVA and LPAR methods lead to an overexposure effect at the bright area as shown in Fig. 9(b2), (c2), (b4), and (c4). MUTE enhances image contrast and some small blood vessels within the optical disk can be observed as shown in Fig. 9(b5) and Fig. 10(a5) to Fig. 10(d5). Some blood vessels that are hard to be observed in the raw images are now can be clearly found, especially in Fig. 9(d5) and Fig. 10(a5).

According to the experimental results, MUTE suppresses the haze effect of the cataractous retinal image and gains better contrast-enhanced results than other SOTA methods.

5.1.2. Objective assessments

To evaluate the performance of the four methods, we also conducted a quantitative evaluation in terms of (1) image sharpness, (2) image local contrast, (3) image entropy, (4) image multiscale contrast, and (5) image fog density. All images were reformatted in 8-bit unsigned integers (uint8, the gray-value is ranged in [0, 255]).

We adapted the Underwater Image Sharpness Measure (UISM) and Underwater Image Contrast Measure (UIConM) for image sharpness and image local contrast since severe blurring occurs in both underwater images and retinal images due to forward scattering (Panetta et al., 2016). More importantly, both the UISM and the UIConM do not rely on

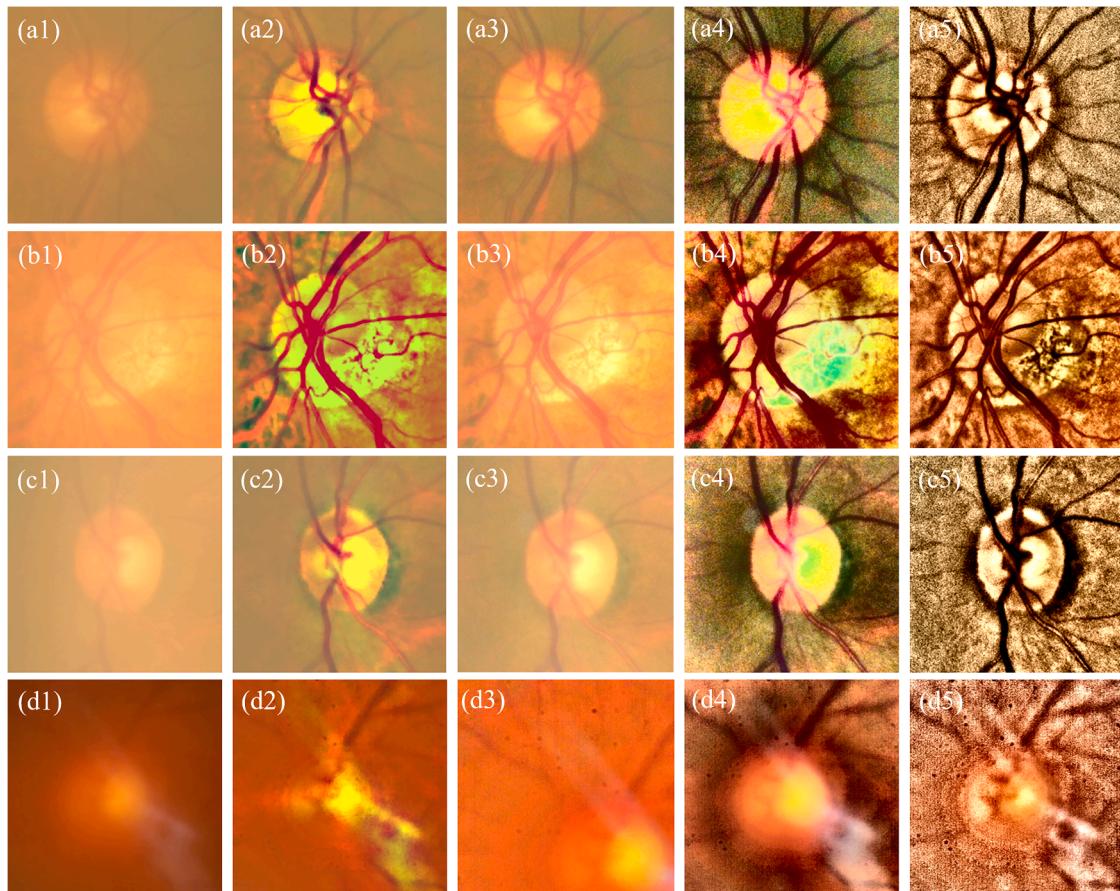


Fig. 9. Partially enlarged areas of raw images in the white boxes. Columns from left to right are raw images, enhanced results of DVA, DRC, LPAR, and MUTE methods.

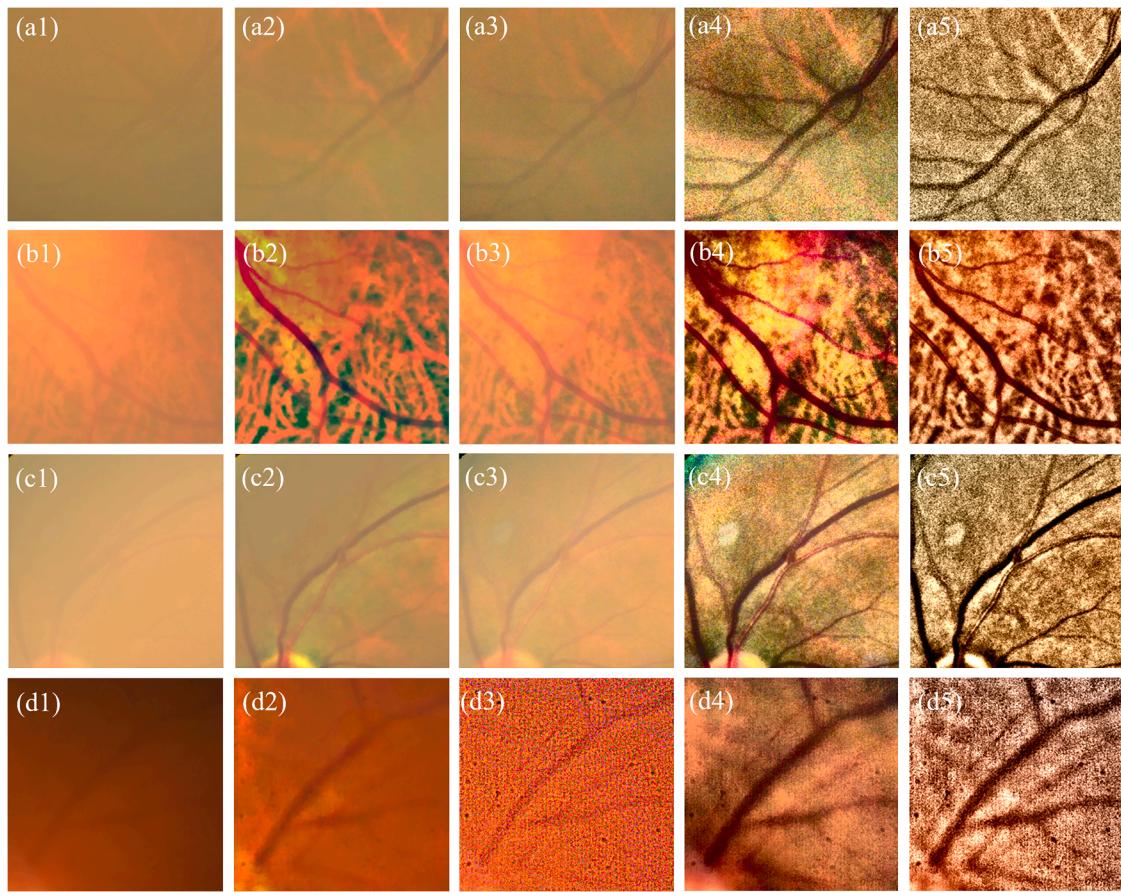


Fig. 10. Partially enlarged areas of raw images in the yellow boxes. Columns from left to right are raw images, enhanced results of DVA, DRC, LPAR, and MUTE methods.

the statistical property of images and thus can be applied to retinal images, regardless of the statistical difference between retinal images and underwater images. The UISM and the UIConM are calculated by

$$UISM = \sum_{c \in \{R,G,B\}} \lambda_c \cdot \left[\frac{2}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \log \left(\frac{I_{edge,max,k,l}}{I_{edge,min,k,l}} \right) \right], \quad (15)$$

and

$$UIConM = \sum_{c \in \{R,G,B\}} \lambda_c \cdot \log \left[\frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \times \sum_{k=1}^{k_2} \frac{I_{max,k,l} \Theta I_{min,k,l}}{I_{max,k,l} \oplus I_{min,k,l}} \log \left(\frac{I_{max,k,l} \Theta I_{min,k,l}}{I_{max,k,l} \oplus I_{min,k,l}} \right) \right]. \quad (16)$$

In Eq. (15), I_{edge} denotes the product between the Sobel edge map and the original image in the corresponding color channel. λ_c is the weight that is assigned to each color channel according to the visual response. $\lambda_R = 0.299$, $\lambda_G = 0.587$ and $\lambda_B = 0.114$ are used according to the relative visual responses of the red, green, and blue channels (Panetta et al., 2016). The image is divided into $k_1 k_2$ blocks, $I_{max,k,l}/I_{min,k,l}$ indicating the relative contrast ratio within each block.

In Eq. (16), \oplus , \otimes , and Θ are the PLIP operations in which $A \oplus B = A + B - A \cdot B / \gamma(M)$, $A \otimes B = \gamma(M) - \gamma(M) \cdot [1 - B / \gamma(M)]^A$, and $A \Theta B = \kappa(M) \cdot (A - B) / [\kappa(M) - B]$. In this research, we choose $\gamma(M) = 1024$ and $\kappa(M) = 600$ are chosen and the image is divided into 10×10 blocks, viz $k_1 = k_2 = 10$.

Image entropy (IE) describes the randomness distribution of the image and its value denotes the amount of image information (Hong et al., 2021; Wang et al., 2018), which is given by

$$IE = \sum_{c \in \{R,G,B\}} \lambda_c \sum_{g=0}^{255} P(x_g) \log(x_g), \quad (17)$$

where $P(x_g)$ is the probability of the appearance of the pixels that have gray-value g in the gray-scaled image. Image entropy can be used to characterize the texture of the image [40] and determine the amount of image information. Images affected by haze tend to have low IE values due to the biased brightness distribution. In contrast, haze-free images have a relatively high IE.

The multiscale contrast of the image, C_{RAMM} , was calculated with a pyramidal multi-resolution representation of luminance (Rizzi et al., 2004). C_{RAMM} is defined as

$$C_{RAMM} = \frac{1}{8(\#pixel)(\#levels)} \sum_{\forall level} \left[\sum_{\forall pixel} \left(\sum_{8-neigh} |P_i - P_j| \right) \right], \quad (18)$$

where $\#pixel$ denotes the total pixel numbers of the image, $\forall level$ denotes the total level of down-sampling, in each level the image is halved without pre-filtering. In this paper, we use 6 levels of down-sampling. The pixel numbers of retinal images in the databases are large enough for 6 levels of down-sampling.

Lastly, the fog-aware density evaluator (FADE) (Choi et al., 2015b) (Choi et al., 2015a) was used to numerically predict perceptual hazy density.

Table 1 demonstrates the enhancement results of the quantitative evaluations using objective metrics. Here, High values of UISM, UIConM, IE, and C_{RAMM} indicate high image quality, while low values of FADE indicate a better dehazing performance.

Numbers marked in bold denote the best performance, while

Table 1

Quantitative assessment of the enhanced images among Fig. 8(a1) to (d1) with five metrics.

Raw image	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Fig. 8(a1)	UISM↑	0.7782	0.9069	2.9260	4.6992	6.7863
	UIConM↑	0.0578	0.0938	0.0780	0.2909	0.3851
	IE↑	5.2354	5.1692	5.6260	7.4339	7.5100
	C _{RAMM} ↑	1.1207	1.5072	1.4048	4.7888	7.5293
	FADE↓	0.5127	1.1202	0.6086	0.2214	0.1554
Fig. 8(b1)	UISM↑	0.8098	2.2411	3.2962	7.0178	5.9409
	UIConM↑	0.0739	0.1950	0.1129	0.3534	0.3759
	IE↑	5.8075	6.2273	6.0271	7.2154	7.3425
	C _{RAMM} ↑	1.3705	3.7879	1.9448	7.1645	8.8383
	FADE↓	0.4347	0.6151	0.4590	0.1746	0.1191
Fig. 8(c1)	UISM↑	0.7715	0.9615	2.9473	4.4087	7.1608
	UIConM↑	0.0745	0.0968	0.0914	0.2970	0.3876
	IE↑	5.0473	4.9388	5.4631	7.3920	7.4646
	C _{RAMM} ↑	1.2978	1.5133	1.5083	4.9293	8.0858
	FADE↓	0.6148	1.3645	0.7896	0.2260	0.1427
Fig. 8(d1)	UISM↑	1.1770	1.5114	3.3452	4.3085	8.1445
	UIConM↑	0.0274	0.0805	0.1279	0.2186	0.4020
	IE↑	6.1477	6.2052	6.3507	7.3024	7.5447
	C _{RAMM} ↑	0.7426	1.4215	1.9617	3.8497	8.4174
	FADE↓	0.7134	0.7112	0.2980	0.2129	0.1183
Average	UISM↑	0.8841	1.4052	3.1287	5.1086	7.0081
	UIConM↑	0.0584	0.1165	0.1026	0.2900	0.3877
	IE↑	5.5595	5.6351	5.8667	7.3359	7.4655
	C _{RAMM} ↑	1.1329	2.0575	1.7049	5.1831	8.2177
	FADE↓	0.5689	0.9528	0.5388	0.2087	0.1339

numbers marked in green denote the second-best performance. As listed in Table 1, MUTE performed best in aspects of UISM, UIConM, IE, C_{RAMM}, and FADE, as it gained better scores than the other three methods. The DRC, LPAR, and MUTE decrease the FADE value denoting good dehazing results.

Table 2 compares the average and standard deviation among five quality matrixes for raw and enhanced images from both public and proprietary cataractous retinal image datasets. In general, the performance of MUTE is better than those of the other three methods, as it has the best average scores among all four methods. The robustness of MUTE is also better than the other three SOTA methods which are implied by the small standard deviation of MUTE.

Apart from the cataractous retinal image, we test the performance of MUTE on non-cataractous retinal images from public databases of Dia-Ret- ([DIARETDB0 - Standard](#); [DIARETDB1 - Standard](#)), STARE- ([Hoover et al., 2000](#); [SStructured Analysis of the Retina](#)), MESSIDOR- ([Decencière et al., 2014](#)), and NORMAL- ([cataract](#)) datasets, adding up to a total of 2116 retinal images.

One group of results are shown in Fig. 11 where the MUTE increases

the image's contrast and also maintains the color naturalness as it won't cause over-exposure effect to the optical disk area. More experimental results for visual assessments are available in [Supplementary Note 6](#) and [Note 7](#). The statistical results are listed in Table 3, for non-cataractous retinal images DVA had the best score on IE, while MUTE showed the best score on UISM, UIConM, C_{RAMM}, and FADE.

In general, MUTE increases the contrast of retinal images by 7-fold, while decreasing their haze density by more than about 3-fold. The small standard deviation value of MUTE denotes its algorithm robustness.

