

Article

A Novel Un-Supervised GAN for Fundus Image Enhancement with Classification Prior Loss

Shizhao Chen, Qian Zhou and Hua Zou * 

School of Computer Science, Wuhan University, Wuhan 430072, China; chensz@whu.edu.cn (S.C.); zhouqian@whu.edu.cn (Q.Z.)

* Correspondence: zouhua@whu.edu.cn

Abstract: Fundus images captured for clinical diagnosis usually suffer from degradation factors due to variation in equipment, operators, or environment. These degraded fundus images need to be enhanced to achieve better diagnosis and improve the results of downstream tasks. As there is no paired low- and high-quality fundus image, existing methods mainly focus on supervised or semi-supervised learning methods for color fundus image enhancement (CFIE) tasks by utilizing synthetic image pairs. Consequently, domain gaps between real images and synthetic images arise. With respect to existing unsupervised methods, the most important low scale pathological features and structural information in degraded fundus images are prone to be erased after enhancement. To solve these problems, an unsupervised GAN is proposed for CFIE tasks utilizing adversarial training to enhance low quality fundus images. Synthetic image pairs are no longer required during the training. A specially designed U-Net with skip connection in our enhancement network can effectively remove degradation factors while preserving pathological features and structural information. Global and local discriminators adopted in the GAN lead to better illumination uniformity in the enhanced fundus image. To better improve the visual quality of enhanced fundus images, a novel non-reference loss function based on a pretrained fundus image quality classification network was designed to guide the enhancement network to produce high quality images. Experiments demonstrated that our method could effectively remove degradation factors in low-quality fundus images and produce a competitive result compared with previous methods in both quantitative and qualitative metrics.

Keywords: fundus image enhancement; unsupervised learning; classification prior loss; U-Net; GAN



Citation: Chen, S.; Zhou, Q.; Zou, H. A Novel Un-Supervised GAN for Fundus Image Enhancement with Classification Prior Loss. *Electronics* **2022**, *11*, 1000. <https://doi.org/10.3390/electronics11071000>

Academic Editor: Kenji Suzuki

Received: 19 January 2022

Accepted: 16 March 2022

Published: 24 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Color fundus images are captured by a specially designed camera system which records the scene observed by human beings through an ophthalmoscope [1]. Ophthalmologists can identify specific pathological features and diagnose diseases by observing morphological changes of the entire retina in the image. Due to the low cost of imaging, ease of use, and high safety, color fundus images have been widely used in clinical screening and diagnosis of ophthalmic diseases and computer-aided diagnosis systems. However, the fundus image capturing process may introduce some annoying degradation factors to the image, causing uneven illumination, artifacts, blurs, etc. These mixed degradations are equivalent to dark clouds obscuring the retina, which severely reduce the visibility of anatomical retinal structures (e.g., vessel, optic disc, and cup), especially for small pathological features. Thus, the degraded fundus image needs to be enhanced to meet practical clinical needs and downstream task requirements. For example, CFIE techniques can be integrated into clinical fundus image capture systems and can improve the visual quality of degraded images for better diagnosis. Downstream tasks, such as diabetic retinopathy (DR) classification and retinal vessel segmentation, can also incorporate the CFIE technique as a pre-process task in the whole task pipeline for enhanced performance.

For low-level tasks, image enhancement algorithms (e.g., denoise [2], dehaze [3–5], derain [6], super-resolution [7–12], low-light enhancement [13,14], etc.) have evolved

over generations and yielded many promising results in recent years. Most supervised image enhancement methods can reach a satisfying score by simply utilizing an L_1 , L_2 , or other referenced loss, with paired low- and high- quality image data as supervision. These successes are usually attributed to the powerful capability of the elaborated neural network. With respect to domains without paired data, supervised methods are prone to use synthesized image pairs. However, synthesized methods can only simulate a few limited kinds of degradation, which may fail when unseen or more complicated degradations appear. Regarding the fundus image domain, it is impractical to capture paired low- and high-quality fundus image data. Degradations in low-quality fundus images are too varied in quantity and degree to synthesize, so unsupervised image enhancement approaches represent another option.

With the great success of GAN in image enhancement [15], more and more papers are adopting GAN networks as their training scheme when there is a lack of paired training data. GAN networks can generate realistic images while not requiring paired images but using paired domain images instead, which are quite suitable for fundus image enhancement tasks. However, there are still several critical questions which need to be solved. Firstly, the GAN generator is trained by minimizing the distribution of generated images and real target images while maximizing the ability of the discriminator to distinguish the two kinds of images. GAN loss is not accurate as pixel level loss occurs, so it may generate undesired features in the enhanced image. This is not acceptable for a fundus image as these features may interfere with the recognition of pathology and anatomical retinal structures. Secondly, the generator does not know which enhancement is optimal. To achieve this goal, extra constraint should be applied on the generator network.