5.2. Comparing to SOTA deep-learning methods

In this section, we compare our proposed MUTE against three SOTA deep-learning methods including ArcNet ([Li et al., 2022b](#)), ScrNet ([Li et al., 2022a](#)), and ArSrNet ([Yang et al., 2023](#)). All three methods are published in 2022. We test the performance of these methods on the cataractous retinal image dataset in order to show their ability on cataract retinal image dehazing. For ArcNet and ScrNet, we used the published pre-trained model in <https://github.com/liamheng/Annotation-free-Fundus-Image-Enhancement>. For ArSrNet, we asked the author to process the images in their local machine and sent them back to us.

5.2.1. Visual assessment

The raw images are resized into 512 by 512 pixels in order to produce good results since these three networks are trained based on simulated cataract retinal images of the same size. Where the ArcNet is trained based on images of 256 by 256 pixels. For the larger size of input images, the network should be trained again due to their lack of generalization. While our proposed MUTE can be applied to any size of input images.

Fig. 12 and Fig. 13 show the output images of four methods. Except for Fig. 12(a1), where the cataract is mild, the cataracts in Fig. 12(b1) to (d1) are rather severe. The enhancements for all three SOAT Networks can merely improve the contrast of the image to a limited degree. The cataract effect as shown in Fig. 12(b2) for ArcNet and Fig. 12(b4) for ArSrNet are still pronounced. While the MUTE significantly improves the image contrast as shown in Fig. 12(b5). Similar results can be also found in the zoomed-in images in Fig. 13.

Note that the ArcNet is trained on images of 256 by 256 pixels, when it is applied to larger images, it will introduce unexpected problematic pixels as can be seen in the black spot in Fig. 12(b2), and Fig. 13(b2). In practical cases, the size of retinal images can be even larger than 2000 by 2000 pixels and vary between different brands of fundus cameras from different companies. The input image size does not influent the output quality of MUTE thanks to the down-up-sampling strategy proposed in [Supplementary Note 10](#).

Table 2

Averages and standard deviations of enhancement results from both public and proprietary cataractous retinal image datasets.

Databases mean ± std.	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Cataract (100 images)	UISM↑	0.9002 ± 0.1432	2.2829 ± 0.9623	3.4599 ± 1.4614	5.5883 ± 0.7461	6.7884 ± 0.7869
	UIConM↑	0.0573 ± 0.0123	0.1752 ± 0.0532	0.1097 ± 0.0437	0.3257 ± 0.0294	0.3890 ± 0.0093
	IE↑	5.8443 ± 0.4449	6.1823 ± 0.4350	6.2841 ± 0.3819	7.4230 ± 0.1960	7.4151 ± 0.1338
	C _{RAMM} ↑	1.2013 ± 0.1741	2.9870 ± 1.0950	1.9515 ± 0.4257	6.2711 ± 0.8769	8.1483 ± 0.6035
	FADE↓	0.4715 ± 0.1155	0.6769 ± 0.2548	0.4502 ± 0.1388	0.2031 ± 0.0321	0.1469 ± 0.0228
Proprietary (94 images)	UISM↑	1.2065 ± 0.3028	2.2897 ± 0.9169	3.8239 ± 0.7211	3.9118 ± 0.9212	6.8936 ± 1.1603
	UIConM↑	0.0446 ± 0.0133	0.1477 ± 0.0549	0.1983 ± 0.0640	0.2140 ± 0.0511	0.3811 ± 0.0646
	IE↑	6.8377 ± 0.5775	6.8649 ± 0.4550	6.7774 ± 0.4714	7.3770 ± 0.1914	7.4889 ± 0.4049
	C _{RAMM} ↑	1.0378 ± 0.1741	2.4631 ± 0.9972	2.8145 ± 0.7723	3.8878 ± 1.2095	7.9087 ± 1.5078
	FADE↓	1.0455 ± 0.3905	0.9345 ± 0.5244	0.2070 ± 0.2112	0.3351 ± 0.1118	0.1857 ± 0.1177
Average (194 images)	UISM↑	1.0486 ± 0.2798	2.2862 ± 0.9381	3.6363 ± 0.6049	4.7760 ± 1.1832	6.8393 ± 0.9844
	UIConM↑	0.0512 ± 0.0143	0.1619 ± 0.0557	0.1527 ± 0.0657	0.2716 ± 0.0695	0.3852 ± 0.0455
	IE↑	6.3164 ± 0.7181	6.5057 ± 0.5639	6.5231 ± 0.4929	7.4045 ± 0.1952	7.4568 ± 0.2996
	C _{RAMM} ↑	1.1221 ± 0.1920	2.7332 ± 1.0785	2.3696 ± 0.7532	5.1163 ± 1.5891	8.0322 ± 1.1388
	FADE↓	0.7496 ± 0.4038	0.8017 ± 0.4272	0.4528 ± 0.1772	0.2671 ± 0.1045	0.1657 ± 0.0856

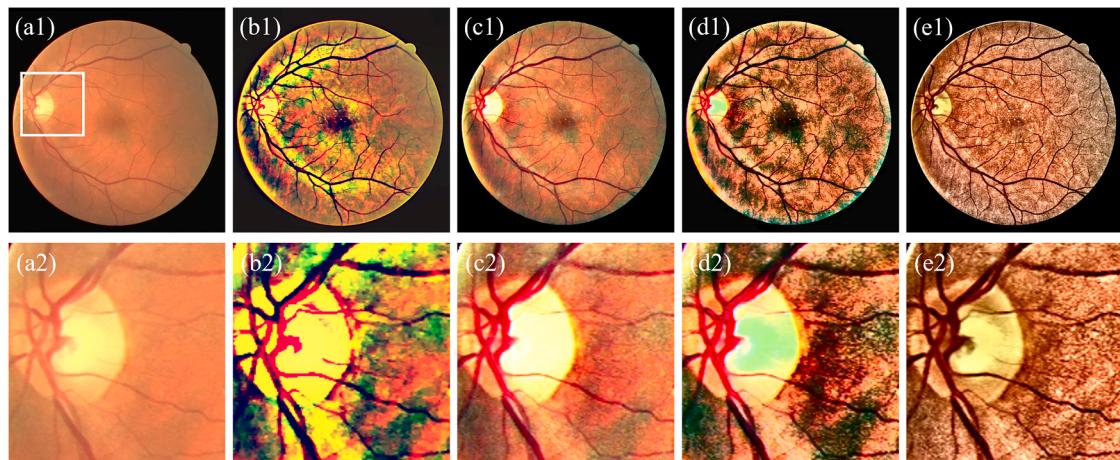


Fig. 11. Group 1 for visual assessment. The raw image is from DRIVE dataset (Staal et al., 2004). Left to right are raw images, results of DVA, DRC, LPAR, and MUTE methods.

Table 3

Averages and standard deviations of enhancement results from non-cataractous retinal image databases.

Databases mean \pm std.	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
DiaRetDB0+DB1 (219 images)	UISM↑	3.2857 \pm 0.2107	3.5914 \pm 0.5870	4.2797 \pm 0.3261	5.3378 \pm 0.8601	6.7373 \pm 0.3610
	UIConM↑	0.0396 \pm 0.0140	0.1793 \pm 0.0393	0.0989 \pm 0.0389	0.2419 \pm 0.0471	0.3977 \pm 0.0068
	IE↑	5.6481 \pm 0.3389	6.2941 \pm 0.2314	5.7972 \pm 0.3434	5.8912 \pm 1.8016	6.6910 \pm 0.1443
	C _{RAMM} ↑	0.7169 \pm 0.1332	2.4855 \pm 0.7205	1.4165 \pm 0.4100	3.7898 \pm 1.2053	8.0577 \pm 0.5667
	FADE↓	0.4752 \pm 0.0395	0.4188 \pm 0.0648	0.3094 \pm 0.0461	0.2495 \pm 0.1344	0.1278 \pm 0.0156
STARE (397 images)	UISM↑	2.6542 \pm 0.6460	4.1758 \pm 1.0861	4.6212 \pm 0.5306	5.5971 \pm 1.0953	6.5124 \pm 0.8229
	UIConM↑	0.0879 \pm 0.0195	0.2440 \pm 0.0458	0.1930 \pm 0.0444	0.2863 \pm 0.0499	0.3693 \pm 0.0180
	IE↑	5.4185 \pm 0.3468	6.3837 \pm 0.2481	5.5470 \pm 0.3198	5.5085 \pm 1.2873	5.8185 \pm 0.2760
	C _{RAMM} ↑	1.9010 \pm 0.2998	5.1454 \pm 1.4516	3.3505 \pm 0.7190	5.4806 \pm 1.2965	7.8314 \pm 0.7522
	FADE↓	0.4744 \pm 0.1274	0.3947 \pm 0.1296	0.2367 \pm 0.0617	0.2739 \pm 0.1453	0.1331 \pm 0.0236
MESSIDOR (1200 images)	UISM↑	1.8419 \pm 0.3224	4.0235 \pm 0.7075	5.0035 \pm 0.4089	6.2763 \pm 0.6584	7.0831 \pm 0.4933
	UIConM↑	0.0746 \pm 0.0185	0.2383 \pm 0.0318	0.1608 \pm 0.0418	0.2777 \pm 0.0280	0.3660 \pm 0.0073
	IE↑	5.1908 \pm 0.3454	6.2726 \pm 0.1881	5.6377 \pm 0.3846	5.9916 \pm 0.7165	6.0725 \pm 0.2111
	C _{RAMM} ↑	1.1810 \pm 0.1639	3.9134 \pm 0.7737	2.2784 \pm 0.4610	5.7179 \pm 0.7411	7.6652 \pm 0.5981
	FADE↓	0.4652 \pm 0.0673	0.4361 \pm 0.0842	0.2320 \pm 0.0410	0.1812 \pm 0.0726	0.1252 \pm 0.0136
NORMAL (300 images)	UISM↑	1.5512 \pm 0.2604	3.4903 \pm 0.7928	4.7259 \pm 0.3344	5.7581 \pm 0.6043	6.9285 \pm 0.4592
	UIConM↑	0.0768 \pm 0.0152	0.2389 \pm 0.0342	0.1700 \pm 0.0336	0.3092 \pm 0.0267	0.3733 \pm 0.0047
	IE↑	5.5313 \pm 0.2497	6.4120 \pm 0.1878	5.8949 \pm 0.2886	6.3338 \pm 0.7880	6.3062 \pm 0.1898
	C _{RAMM} ↑	1.2468 \pm 0.1495	3.9861 \pm 0.8219	2.5810 \pm 0.4295	6.3165 \pm 0.8719	8.3668 \pm 0.4882
	FADE↓	0.6671 \pm 0.1226	0.5629 \pm 0.1132	0.3143 \pm 0.0622	0.2096 \pm 0.0876	0.1497 \pm 0.0186
Average (2116 images)	UISM↑	2.1025 \pm 0.6595	3.9318 \pm 0.8274	4.8175 \pm 0.4819	5.9783 \pm 0.8516	6.9183 \pm 0.5981
	UIConM↑	0.0738 \pm 0.0219	0.2334 \pm 0.0404	0.1618 \pm 0.0478	0.2801 \pm 0.0390	0.3709 \pm 0.0137
	IE↑	5.3291 \pm 0.3735	6.3154 \pm 0.2130	5.6736 \pm 0.3731	5.9391 \pm 1.0404	6.1220 \pm 0.3221
	C _{RAMM} ↑	1.2774 \pm 0.3847	4.0071 \pm 1.1709	2.4333 \pm 0.7364	5.5587 \pm 1.1420	7.8364 \pm 0.6612
	FADE↓	0.4966 \pm 0.1128	0.4445 \pm 0.1094	0.2525 \pm 0.0599	0.2097 \pm 0.1062	0.1304 \pm 0.0188

5.2.2. Objective assessment

The score for objective assessments is listed in Table 4 for images in Fig. 12, and in Table 5, for statistic results for 100 images in the cataractous retinal images dataset. Accordingly, our MUTE increases the image edge information (UISM) larger than SOTA deep-learning method by 2-folds, and increases the image contrast (UIConM, and C_{RAMM}) by 3-folds. The MUTE also decreases the haze effect of cataract retinal image (FADE) by 2-folds less than SOTA deep-learning methods.

The failure of deep-learning methods on cataract retinal image dehazing can be attributed to the training data. Those methods trained the network based on simulated cataract images due to lack of paired cataract and cataract-free retinal images. The simulated cataract is way simple that cannot fully represent the real cataracts happened in practical/clinical situations. As such, there is a high possibility that the model overfitted on the simulation data and lack of generalization on real cataractous retinal image.

In the following subsection, we present demonstrates of MUTE's potential contribution to three clinical applications including blood

vessel segmentations, retinal image registration, and retinopathy diagnosis.

5.3. Potential clinical applications

5.3.1. Retinal image blood vessels segmentations

Retinal image blood vessels segmentation provides the shape distributions of blood vessels, which is important for clinical diagnosis as the morphological changes of blood vessels are biomarkers for diseases such as lacunar stroke (Doubal et al., 2010), cognitive dysfunction (Cheung et al., 2014), cardiovascular risk (Zhu et al., 2014), diabetes (Huang et al., 2016) and glaucoma (Ciancaglini et al., 2015). Blood vessel segmentation can be achieved by either human specialists or computer software. The former provides accurate results but is time-consuming. The latter option provides fast segmentation results but is less accurate than human specialists. Moreover, due to poor image contrast of the cataractous retinal image, hand-based segmentation is even more time-consuming, and automatic segmentation for hazy retinal images

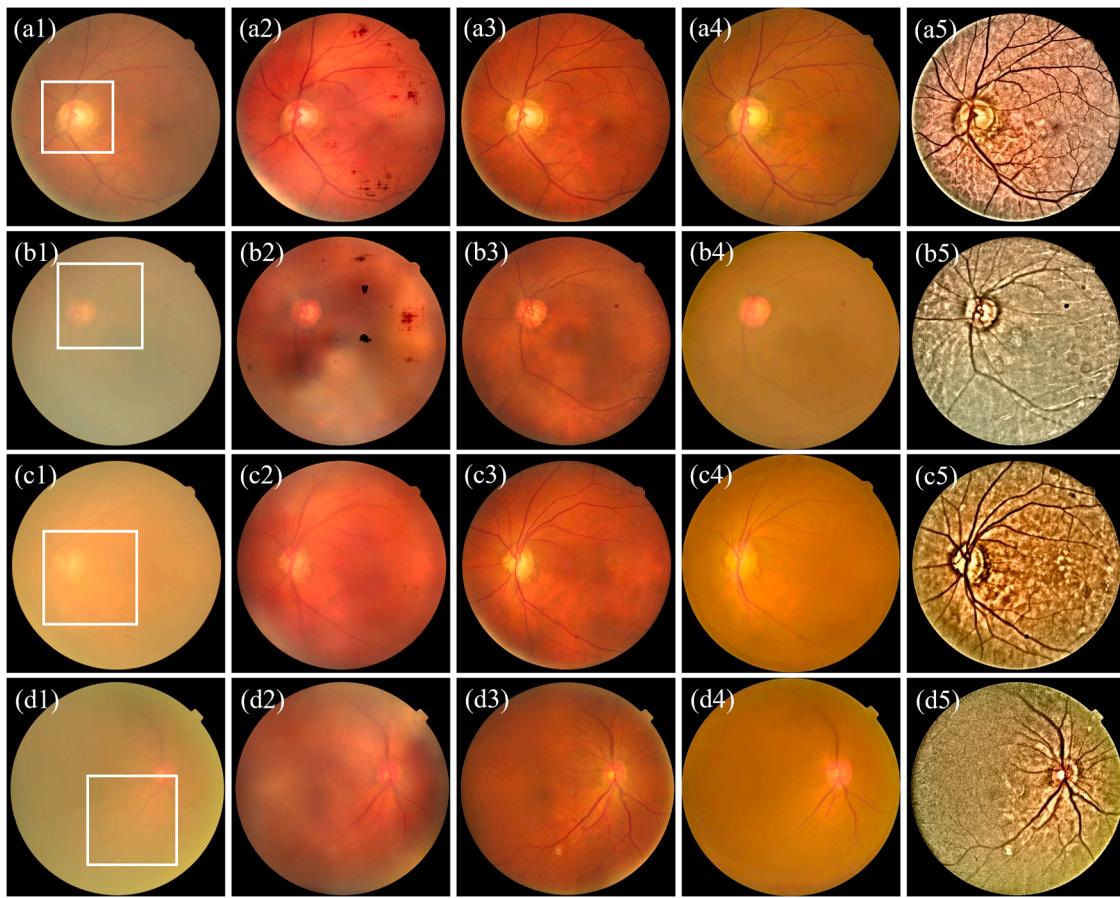


Fig. 12. Comparison against SOTA deep learning methods. (a1)-(d1) raw images. (a2)-(d2) outputs of ArcNet. (a3)-(d3) outputs of ScrNet. (a4)-(d4) outputs of ArSrNet. (a5)-(d5) output of MUTE. Images are sized to 512 by 512 pixels.

can be error-prone.

Fig. 14(a1) to (c1) demonstrates the raw image and segmentation results given by two filtering-based methods including the Frangi filter (Frangi et al., 1998) [Fig. 14(b1)] and the matched filter (Chakraborti et al., 2015) [Fig. 14(c1)]. As shown in Fig. 14(b1) and Fig. 14(c1), the blood vessels have a very weak response to the filters due to the hazy effect, resulting in the failure of blood vessels segmentation.

Without other image preprocessing, enhancing the raw image using the DVA, DRC, LPAR, and MUTE methods before segmenting increases the performance of segmentation results in different degrees. The results for DVA and DRC are still limited as shown in Fig. 14(b2) to (c3), while the two methods are able to detect the vessels' brunches for image enhanced by LPAR and MUTE as shown in Fig. 14(b4) to (c5).

The response of blood vessels to the Frangi filters in Fig. 14(a5) is higher than that of Fig. 14(a5) as the intensity of blood vessels in Fig. 14(b5) are higher than that in Fig. 14(b4). This implies that the segmentation performance in Fig. 14(a5) is better than Fig. 14(a4). Note that the Frangi filter fails to detect the blood vessels near the edge of the ROI for LPAR enhanced image due to the halo effect as shown in Fig. 14(a4).

Another clinical application is shown in Fig. 15, reporting a cataractous retinal image with poor illumination conditions. Both Frangi filter and match filter are able to detect more blood vessels' brunches for Fig. 15(a5) than other enhanced images shown in Fig. 15(b1) to (c4).

As there is no ground-truth blood vessel map for real cataractous retinal images since manual blood vessel segmentation is extremely hard due to the lack of image contrast for cataractous retinal images, we can only present the visual assessment to compare the segmentation results. In order to quantitatively present that MUTE benefits blood vessel segmentation, we perform simulation experiments based on simulated cataractous retinal images with well-defined ground-truth blood vessel

maps in **Supplementary Note 8** in detail. In general, our experimental and simulation results show that MUTE improves the blood vessel segmentation in both visual and objective aspects.

5.3.2. Retinal image registration

Image registration is an important application as it provides in the fields of computer vision, pattern recognition, and medical image analysis. It aligns two or more retinal images together in the same spatial axis to provide an overall comprehensive understanding. A promising retinal image registration relies on precise feature detecting and matching for images to be registered. Registration of cataractous retinal images can be failed due to poor feature paring results as the features are hidden by the haze effect of cataracts.

Fig. 16 compares the registration results before and after being processed by MUTE. Fig. 16(a) is the retinal image before cataract surgery, while Fig. 16(b) is the enhanced image. Fig. 16(c) is the retinal image after cataract surgery. We registered retinal image using Gaussian Field Estimator with Manifold Regularization (GFMR) method (Wang et al., 2019), Fig. 16(d) shows the feature matching results between images in Fig. 16(a) and Fig. 16(c). The feature matching results are problematic as the algorithm failed to find the correct paired features in both images (indicated by green lines), resulting in a distorted registration as shown in Fig. 16(f). Fig. 16(e) shows the feature matching results between images in Fig. 16(b) and Fig. 16(c) where the algorithm pinpoints the paired features as indicated by green lines. Successfully registered images are shown in Fig. 16(g).

Fig. 17 shows the comparison of registration results for images enhanced by different methods. The GFMR can find sufficient paired features to perform registration based on MUTE, LPAR, and DVA outputs, however, according to the visual assessment in Fig. 17(a3) to

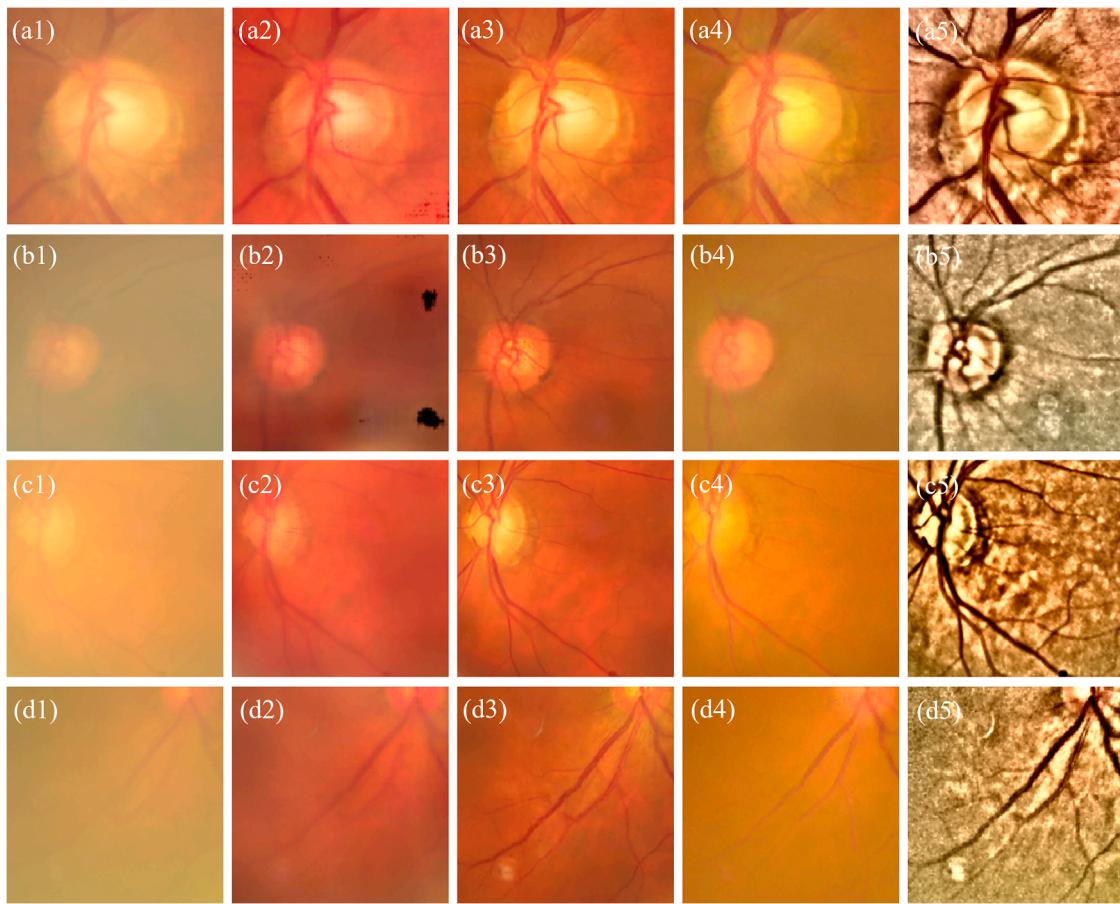


Fig. 13. Zoomed-in images for the region in the white boxes in Fig. 12. (a1)-(d1) raw images. (a2)-(d2) outputs of ArcNet. (a3)-(d3) outputs of ScrNet. (a4)-(d4) outputs of ArSrNet. (a5)-(d5) output of MUTE.

Table 4

Quantitative assessment for retinal image enhancement results in Fig. 12. σ is the noise level in Section 7.

Raw image	Matrixes	Raw	Methods			
			ArcNet	ScrNet	ArSrNet	MUTE
Fig. 12(a1)	UIISM↑	1.1599	1.7425	1.8711	1.9078	3.8646
	UIConM↑	0.0797	0.0826	0.1063	0.0727	0.3476
	IE↑	5.7880	5.9789	6.3074	5.8403	7.3208
	C _{RAMM} ↑	2.0155	2.0973	2.2553	1.9387	7.7648
	FADE↓	0.3924	0.2695	0.2837	0.2841	0.1352
	$\sigma \downarrow$	0.5862	0.4426	0.9217	1.0030	3.5575
Fig. 12(b1)	UIISM↑	1.1188	1.5567	1.3384	1.4918	4.0252
	UIConM↑	0.0946	0.0794	0.0736	0.0757	0.3265
	IE↑	5.2076	5.8317	6.0489	5.5855	7.2288
	C _{RAMM} ↑	2.1192	1.9930	1.8268	1.8556	7.2337
	FADE↓	0.5572	0.3092	0.3629	0.3070	0.1489
	$\sigma \downarrow$	0.5725	0.3590	0.6832	1.1588	1.3042
Fig. 12(c1)	UIISM↑	1.3660	1.6098	1.3196	1.5484	4.2620
	UIConM↑	0.0928	0.0760	0.0721	0.0773	0.2999
	IE↑	5.5182	5.7829	6.2343	5.4319	7.0442
	C _{RAMM} ↑	2.0319	1.8701	1.7859	1.7791	5.8601
	FADE↓	0.7310	0.3168	0.3788	0.3362	0.1505
	$\sigma \downarrow$	0.6837	0.4354	0.5644	1.0740	2.9904
Fig. 12(d1)	UIISM↑	1.1144	1.5886	1.9574	1.2968	3.6856
	UIConM↑	0.0819	0.0658	0.1019	0.0676	0.2991
	IE↑	5.2409	5.5612	6.4547	4.8506	7.0558
	C _{RAMM} ↑	1.9275	1.8016	2.1814	1.6819	5.9793
	FADE↓	0.6085	0.3310	0.3409	0.3881	0.1842
	$\sigma \downarrow$	0.5604	0.3427	0.7253	0.8832	1.5557

Table 5

Quantitative assessment for retinal image enhancement results in cataract dataset. The images are resized to 512 by 512 pixels.

Databases	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Cataract (100 images)	UIISM↑	1.2050	1.7907	1.7668	1.7658	3.7250
	UIConM↑	0.1054	0.1609	0.2522	0.2260	0.3782
	IE↑	0.0691	0.0772	0.0876	0.0678	0.3315
		±	±	±	±	±
		0.0163	0.0066	0.0091	0.0082	0.0224
		5.8285	5.8318	6.2697	5.8190	7.2347
C _{RAMM} ↑	IE↑	±	±	±	±	±
		0.4496	0.1535	0.1655	0.3155	0.1682
		1.8163	2.0137	2.1168	1.8042	7.1012
		±	±	±	±	±
		0.2616	0.1099	0.1669	0.1519	0.8085
		0.4481	0.2827	0.3023	0.3054	0.1460
FADE↓	IE↑	±	±	±	±	±
		0.1050	0.0230	0.0253	0.0295	0.0143
		0.5825	0.4384	0.8012	0.9835	3.6337
		±	±	±	±	±
		0.0577	0.0762	0.1473	0.0922	1.3543

[Fig. 17\(d4\)](#) the registration accuracy for MUTE is better than other methods. For example, as shown in [Fig. 17\(a4\)](#) to [Fig. 17\(c4\)](#), the blood vessels are correctly registered for MUTE [[Fig. 17\(a4\)](#)] where the blood vessels overlapped in both enhanced image and post-surgery images, while the blood vessel is not completely overlapped for LPAR, DVA, and

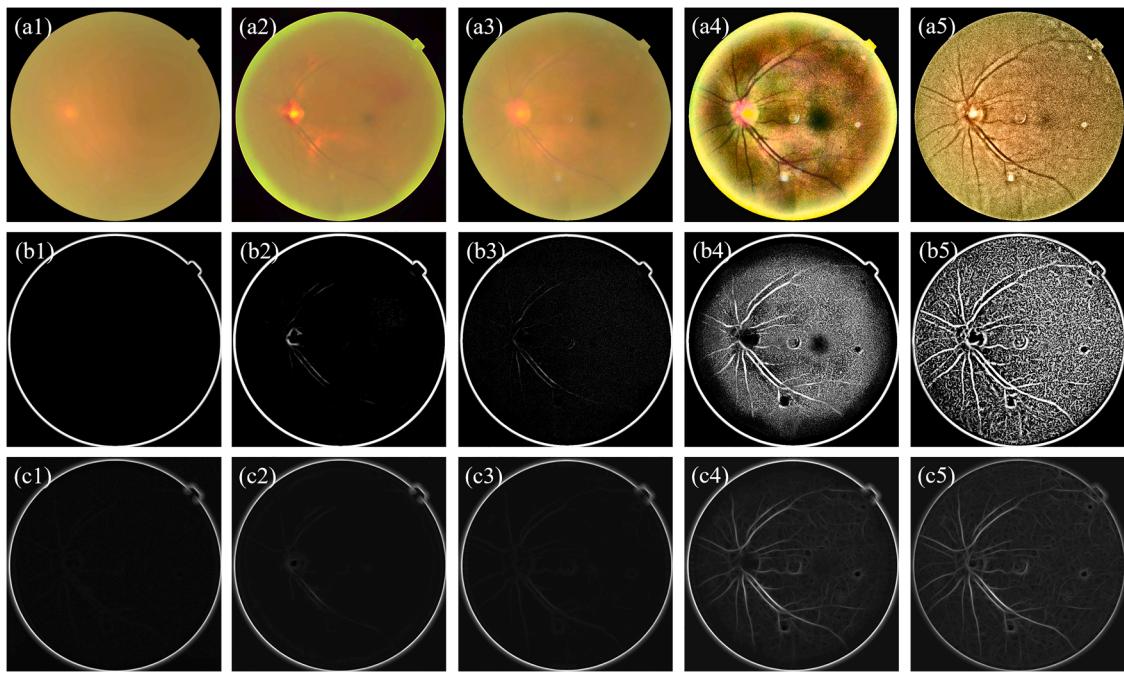


Fig. 14. Demonstration of retinal blood vessels segmentation for the cataractous retinal image. (a1) to (a5) are raw, and enhanced images by DVA, DRC, LAPR, and MUTE, respectively. (b1) to (b5) are the segmentation results of the Frangi filter. (c1) to (c5) are results for the Match filter.

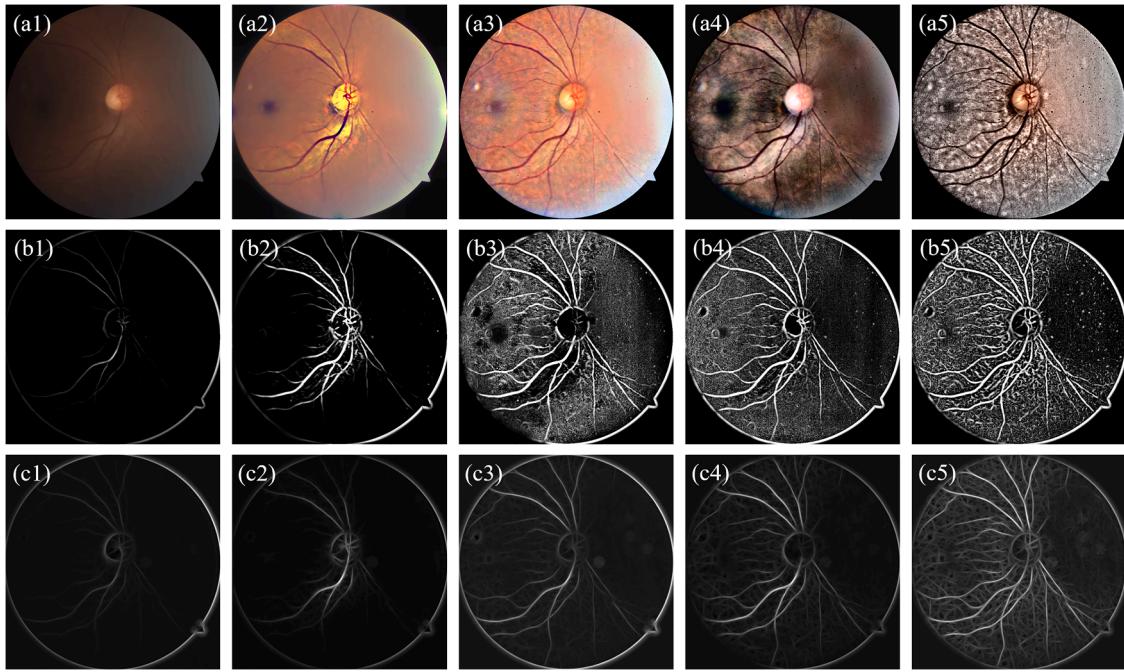


Fig. 15. Demonstration of retinal blood vessels segmentation for cataractous retinal image with poor illumination condition. (a1) to (a5) are raw, and enhanced images by DVA, DRC, LAPR, and MUTE, respectively. (b1) to (b5) are segmentation results of the Frangi filter. (c1) to (c5) are results for the Match filter.

DRC as shown in Fig. 17(b4) to (d4), respectively.

As there is no ground-truth registration image to quantitatively evaluate the registration quality, we perform a simulation experiment in **Supplementary Note 9** in detail, and the results show that MUTE improves the registration quality also in objective aspects.

In general, MUTE enhanced the contrast of the image by suppressing the haze effect of cataractous retinal images. As more retinal structures are revealed during the dehazing process, more paired features can be detected by GFMR, increasing the success rate of cataractous retinal

image registration.

5.3.3. Retinopathy diagnosis

Finally, we focus on the enhancement results in the areas that indicate retinopathy like hard exudates and hemorrhages. The enhancement results should increase the visual quality of the retinopathy area while do not create unexpected artifacts to guarantee structure fidelity. We first test MUTE on the DiaRetDB1 dataset ([DIARETDB1 - Standard](#)), since the retinal images in this dataset are not as severely degraded as

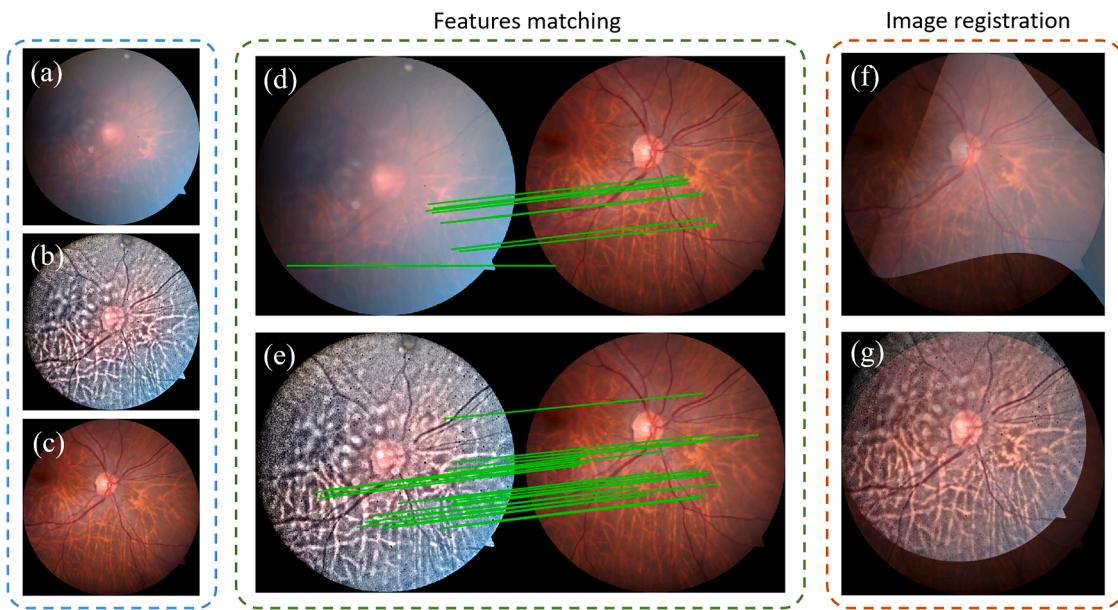


Fig. 16. Demonstration of retinal image registration with and without enhanced by MUTE. (a) to (c) are raw images, enhanced images, and post-surgery images, respectively. (d) indicates the matched features for (a) and (c), while (e) is the matched features for (b) and (c). (f) and (g) are montages of retinal image registration.

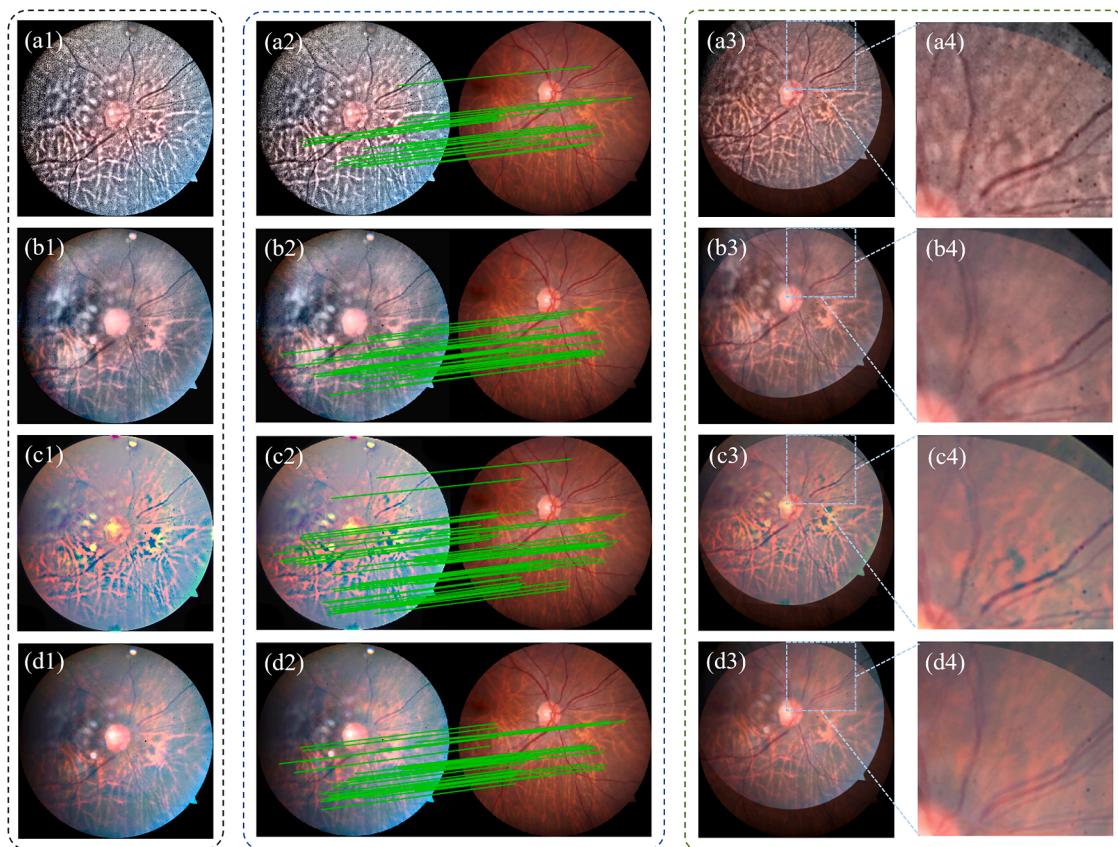


Fig. 17. Comparison of registration results for (a1-a3) MUTE, (b1-b3) LPAR, (c1-c3) DVA and (d1-d3) DRC methods. (a1), (b1), (c1), and (d1) are enhancements. (a2), (b2), (c2), and (d2) are paired features (a3), (b3), (c3), and (d3) are registration results. (a4) to (d4) are zoomed-in part in the blue boxes.

that of cataractous retinal images, and retinopathies are already marked by human specialists shown in Fig. 18(b) and Fig. 18(c). As shown in Fig. 18(d1) to (d4), MUTE increases the visual quality of the hard exudates. Some hard exudates which can hardly be observed in the raw image shown in Fig. 18(d3) can be clearly found in the enhanced images

in Fig. 18(d4) due to the increase of image contrast. MUTE also increases the visual quality of hemorrhages as shown in Fig. 18(e1) to (e4), as the contrast between hemorrhage areas and the background increases in enhanced images [see Fig. 18(e2) and Fig. 18(e4)].

Then we tested MUTE on cataractous retinal images before and after

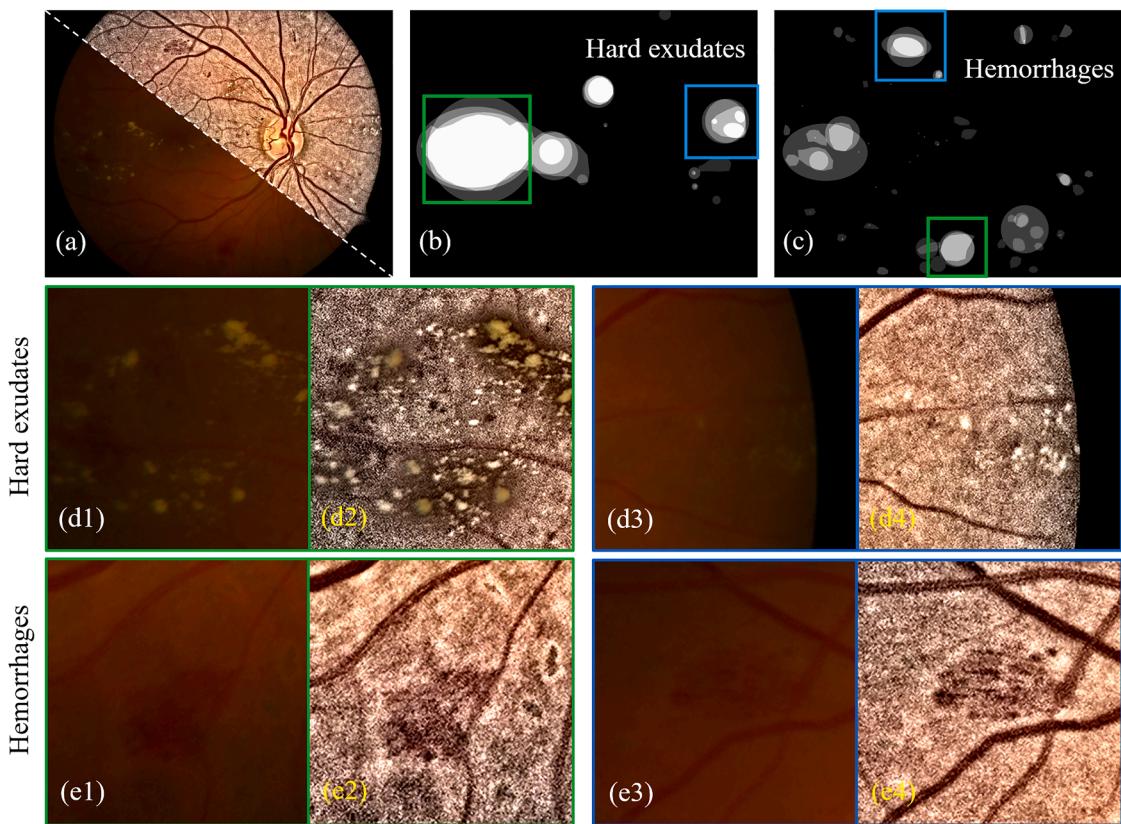


Fig. 18. Enhancement of retinopathy areas of MUTE. (a) Montage of raw and enhanced images. (b) labels of hard exudates areas. (b) labels of hemorrhage areas. (d1) to (d4) are enlarged parts of raw and enhanced images corresponding to green and blue boxes in (b). (e1) to (e4) are enlarged parts for (c).

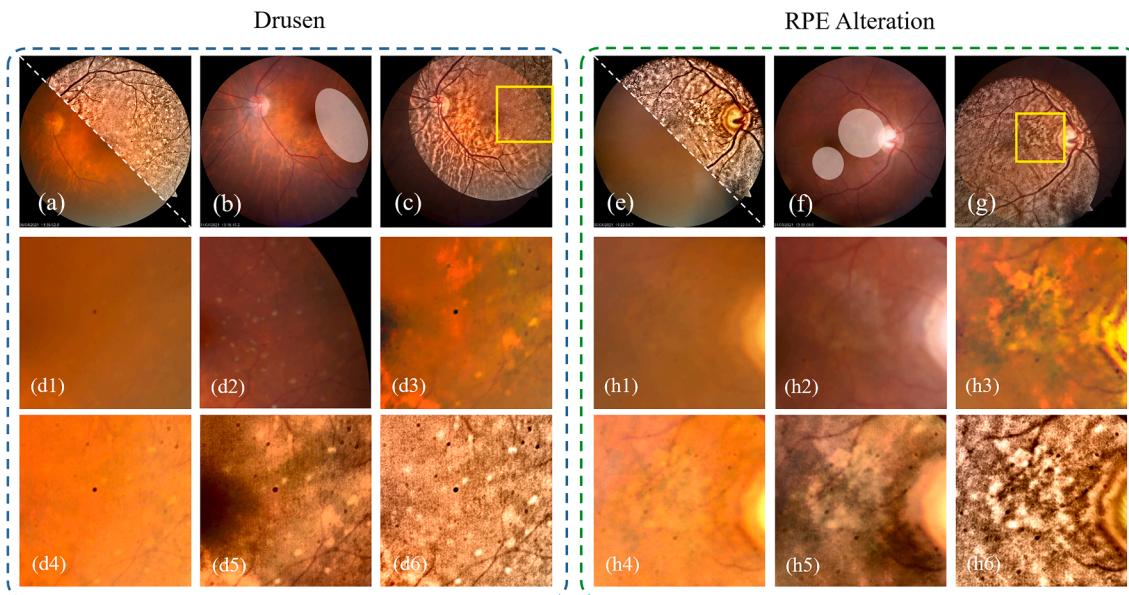


Fig. 19. Demonstration of enhancement for cataractous retinal images, with Drusen and RPE alteration. (a) and (e) are Montages of raw and enhanced images by MUTE. (b) and (f) are post-surgery retinal images marked by specialists. (c) and (g) denotes the image registration between post-surgery images and enhanced images. (d1) to (d6) are enlarged parts in the yellow box in (c) for raw image, and post-surgery image, enhanced by DVA, DRC, LPAR and MUTE respectively. (h1) to (h6) are zoomed-in for yellow box in (g).

cataract surgery. The retinopathy is manually marked based on the post-surgery retinal images as shown in Fig. 19(b) and Fig. 19(f) for Groups Drusen and RPE Alteration, respectively. After image registrations, we enlarged the area in the yellow box where there is retinopathy (Drusen, and RPE alteration in these cases) for raw images and enhanced images.

As shown in Fig. 19(d1), the drusen in the raw images have visual quality due to cataracts, while it can be clearly observed in the post-surgery image shown in Fig. 19(d2), and so does the RPE alteration shown in Fig. 19(e) and Fig. 19(f). After image enhancement by different methods, the visual quality of the retinopathy areas increases, and the

MUTE gains a better performance than the other three methods as it significantly increases the image contrast which has also been shown in Table 2 and Table 3.

MUTE increases the contrast of cataractous retinal images which may potentially facilitate the detection of retinopathy. Meanwhile, by comparing MUTE against the raw image, post-surgery image, and enhanced image by the other three methods, we found that MUTE won't lead to additional artifact structure on enhanced images and guarantees the information fidelity.

6. Ablation study

In this subsection, we perform ablation studies on the number of denoising layers, the type of denoising layers, the selection of I_0 values, the function of illumination correction, and the treatment in Pb/Pr channels.

6.1. On the denoising layers

Fig. 20 shows dehazing results of the image in Fig. 3(d3) with different total numbers of Gaussian denoising levels. More levels give a more refined estimation of T_{sc} but increase the computational complexity. The visual quality of the dehazed image in Fig. 20(b2) is better than that in Fig. 20(d2), which still has a slight haze effect. The cost function for Fig. 20(d2) is oscillating and decreasing with respect to iteration numbers. This means that the optimization process for Fig. 20(d2) is unstable due to large numbers of undermined parameters. According to the experimental trial, we found that $\mathcal{L} \in [4, 6]$ was appropriate for most of the cataractous retinal images.

Our proposed dehazing method is not sensitive to the choice of \mathcal{I}_l . One can use total variance regularization (Rudin et al., 1992) and its variants, or image gradient L_0 -norm regularization (Xu et al., 2011) for alternative methods when calculating \mathcal{I}_l . Supplementary Note 11

compares the results of MUTE with different denoisers including Domain transform recursive filter and total-variations. The intermedia output of the parameter \mathcal{A} and $\varphi_l(\mathbf{r})$ are also shown in detail.

In an extreme case, we can even simply let $\mathcal{L} = 1$ and $\mathcal{I}_0(\mathbf{r}) = \mathcal{I}_1(\mathbf{r}) = 1$ since the uniform pattern can be regarded as an extremely denoised image. In this case, as shown in Fig. 21, the adaptive pixel-wise stimulated algorithm still returns an appropriate estimation of T_{sc} together with dehazed \mathbf{O}_Y .

6.2. On the selection of I_0

To show the reason why $I_0 = 0.52$ is used in our experiments, we perform a statistical analysis of the Y channel average pixel value for raw images, images after illumination correction, and final outputs of MUTE. The analysis is applied to 194 cataractous retinal images (100 images in the CATARACT dataset, and 95 images in our proprietary dataset), and 1200 retinal images from the MESSIDOR dataset.

As shown in Fig. 22(a1) and Fig. 22(a2), the intensity of raw images from both the cataract dataset and the MESSIDOR dataset varies from 0.2 to 0.6. The normalized probability density functions (PDF) are fitted with Gaussian distribution (plotted in red curves) which has large standard variations. The distribution pattern implies the fact that not all retinal images are collected under good conditions, some of them suffer from bad and insufficient illumination problems.

The distribution patterns after the illumination correction process are shown in Fig. 22(b1) and (b2), where the distributions are concentrated near 0.95. Since the illumination correction is performed in an adaptive manner, it is suitable for retinal images of different illumination conditions. Moreover, the adaptive illumination correction process has a similar function as intensity normalization - that is adjusting the intensity of raw images to similar levels. This feature will benefit the subsequent dehazing process.

As the dehazing process is performed based on the intensity cost

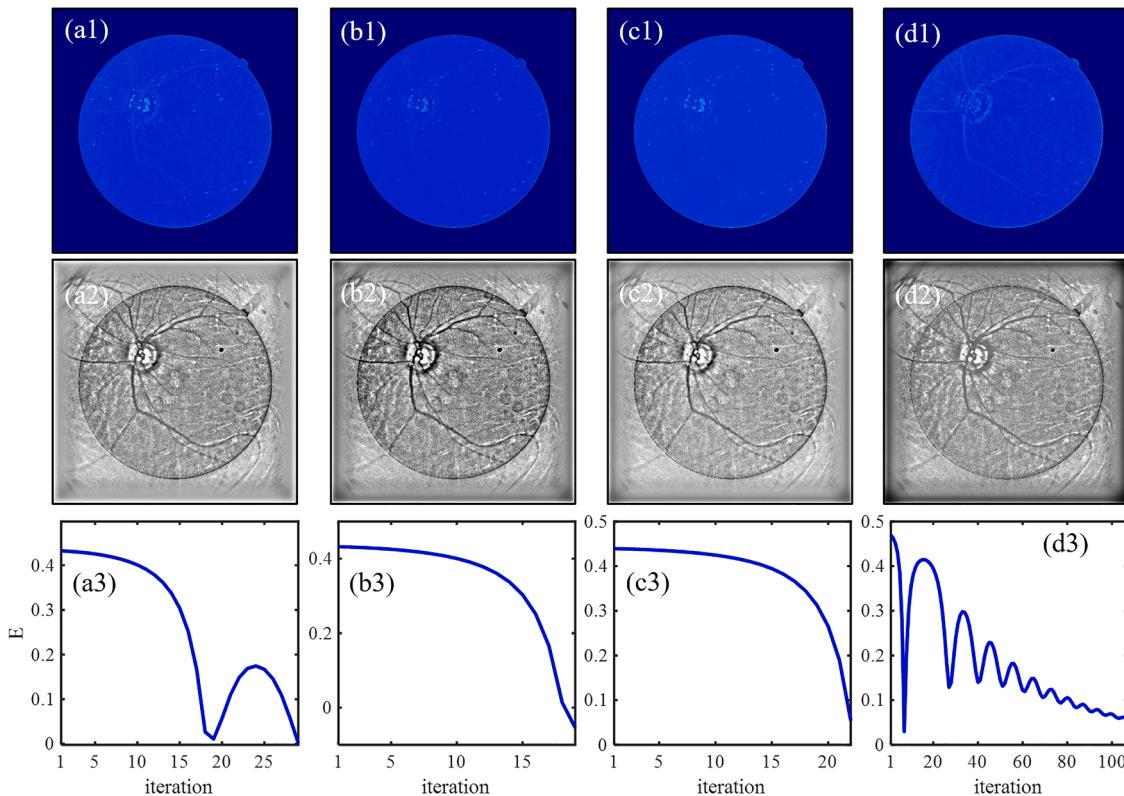


Fig. 20. Dehazing of Fig. 3(d3) with different levels of Gaussian denoising. (a1) to (d1) are T_{sc} for $\mathcal{L} = 2$, $\mathcal{L} = 4$, $\mathcal{L} = 6$, and $\mathcal{L} = 8$. (a2) to (d2) are corresponding dehazed Y channel. (a3) to (d3) are evolution of cost function.

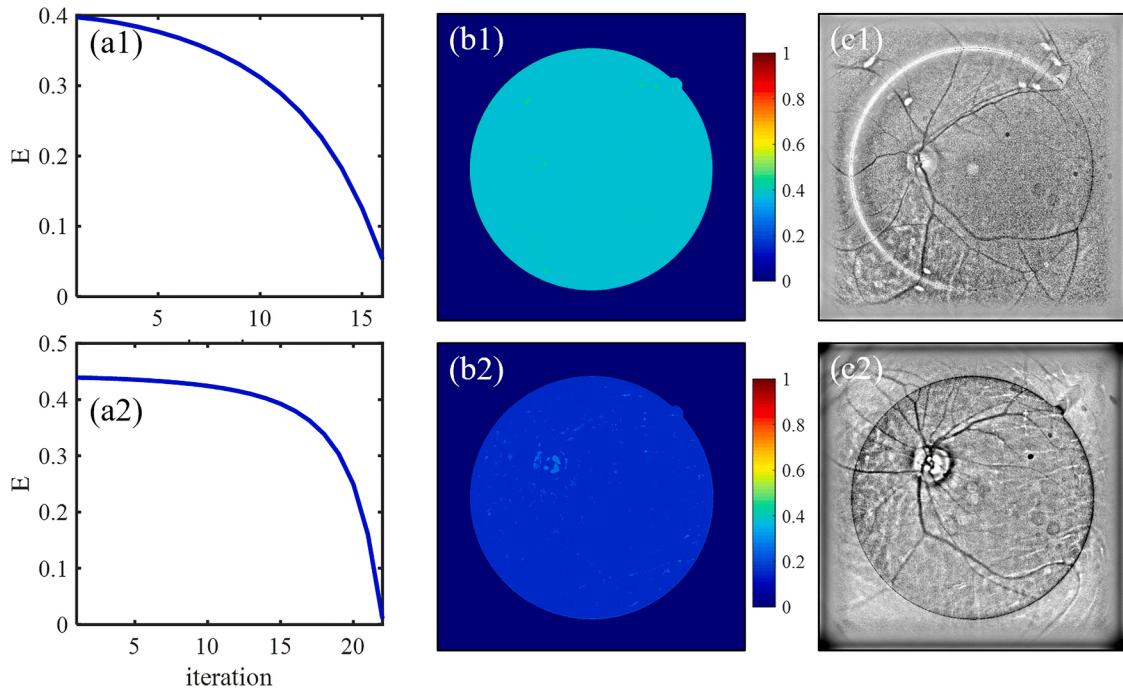


Fig. 21. Results for Y channel dehazing when $\mathcal{L} = 2$ and $\mathcal{S}_1(r) = \mathcal{S}_2(r) = 1$ for the image in Figs 3(c3) and (d3). (a1) and (a2) are evolution of cost functions. (b1) and (b2) are T_{sc} . (c1) and (c2) are O_Y .

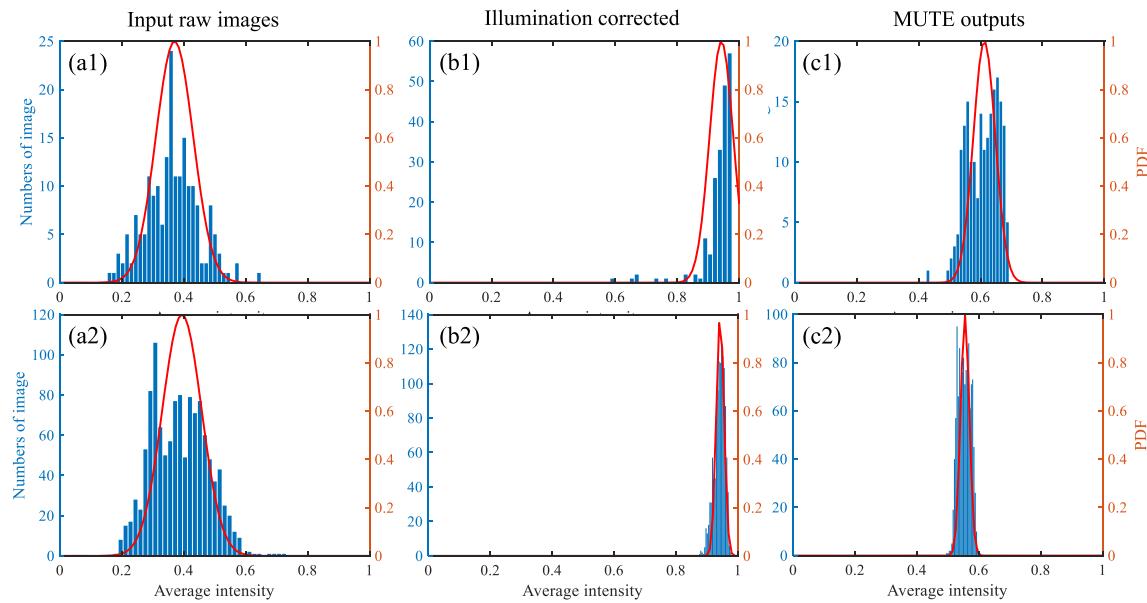


Fig. 22. Distribution of Y channel average value for (a1), (a2) raw images; (b1), (b2) illumination corrected images, and (c1), (c2) MUTE output images. (a1) to (c1) are results for cataractous retinal images (194 images included). (a2) to (c2) are results for MESSIDOR dataset (1200 images included).

function and the dehazing always decreases the intensity level, we consider the target intensity level being a constant ratio I_0 between the corrected intensity level and the final output intensity level. As shown in Fig. 22(a1) and Fig. 22(a2) the average Y channel value for the raw image approaches 0.4, and the average Y channel value for the illumination corrected images approach 0.95, we, therefore, can let $I_0 \approx 0.4 / 0.95 = 0.4211$.

In our experiment, we choose $I_0 = 0.52$ to let the intensity level for the output images becomes larger than 0.4211. The distribution for the Y channel average value of MUTE outputs is shown in Fig. 22(c1) and (c2). Moreover, Fig. 20 shows the dehazed results with different values of I_0 .

In general, a small value of I_0 increases the degree of dehazing as the hazed effect is significantly suppressed. However, the image intensity also decreases. According to the experimental results in Fig. 23 and statistical analysis, we believe that $I_0 = 0.52$ is an appropriate setting.

Finally, according to Eq. (3), the Pb/Pr channel can be restored through an illumination correction process with known of $L(r) \cdot T_{lens}^2(r)$ from illumination correction in Section 3.2, and $T_{sc}^2(r)$ from dehazing in Section 4. According to the experiments (**Supplementary Note 5**), the Pb/Pc treatment can be skipped since this process won't influent the final dehazing quality of MUTE. Please refer to **Supplementary Note 5** for detailed experiment results.

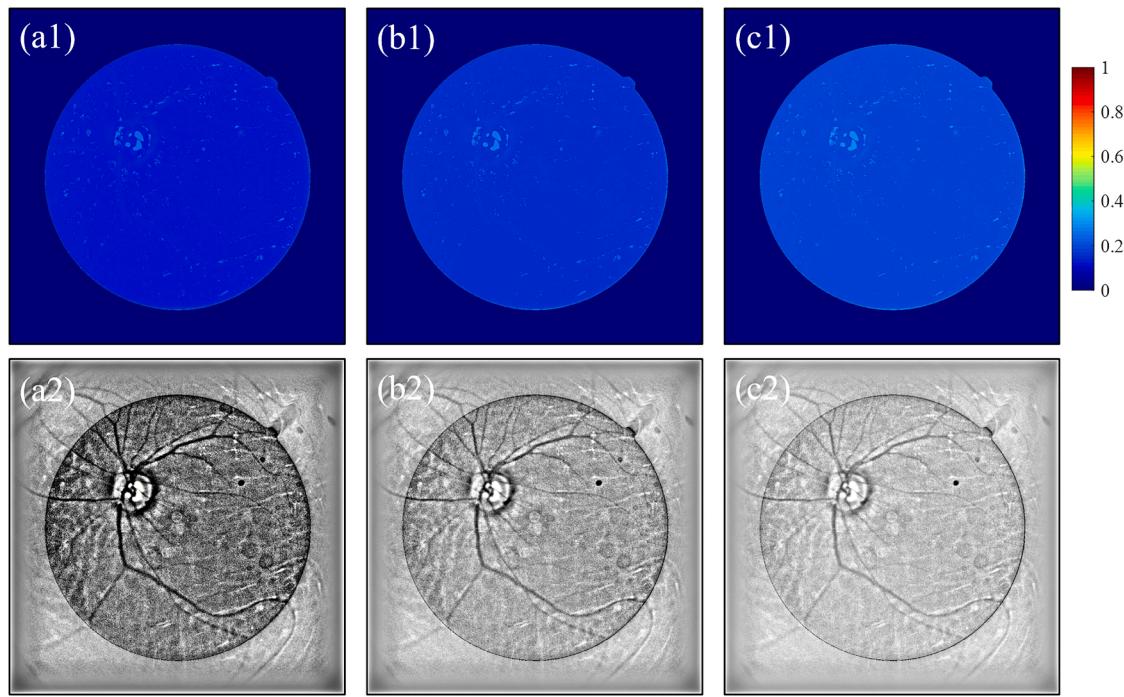


Fig. 23. Results for Y channel dehazing when $\mathcal{L} = 4$ with different values of I_0 (a1) to (c1) are T_{sc} with $I_0 = 0.35$, $I_0 = 0.55$, and $I_0 = 0.65$, respectively. (a2) to (c2) are dehazed Y channels.

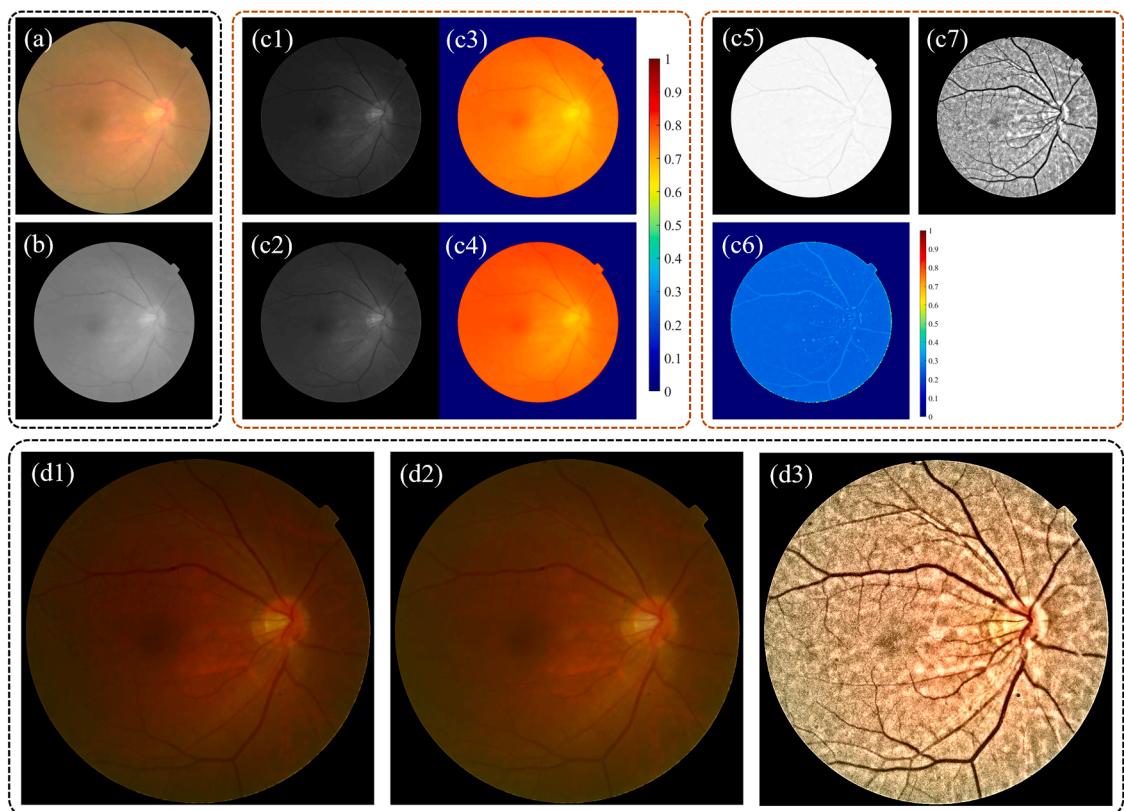


Fig. 24. MUTE dehazing output without and with illumination correction process. (a) raw input image. (b) Y channel of (a). Without the illumination correction process, (c1) and (c3) are dehazed Y channel and T_{sc} with $I_0 = 0.1$. (c1) and (c3) are dehazed Y channel and T_{sc} with $I_0 = 0.2$. The MUTE outputs are in (d1) and (d2). With illumination correction, the illumination-corrected Y channel, T_{sc} and dehazed Y channel are shown in (c5), (c6), and (c7), respectively. (d3) is the corresponding MUTE output.

6.3. On the necessity of illumination correction

As shown in [Section 6.2](#), the illumination correction will normalize the intensity of various input retinal images into similar intensity levels approach 0.95. This high-intensity level gives sufficient intensity adjusting tolerance for the dehazing process since image dehazing always decreases the intensity level of the image. As such, the illumination correction benefits MUTE dehazing. If we skip the illumination correction process, the MUTE may not output good dehazing retinal images.

For example, if the input retinal images have low-intensity levels, the dehazing effect is not significant and may cause bad performance since the intensity level of input raw images is already very low. Further dehazing will decrease the image intensity and will distort image colors.

For retinal images that have relatively good illumination conditions, if we skip the illumination correction, we need to manually adjust parameter I_0 according to individual input images to optimize the dehazing results since the average intensity of the Y channel varies from 0.2 to 0.6. For example, [Fig. 24](#) shows the dehazing results without illumination correction. The input image is shown in [Fig. 21\(a\)](#) which has good and uniform illumination conditions. Its Y channel is shown in [Fig. 21\(b\)](#) where the mean value is 0.5357. [Fig. 24\(c1\)](#) and [Fig. 24\(c2\)](#) show the dehazed Y channel and T_{sc} for $I_0 = 0.1$ and $I_0 = 0.2$, respectively. The MUTE outputs are shown in [Fig. 24\(d1\)](#) for $I_0 = 0.1$ and [Fig. 24\(d2\)](#) for $I_0 = 0.2$, where dehazing effects for both images are not significant, as the MUTE overestimates the transmission matrix T_{sc} as shown in [Fig. 24\(c3\)](#) and [Fig. 24\(c4\)](#).

With illumination correction as shown in [Fig. 24\(c5\)](#), the Y channel is

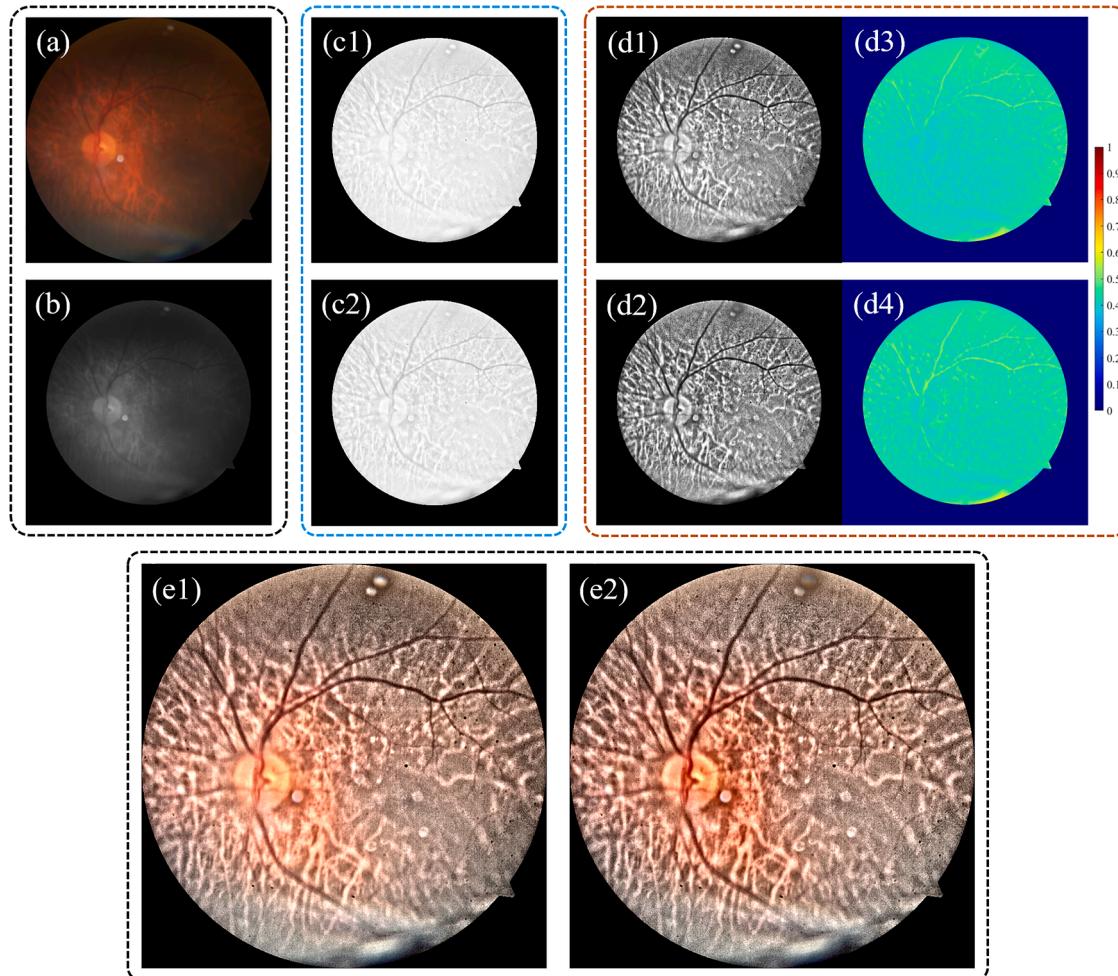


Fig. 25. MUTE dehazing output without and with coarse illumination correction process. (a) raw input image. (b) Y channel of (a). (c1) and (c2) are outputs of illumination correction without and with coarse illumination correction. (d1) to (d4) are MUTE dehazing for (c1) and (c2). (e1) and (e2) are final outputs.

at a high-intensity level where the mean value is 0.9721 and has a sufficient intensity range for the dehazing process to darken the image. The transmission map can be correctly estimated as shown in [Fig. 24\(c6\)](#), $I_0 = 0.52$. The dehazed Y channel is shown in [Fig. 24\(c7\)](#). The final output is shown in [Fig. 24\(d3\)](#), where the contrast of the image is improved.

As proposed in [Supplementary Note 2.2](#), our adaptive illumination correction contains two steps: coarse illumination correction and fine illumination correction. The first step uses Retinex to correct the uneven illumination pattern so that the illumination pattern over the entire image is flat. Without the first step, the dehazing quality of MUTE will be degraded for retinal images that have uneven illumination conditions.

For example, as shown in [Fig. 25\(a\)](#) the raw image has an uneven illumination pattern where the intensity near the optical disk is higher than the areas far from the optical disk. [Fig. 25\(c1\)](#) and [\(c2\)](#) show the output of intensity correction without and with coarse illumination correction. Although the intensity level can still be corrected based on the fine illumination correction step, we can still find that the intensity level near the optical disk in [Fig. 25\(c1\)](#) is higher than the intensity in other places, and so does [Fig. 25\(d1\)](#) for dehazed Y channel. The final outputs are shown in [Fig. 25\(e1\)](#) and [\(e2\)](#) without and with coarse illumination correction, where the dehazing quality near the optical disk in [Fig. 25\(e2\)](#) is better than that of [Fig. 25\(e1\)](#). In general, the illumination correction process is essential for MUTE dehazing, while each step of illumination correction benefits the dehazing quality of MUTE.

7. Suppressing noise effect of MUTE using Hessian regularization

For retinal images that have poor illumination and heavy haze effects, the enhancement algorithm increases the contrast of the image while also enlarging the impact of noises simultaneously. A denoising process can be involved after Eq. (8) is finished. For biologic images, we would recommend using L₁-norm Hessian (second-order gradient) regularization for image denoising, which is given by minimizing the following equation

$$\mathbf{O}_{dn} = \text{argmin} \| \mathbf{O} - \mathbf{O}_{dn} \|_2^2 + \lambda (\| \nabla_{xx} \mathbf{O}_{dn} \|_1 + \| \nabla_{yy} \mathbf{O}_{dn} \|_1 + 2 \| \nabla_{xy} \mathbf{O}_{dn} \|_1) \quad (19)$$

where \mathbf{O} is the image calculated by Eq. (8), and \mathbf{O}_{dn} is the denoised image. The TV regularization is not recommended as it is calculated from the first-order partial derivatives and might over-sharpen the boundary between different regions while reducing the differences within a single region, which would result in staircase-like and aberrant reconstructed images.

In this section, we quantitatively analyze the noise level of MUTE output and propose using Hessian regularization to suppress the noise signals of MUTE. To quantitatively measure the noise level without a reference image, we use the global noise estimation (Immerkær, 1996; Wang et al., 2021):

$$\sigma = \sqrt{\frac{\pi}{2}} \frac{1}{6(M-2)(N-2)} \sum_{x,y} |\mathbf{N} \otimes \mathbf{O}_Y|, \quad \mathbf{N} = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}, \quad (20)$$

where \mathbf{O}_Y denotes the input Y channel. A larger value of σ denotes stronger noise levels.

We use Eq. (20) to measure the noise level of all four methods (DVA, DRC, LPAR, and MUTE) among cataract retinal images (100 images from the CATARACT dataset and 94 images of our proprietary) and the MESSIDOR retinal images (1200 images), the results are listed in Table 6.

As shown in Table 6, the MUTE can also enlarge the noise signals for retinal images corrupted by poor illumination and heavy hazy effects. Therefore, a denoising process can be applied to the MUTE output. Here we choose L₁-norm Hessian regularization for image denoising. The denoising process is only applied to the Y channel since Pb and Pr channels remain unchanged.

As the λ is the key parameter that controls the denoising level and data fidelity, it should be chosen carefully based on the noise level. We consider Eq. (20) as a reference value, and test $\lambda = \sigma/10$, $\lambda = \sigma/20$, and $\lambda = \sigma/30$ for the 194 cataract retinal images. Again, we use IE, Cramm, UISM, UIConM, and FADE to measure the dehazing performance. Results are shown in Table 7.

As listed in Table 7, the Hessian regularization significantly suppresses the noise level of MUTE outputs. When $\lambda = \sigma/30$ the noise level

Table 6
Noise level for 4 methods on cataract retinal image dehazing.

Databases mean \pm std.	Noise level ↓	Raw	Methods			
			DVA	DRC	LPAR	MUTE
Cataract (100 images)	σ	0.5127	0.8148	1.6710	8.5224	11.7200
	\pm	\pm	\pm	\pm	\pm	± 3.0493
Proprietary (94 images)	σ	0.6979	1.1143	2.1283	6.0175	16.0702
	\pm	\pm	\pm	\pm	\pm	± 4.7823
MESSIDOR (1200 images)	σ	0.5445	1.4447	1.8126	5.1842	8.5259
	\pm	\pm	\pm	\pm	\pm	± 2.4833
		0.1710	0.4177	0.6255	1.8195	

of denoised MUTE is less than the original MUTE by 3-fold, and is less than LAPR (Table 4) by 2-fold in the CATARACT dataset. The dehazing performance is also maintained, for example, the FADE score is less than 0.2 for both CATARACT and our proprietary datasets.

We compared the dehazing results for MUTE without and with the denoising process against other methods on 2116 non-cataract retinal images. $\lambda = \sigma/30$ for the Hessian regularization. As listed in Table 8, the Hessian denoising decreases the noise level of MUTE. The noise level is smaller than other methods while the dehazing qualities including UISM, UIConM, Cramm, and FADE scores for Hessian denoised MUTE are maintained as it outperforms other state-of-the-art methods and approaches to un-denoised MUTE.

Fig. 26 shows visual assessments for two cataractous retinal images. As shown in the fourth column of Fig. 26, the Hessian regularization suppresses the noise signals, and the dehazing quality is still similar to those without the denoising process shown in the third column of Fig. 26.

8. Concluding remarks

8.1. Discussion

Due to the existence of the ocular lens and absorption of biological tissues, retinal images have much higher intensity in red/green channels than that in blue channels. Algorithms for cataractous retinal image dehazing should be designed under consideration of the spectral feature of retinal images.

In this research, we applied the YPbPr color space transformation on the double-pass fundus reflection (DPFR) model and obtain the YPbPr-DPFR model, which shows the haze effect of cataracts only impacts the Y channels of the image. For the YPbPr transformation matrix, it assigned the most weight to the green channel and the least weight to the blue channels, which is just corresponding to the prior feature of retinal images that is the green channel contains important retinal structures and the blue channel has less information due to the absorption.

Then we were inspired by the duality between image dehazing and image denoising and proposed a multilevel-stimulated denoising strategy, termed MUTE, that is dedicated to cataractous retinal image dehazing. Through our experiments, we found that preprocessing is important as it eliminates the boundary effect and normalized the intensity levels for retinal images of different problematic illumination conditions.

In general, MUTE has six hyperparameters that need to be manually determined which are I_0 for dehazing degree, and η , α , β , τ , ξ for Adam-gradient descent. However, according to our massive experiment throughout 2318 images, we have found that the default values ($I_0 = 0.52$, $\eta = 0.01$, $\alpha = \tau = 0.9$, $\beta = \xi = 0.99$) given in the main text are appropriate in order to obtain promising enhancement results. Therefore, there is only one parameter I_0 that is needed to be manually adjusted to optimize the performance for MUTE.

The algorithm for MUTE is composed using MATLAB 2018a and executed using a personal laptop with Intel Core i7-8750H CPU and 16 RAM. No GPU acceleration is used. Average executing times for images of different sizes are listed in Table 9. As the image size increases, MUTE takes a longer time to proceed with each image. For image size of 2048 by 2048, it takes about 98.44 s on average for each image. To speed up the MUTE, we propose a down-up-sampling strategy (DUSS) in Supplementary Note 10, which significantly decreases the executing times of MUTE while maintaining the enhancement quality.

8.2. Contributions of MUTE

In this research, we propose the multilevel stimulated denoising strategy (MUTE) to achieve retinal image dehazing through the process of image denoising. The MUTE has the following novelty and

Table 7

MUTE with Hessian regularization denoising on cataractous retinal images.

Databases mean \pm std.	Matrixes	Raw	MUTE, λ value			
			No denoise	$\sigma/10$	$\sigma/20$	$\sigma/30$
Cataract (100 images)	UISM↑	0.9002 \pm 0.1432	6.7884 \pm 0.7869	5.3388 \pm 0.4284	5.5770 \pm 0.3952	5.7044 \pm 0.4106
	UIConM↑	0.0573 \pm 0.0123	0.3890 \pm 0.0093	0.3298 \pm 0.0303	0.3419 \pm 0.0212	0.3467 \pm 0.0189
	IE↑	5.8443 \pm 0.4449	7.4151 \pm 0.1338	7.1078 \pm 0.2255	7.1581 \pm 0.1787	7.1840 \pm 0.1755
	C _{RAMM} ↑	1.2013 \pm 0.1741	8.1483 \pm 0.6035	6.0133 \pm 0.8880	6.3177 \pm 0.8098	6.4882 \pm 0.7938
	FADE ↓	0.4715 \pm 0.1155	0.1469 \pm 0.0228	0.2109 \pm 0.0272	0.1903 \pm 0.0234	0.1790 \pm 0.0215
	$\sigma\downarrow$	0.5127 \pm 0.0724	11.7200 \pm 3.0493	2.4091 \pm 0.1913	2.7044 \pm 0.5011	3.1508 \pm 0.8263
Proprietary (94 images)	UISM↑	1.2065 \pm 0.3028	6.8936 \pm 1.1603	5.8507 \pm 0.5926	6.1099 \pm 0.6901	6.2199 \pm 0.7329
	UIConM↑	0.0446 \pm 0.0133	0.3811 \pm 0.0646	0.3482 \pm 0.0492	0.3548 \pm 0.0497	0.3571 \pm 0.0498
	IE↑	6.8377 \pm 0.5775	7.4889 \pm 0.4049	7.3438 \pm 0.3763	7.3909 \pm 0.3762	7.4144 \pm 0.3760
	C _{RAMM} ↑	1.0378 \pm 0.1741	7.9087 \pm 1.5078	6.1696 \pm 1.2147	6.5879 \pm 1.2283	6.7936 \pm 1.2309
	FADE ↓	1.0455 \pm 0.3905	0.1857 \pm 0.1177	0.2062 \pm 0.0474	0.1865 \pm 0.0473	0.1790 \pm 0.0469
	$\sigma\downarrow$	0.6979 \pm 0.3680	16.0702 \pm 4.7823	3.9869 \pm 1.5204	5.8252 \pm 2.4833	7.2014 \pm 2.9582
Average (194 images)	UISM↑	1.0486 \pm 0.2798	6.8393 \pm 0.9844	5.5869 \pm 0.5737	5.8352 \pm 0.6172	5.9542 \pm 0.6419
	UIConM↑	0.0512 \pm 0.0143	0.3852 \pm 0.0455	0.3390 \pm 0.0418	0.3483 \pm 0.0386	0.3519 \pm 0.0379
	IE↑	6.3164 \pm 0.7181	7.4568 \pm 0.2996	7.2258 \pm 0.3313	7.2745 \pm 0.3161	7.2992 \pm 0.3146
	C _{RAMM} ↑	1.1221 \pm 0.1920	8.0322 \pm 1.1388	6.0915 \pm 1.0642	6.4528 \pm 1.0465	6.6409 \pm 1.0443
	FADE ↓	0.7496 \pm 0.4038	0.1657 \pm 0.0856	0.2086 \pm 0.0386	0.1884 \pm 0.0373	0.1790 \pm 0.0364
	$\sigma\downarrow$	0.5986 \pm 0.2733	12.7902 \pm 3.8432	3.1980 \pm 1.3393	4.2648 \pm 2.3749	5.1761 \pm 2.9691

Table 8MUTE with Hessian regularization denoising $\lambda = \sigma / 30$ on other datasets.

Databases mean \pm std.	Matrixes	Raw	Methods			
			DVA	DRC	LPAR	MUTE
DiaRetDB0+DB1 (219 images)	UISM↑	3.2857 \pm 0.2107	3.5914 \pm 0.5870	4.2797 \pm 0.3261	5.3378 \pm 0.8601	6.7373 \pm 0.3610
	UIConM↑	0.0396 \pm 0.0140	0.1793 \pm 0.0393	0.0989 \pm 0.0389	0.2419 \pm 0.0471	0.3977 \pm 0.0068
	IE↑	5.6481 \pm 0.3389	6.2941 \pm 0.2314	5.7972 \pm 0.3434	5.8912 \pm 1.8016	6.6910 \pm 0.1443
	C _{RAMM} ↑	0.7169 \pm 0.1332	2.4855 \pm 0.7205	1.4165 \pm 0.4100	3.7898 \pm 1.2053	8.0577 \pm 0.5667
	FADE ↓	0.4752 \pm 0.0395	0.4188 \pm 0.0648	0.3094 \pm 0.0461	0.2495 \pm 0.1344	0.1278 \pm 0.0156
	$\sigma\downarrow$	0.6522 \pm 0.0507	0.9402 \pm 0.2096	1.4415 \pm 0.2781	4.6964 \pm 1.5661	12.3944 \pm 1.6480
STARE (397 images)	UISM↑	2.6542 \pm 0.6460	4.1758 \pm 1.0861	4.6212 \pm 0.5306	5.5971 \pm 1.0953	6.5124 \pm 0.8229
	UIConM↑	0.0879 \pm 0.0195	0.2440 \pm 0.0458	0.1930 \pm 0.0444	0.2863 \pm 0.0499	0.3693 \pm 0.0180
	IE↑	5.4185 \pm 0.3468	6.3837 \pm 0.2481	5.5470 \pm 0.3198	5.5085 \pm 1.2873	5.8185 \pm 0.2760
	C _{RAMM} ↑	1.9010 \pm 0.2998	5.1454 \pm 1.4516	3.3505 \pm 0.7190	5.4806 \pm 1.2965	7.8314 \pm 0.7522
	FADE ↓	0.4744 \pm 0.1274	0.3947 \pm 0.1296	0.2367 \pm 0.0617	0.2739 \pm 0.1453	0.1331 \pm 0.0236
	$\sigma\downarrow$	0.8479 \pm 0.1725	1.7854 \pm 0.6727	2.5389 \pm 0.6197	5.1547 \pm 2.0209	8.5333 \pm 2.3205
MESSIDOR (1200 images)	UISM↑	1.8419 \pm 0.3224	4.0235 \pm 0.7075	5.0035 \pm 0.4089	6.2763 \pm 0.6584	7.0831 \pm 0.4933
	UIConM↑	0.0746 \pm 0.0185	0.2383 \pm 0.0318	0.1608 \pm 0.0418	0.2777 \pm 0.0280	0.3660 \pm 0.0073
	IE↑	5.1908 \pm 0.3454	6.2726 \pm 0.1881	5.6377 \pm 0.3846	5.9916 \pm 0.7165	6.0725 \pm 0.2111
	C _{RAMM} ↑	1.1810 \pm 0.1639	3.9134 \pm 0.7737	2.2784 \pm 0.4610	5.7179 \pm 0.7411	7.6652 \pm 0.5981
	FADE ↓	0.4652 \pm 0.0673	0.4361 \pm 0.0842	0.2320 \pm 0.0410	0.1812 \pm 0.0726	0.1252 \pm 0.0136
	$\sigma\downarrow$	0.5445 \pm 0.1710	1.4447 \pm 0.4177	1.8126 \pm 0.6255	5.1842 \pm 1.8195	8.5259 \pm 2.4833
NORMAL (300 images)	UISM↑	1.5512 \pm 0.2604	3.4903 \pm 0.7928	4.7259 \pm 0.3344	5.7581 \pm 0.6043	6.9285 \pm 0.4592
	UIConM↑	0.0768 \pm 0.0152	0.2389 \pm 0.0342	0.1700 \pm 0.0336	0.3092 \pm 0.0267	0.3733 \pm 0.0047
	IE↑	5.5313 \pm 0.2497	6.4120 \pm 0.1878	5.8949 \pm 0.2886	6.3338 \pm 0.7880	6.3062 \pm 0.1898
	C _{RAMM} ↑	1.2468 \pm 0.1495	3.9861 \pm 0.8219	2.5810 \pm 0.4295	6.3165 \pm 0.8719	8.3668 \pm 0.4882
	FADE ↓	0.6671 \pm 0.1226	0.5629 \pm 0.1132	0.3143 \pm 0.0622	0.2096 \pm 0.0876	0.1497 \pm 0.0186
	$\sigma\downarrow$	0.4923 \pm 0.0778	1.1698 \pm 0.3010	1.7765 \pm 0.2532	5.4016 \pm 0.9604	8.8852 \pm 1.5514
Average (2116 images)	UISM↑	2.1025 \pm 0.6595	3.9318 \pm 0.8274	4.8175 \pm 0.4819	5.9783 \pm 0.8516	6.9183 \pm 0.5981
	UIConM↑	0.0738 \pm 0.0219	0.2334 \pm 0.0404	0.1618 \pm 0.0478	0.2801 \pm 0.0390	0.3709 \pm 0.0137
	IE↑	5.3291 \pm 0.3735	6.3154 \pm 0.2130	5.6736 \pm 0.3731	5.9391 \pm 1.0404	6.1220 \pm 0.3221
	C _{RAMM} ↑	1.2774 \pm 0.3847	4.0071 \pm 1.1709	2.4333 \pm 0.7364	5.5587 \pm 1.1420	7.8364 \pm 0.6612
	FADE ↓	0.4966 \pm 0.1128	0.4445 \pm 0.1094	0.2525 \pm 0.0599	0.2097 \pm 0.1062	0.1304 \pm 0.0188
	$\sigma\downarrow$	0.6052 \pm 0.1961	1.4174 \pm 0.5090	1.9053 \pm 0.6446	5.1590 \pm 1.7480	8.9786 \pm 2.5470

contributions.

First, we adapted our recently developed image formation model, the double-pass fundus reflection model, to formulate the optical process of retinal image formation. This model is physically self-consistent and can better model the retinal image formation than traditional models. We further revealed the mathematical insight that the haze effect in RGB color space only appears in its Y channel in the YPbPr color space. This allows us to separate the haze component from RGB-colored images and thus better handle the dehazing problem for retinal images since they have different statistical features from natural scenes/objects.

Secondly, we mathematically unify the image dehazing task into the framework of image denoising which was historically thought to be unrelated (Supplementary Note 3). We discover that an accurate

estimation of the haze transmission matrix plays a crucial role in successful image dehazing, and this estimation can be achieved through proper image denoising techniques. This finding has the potential to advance the combination of research in image denoising and dehazing, fostering new insights and techniques in both traditional and deep learning-based approaches. We employ an intensity-based cost function that exploits the observation that the dehazing process consistently decreases the image's intensity. This idea facilitates the development of adaptive and unsupervised tuning of dehazing parameters.

Thirdly, MUTE outperforms SOTA traditional and learning-based methods, particularly in increasing the contrast of cataractous retinal images as presented in Section 5. The simplicity of the MUTE model structure allows for full monitoring and understanding of every

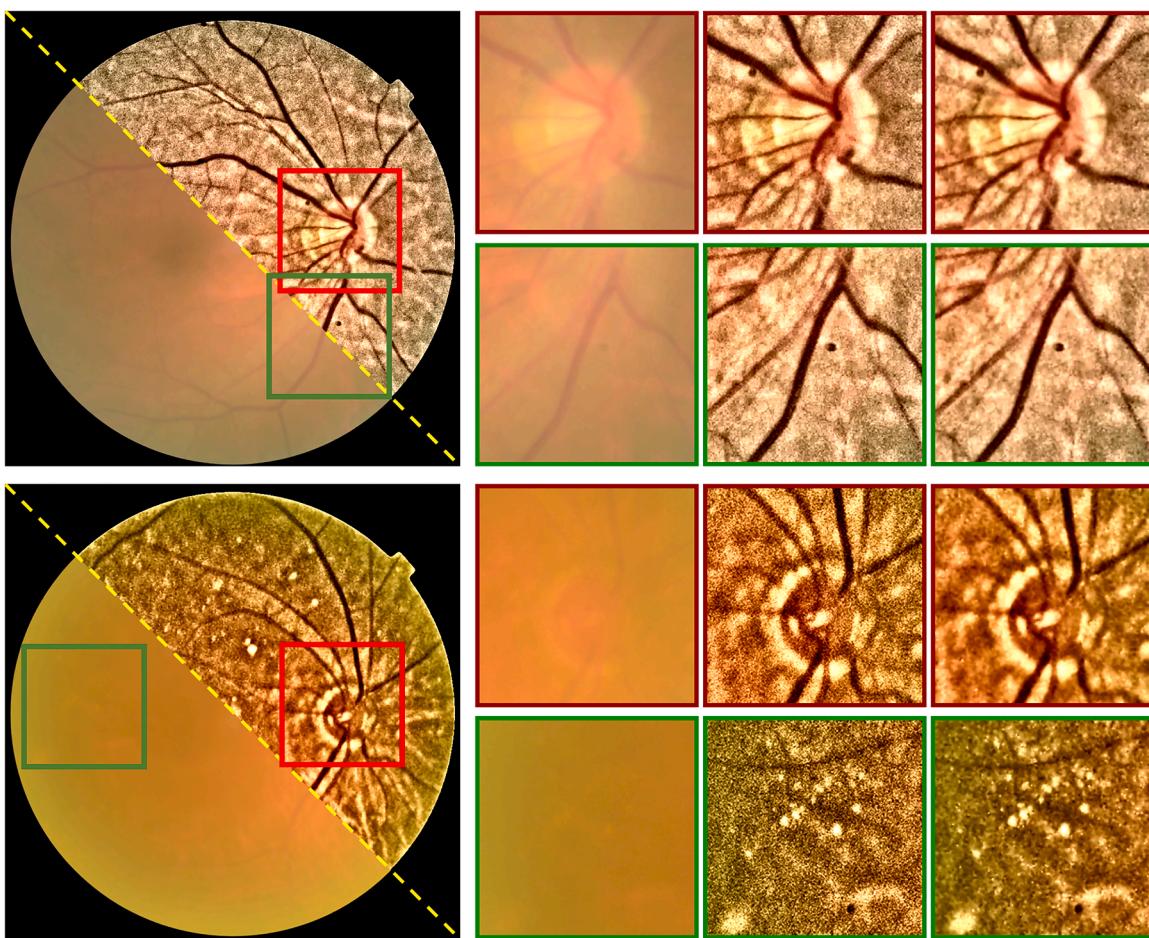


Fig. 26. Demonstration of MUTE with and without Hessian denoising process. $\lambda = \sigma/30$. Images in the first column are montages of raw and dehazed images. The third column and fourth column are original output and denoised output of MUTE for the zoomed-in area of red and green boxes.

Table 9
Average executing time for DVA, DRC, LPAR and MUTE.

Image Size	DVA (s)	DRC (s)	LPAR (s)	MUTE ($\mathcal{L} = 5$) (s)	MUTE with DUSS ($\mathcal{L} = 5$) (s)
512^2	7.21	1.71	2.73	6.23	6.23
$10,24^2$	16.22	3.11	8.22	23.05	16.33
$20,48^2$	48.13	15.76	38.96	98.44	16.54

intermediate step by human specialists such as mathematicians and ophthalmologists. This interpretability aspect ensures reliable and safe clinical trials and future applications of MUTE.

9. Conclusion

MUTE significantly improves the visual quality and the contrast of cataractous retinal images and reveals retinal structures that might not be detected by clinical doctors due to insufficient contrast of the raw images. It enables and facilitates the detection of features like drusen, hard exudates, bleeding hemorrhages, or microaneurysms, as these tasks may use more color or local detail information. Early detection of these findings through cataractous retinal images may improve the outcome of treatment of retinal diseases, and influent the decision-making on surgery that might be actually not necessary. For example, patients with severe cataracts might also have macular degeneration. In these cases, if the macular degeneration can be observed through enhanced cataractous retinal image, then ophthalmologists are able to choose a better treatment plan to minimize risks and stress for patients, for example, the

patients can be counseled on the guarded visual prognosis and a less expensive lens can be used.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Code will be available on request. Most data used in this research can be found in public datasets, while the authors have no permission to share the private cataractous retinal image data.

Acknowledgment

This research is supported by China Scholarship Council (CSC) (201908340078). Shuhe Zhang acknowledges his wife, Yanlin Zhuang's help during the preparation of this research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2023.102848.

References

- Ao, M., Li, X., Qiu, W., Hou, Z., Su, J., Wang, W., 2019. The impact of age-related cataracts on colour perception, postoperative recovery and related spectra derived from test of hue perception. *BMC Ophthalmol.* 19, 1–9.
- Arias, A., Artal, P., 2020. Wavefront-shaping-based correction of optically simulated cataracts. *Optica* 7, 22–27.
- Artal, P., Iglesias, I., López-Gil, N., 1995a. Double-pass measurements of the retinal-image quality with unequal entrance and exit pupil sizes and the reversibility of the eye's optical system. *J. Opt. Soc. Am. A* 12, 2358–2366.
- Artal, P., Marcos, S., Navarro, R., Williams, D.R., 1995b. Odd aberrations and double-pass measurements of retinal image quality. *J. Opt. Soc. Am. A* 12, 195–201.
- Berendschot, T.T.J.M., DeLint, P.J., Norren, D.v., 2003. Fundus reflectance—Historical and present ideas. *Prog. Retin. Eye Res.* 22, 171–200.
- cataract dataset. <https://www.kaggle.com/datasets/jr2ngb/cataractdataset>.
- Cao, L., Li, H., 2020. Enhancement of blurry retinal image based on non-uniform contrast stretching and intensity transfer. *Med. Biol. Eng. Comput.* 59.
- Cao, L., Li, H., 2021. Detail-richest-channel based enhancement for retinal image and beyond. *Biomed. Signal Process. Control* 69, 102933.
- Cao, L., Li, H., Zhang, Y., 2020. Retinal image enhancement using low-pass filtering and α -rooting. *Signal Processing* 170, 107445.
- Chakraborti, T., Jha, D.K., Chowdhury, A.S., Jiang, X.J.M.V., 2015. A self-adaptive matched filter for retinal blood vessel detection. *Applications* 26, 55–68.
- Cheung, C.Y., Ong, S., Ikram, M.K., Ong, Y.T., Chen, C.P., Venketasubramanian, N., Wong, T.Y., 2014. Retinal vascular fractal dimension is associated with cognitive dysfunction. *J. Stroke Cerebrovasc. Dis.* 23, 43–50.
- Choi, L.K., You, J., Bovik, A.C., 2015a. LIVE Image Defogging Database.
- Choi, L.K., You, J., Bovik, A.C., 2015b. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Trans. Image Process.* 24, 3888–3901.
- Ciancaglini, M., Guerra, G., Agnifili, L., Mastropasqua, R., Fasanella, V., Cinelli, M., Costagliola, C., Ambrosone, L., 2015. Fractal dimension as a new tool to analyze optic nerve head vasculature in primary open angle glaucoma. *In Vivo (Brooklyn)* 29, 273–279.
- DIARETDB0 - Standard Diabetic Retinopathy Database.
- DIARETDB1 - Standard Diabetic Retinopathy Database Calibration level 1. <http://www2.itlut.fi/project/imageret/diaretbdb1/>.
- Dai, P., Sheng, H., Zhang, J., Li, L., Wu, J., Fan, M., 2016. Retinal fundus image enhancement using the normalized convolution and noise removing. *Int. J. Biomed. Imaging* 2016, 1–12.
- Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., Erginay, A., Charton, B., Klein, J.-C., 2014. Feedback on a publicly distributed database: the Messidor database. *2014 33, 4%J Image Analysis & Stereology*.
- Dougal, F.N., MacGillivray, T.J., Patton, N., Dhillon, B., Dennis, M.S., Wardlaw, J.M., 2010. Fractal analysis of retinal vessels suggests that a distinct vasculopathy causes lacunar stroke. *Neurology* 74, 1102–1107.
- Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A., 1998. Multiscale Vessel Enhancement filtering. *International conference On Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 130–137.
- Galdran, A., Alvarez-Gila, A., Bria, A., Vazquez-Corral, J., Bertalmio, M., 2018. On the duality between retinex and image dehazing. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8212–8221.
- Gaudio, A., Smailagic, A., Campilho, A.e., 2020. Enhancement of retinal fundus images via pixel color amplification. In: *Proceedings of the International Conference on Image Analysis and Recognition*.
- Gupta, B., Tiwari, M., 2019. Color retinal image enhancement using luminosity and quantile based contrast enhancement. *Multidimens. Syst. Signal. Process.* 30, 1829–1837.
- Hong, S., Kim, M., Kang, M., 2021. Single image dehazing via atmospheric scattering model-based image fusion. *Signal Processing* 178, 107798.
- Hoover, A., Kouznetsova, V., Goldbaum, M., 2000. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. *IEEE Trans. Med. Imaging* 19, 203–210.
- Huang, F., Dashtbozorg, B., Zhang, J., Bekkers, E., Abbasi-Sureshjani, S., Berendschot, T. T., Ter Haar Romeny, B.M., 2016. Reliability of using retinal vascular fractal dimension as a biomarker in the diabetic retinopathy detection. *J. Ophthalmol.* 2016, 6259047.
- Immerkær, J., 1996. Fast Noise Variance Estimation. *Comput. Vis. Image Understand* 64, 300–302.
- Ju, M., Ding, C., Guo, Y.J., Zhang, D., 2021. IDGCP: image dehazing based on gamma correction prior. *IEEE Trans. Image Process.* 29, 3104–3118.
- Kingma, D.P., Ba, J.L., 2015. Adam: a method for stochastic optimization. In: *Proceedings of the International Conference on Learning Representations*.
- Koschmieder, H., 1938. Luftlicht und sichtweite. *Naturwissenschaften* 26, 521–528.
- Li, H., Liu, H., Fu, H., Shu, H., Zhao, Y., Luo, X., Hu, Y., Liu, J., 2022a. Structure-consistent restoration network for cataract fundus image enhancement. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Springer Nature Switzerland, Cham, pp. 487–496.
- Li, H., Liu, H., Hu, Y., Fu, H., Zhao, Y., Miao, H., Liu, J., 2022b. An annotation-free restoration network for cataractous fundus images. *IEEE Trans. Med. Imaging* 41, 1699–1710.
- 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 J. Biomed. Health Inform.*
- Mary, M.C.V.S., Rajsingh, E.B., Naik, G.R., 2016. Retinal fundus image analysis for diagnosis of glaucoma: a comprehensive survey. *IEEE Access* 4, 4327–4354.
- He, K., Sun, J., Tang, X., 2009. Single image haze removal using dark channel prior, 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 1956–1963.
- McCartney, E.J., 1976. Optics of the Atmosphere: scattering by molecules and particles. New york, 408.
- Mitra, A., Roy, S., Roy, S., Setua, S.K., 2018. Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract. *Comput. Methods Programs Biomed.* 156, 169–178.
- Mujbaile D., and Rojatkar D., 2020. Model based Dehazing Algorithms for Hazy Image Restoration – A Review, 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 2020, pp. 142–148.
- Panetta, K., Gao, C., Agian, S., 2016. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* 41, 1–11.
- Peli, E., Peli, T., 1989. Restoration of retinal images obtained through cataracts. *IEEE Trans Med Imaging* 8, 401–406.
- Rizzi, A., Algeri, T., Medeghini, G., Marini, D., 2004. A proposal for contrast measure in digital images. In: *proceedings of the Conference on colour in graphics, imaging, and vision*, pp. 187–192.
- Rudin, L.I., Osher, S., Fatemi, E.J.P.D.N.P., 1992. Nonlinear total variation based noise removal algorithms. *Physica D* 60, 259–268.
- Savelli, B., Bria, A., Galdran, A., Marrocco, C., Molinara, M., Campilho, A.e., Tortorella, F., 2017. Illumination correction by dehazing for retinal vessel segmentation. In: *Proceedings of the IEEE 30th International Symposium on Computer-Based Medical Systems*.
- Saxena, G., Bhaduria, S.S., 2020. An efficient single image haze removal algorithm for computer vision applications. *Multimed. Tools Appl.* 79, 28239–28263.
- Senthilkumar, K.P., Sivakumar, P., 2019. A review on haze removal techniques. *Comput. Aided Intervent. Diagnostics Clinic. Med. Images* 113–123.
- Siegrist, M., Hanssen, H., Neidig, M., Fuchs, M., Lechner, F., Stetten, M., Blume, K., Lammel, C., Haller, B., Vogeser, M., Parhofer, K.G., Halle, M., 2014. Association of leptin and insulin with childhood obesity and retinal vessel diameters. *Int. J. Obes.* 38, 1241–1247.
- Singh, D., Kumar, V., 2018. A comprehensive review of computational dehazing techniques. *Arch. Comput. Methods. Eng.*
- Staal, J., Abramoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B., 2004. Ridge-based vessel segmentation in color images of the retina. *EEE Trans. Med. Imagin.* 23 (4), 501–509.
- STructured Analysis of the Retina. <https://cecas.clemson.edu/~ahoover/stare/>.
- Wang, J., Chen, J., Xu, H., Zhang, S., Mei, X., Huang, J., Ma, J.J.S.P., 2019. Gaussian field estimator with manifold regularization for retinal image registration. *Signal Processing* 157, 225–235.
- Wang, J., Li, Y.-J., Yang, K.-F., 2021. Retinal fundus image enhancement with image decomposition and visual adaptation. *Comput. Biol. Med.* 128, 104116.
- Wang, J., Lu, K., Xue, J., He, N., Shao, L., 2018. Single image dehazing based on the physical model and MSRCR algorithm. *IEEE Trans. Circuits Syst. Video Technol.* 29, 2190–2199.
- Xiong, L., Li, H., Xu, L., 2017. An enhancement method for color retinal images based on image formation model. *Comput. Methods Programs Biomed.* 143, 137–150.
- Xu, L., Lu, C., Xu, Y., Jia, J., 2011. Image smoothing via L 0 gradient minimization. In: *Proceedings of the 2011 SIGGRAPH Asia conference*, pp. 1–12.
- Yang, B., Zhao, H., Cao, L., Liu, H., Wang, N., Li, H., 2023. Retinal image enhancement with artifact reduction and structure retention. *Pattern Recognit.* 133, 108968.
- Zafar, S., McCormick, J., Giancardo, L., Saidha, S., Abraham, A., Channa, R., 2019. Retinal Imaging for Neurological Diseases: A Window into the Brain. *Int Ophthalmol Clin* 59, 137–154.
- Zhang, J., Dashtbozorg, B., Huang, F., Berendschot, T.T.J.M., ter Haar Romeny, B.M., 2018. Analysis of retinal vascular biomarkers for early detection of diabetes. In: Tavares, J.M.R.S., Natal Jorge, R.M. (Eds.), *VipIMAGE 2017*. Springer International Publishing, Cham, pp. 811–817.
- Zhang, S., Webers, C.A.B., Berendschot, T.T.J.M., 2022. A double-pass fundus reflection model for efficient single retinal image enhancement. *Signal Processing* 192, 108400.
- Zhou, M., Jin, K., Wang, S., Ye, J., Qian, D., 2018. Color retinal image enhancement based on luminosity and contrast adjustment. *IEEE Trans. Biomed. Eng.*
- Zhu, P., Huang, F., Lin, F., Li, Q., Yuan, Y., Gao, Z., Chen, F., 2014. The relationship of retinal vessel diameters and fractal dimensions with blood pressure and cardiovascular risk factors. *PLoS ONE* 9, e106551.
- Zhu, Q., Mai, J., Shao, L., 2015. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* 24, 3522–3533.