To address the issues above, a U-Net-based [16] GAN network is proposed to enhance low quality fundus images. Our method treats the degradation removal goal as a reverse manipulation of adding degradation features to the clean images. That is, inversed degradation features in low quality images are extracted from the U-Net and are then applied to the original low quality images to obtain enhanced images. The pathological and anatomical features remain unchanged. In addition to the GAN loss, a classification prior loss (CPL) is proposed based on a pretrained fundus image quality classification model, which can evaluate the quality of enhanced images and provide a gradient to the generator to produce high quality images.

2. Related Work

Image enhancement techniques are closely related to their target domains. To date, no enhancement method that can be generalized to every image domain exists. In this section, image enhancement methods, from natural image domains to fundus image domains, will be introduced to highlight the common underlying ideas and domain-specific differences that have stimulated our work.

2.1. Natural Image Enhancement

With the development of deep learning models, many tasks have used learning-based methods for natural image enhancement, such as SR-GAN [9] for image super-resolution, [4,17] for dehazing, [18] for denoising, [13] and for low-light enhancement. In this section, a low-light enhancement task is introduced as an example. Lighting conditions have a great influence on image quality. Images captured under low light conditions have poor visibility. Wei et al. proposed Retinex-Net based on Retinex theory [19] which separately decomposes low and high images into a reflectance map and an illumination map, and then applies a traditional denoising operation BM3D [20] to the reflectance map to remove the noise. In the meantime, the illumination map is enhanced by a deep neural network. Finally, the well-lit image is reconstructed with the two adjusted maps by a simple element-wise product. Considering the noise variance related to the reflectance map, Zhang et al. [14] introduced Restoration-Net in substitution for BM3D in Retinex-Net to handle the reflectance map in Retinex-Net. Instead of using Retinex theory,

Wang et al. [21] designed GLAD-Net which adopted an encoder-decoder network to generate global a priori knowledge of illumination from low light images to guide the enhancement process. However, the low-light enhancement methods mentioned above still depend on the supervision of specially captured or synthetic low and normal light image pairs. What about no paired training data? Zhao et al. [22] proposed RetinexDIP which adopted a novel generative strategy for the Retinex decomposition process, that is, the reflection map and illumination map were generated by two vanilla GAN networks, and the noise hidden in the reflectance map was removed by the generator. However, this kind of generative strategy requires significant optimization time. Guo et al. [23] treat the low light enhancement task as a curve adjustment task which has been widely used in business image processing software (e.g., Photoshop). The curve parameter was learned from the low light images with a deep neural network, combined with a set of carefully formulated non-reference loss functions which implicitly measure the quality of the enhanced image, such that a low light image is successfully enhanced without any paired or unpaired dataset. Jiang et al. [15] proposed EnlightenGAN to achieve this by incorporating a generative adversarial network and regularizing the unpaired training using information extracted from the input itself. With the help of an illumination attention map and perceptual loss, Jiang's methods not only outperformed their recent methods, but can also be easily adapted to enhance real world images from various domains. For image domains such as fundus images, for which it is impossible to obtain a paired image dataset, adversarial training is suitable for enhancing low quality fundus images to match the real distribution of high-quality fundus images. In this paper, the advantages of GAN training and non-reference loss are combined for better results.

2.2. Fundus Image Enhancement

Color fundus image quality is often degraded by uneven illumination, artifacts, low contrast, etc. For improved analysis, degraded color fundus images should be enhanced. Chen et al. [24] proposed a structure-preserving guided retinal image filtering method based on an attenuation and scattering model, which can improve the contrast of color fundus images. Since it is quite difficult to obtain paired medical images, You et al. [25] proposed a retinal enhancement method called Cycle-CBAM which requires no paired training data. Cycle-CBAM utilizes Cycle-GAN [26] as its main framework in which CBAM [27] is embedded, and yields better results than the original Cycle-GAN. However, this method suffers from the critical defect that it may introduce some fake features to the enhanced image. Ma et al. [28] proposed StillGAN, which also adopted Cycle-GAN, to improve medical image quality without paired data. He argued that CycleGAN-based methods only focus on global appearance without imposing constraints on structure or illumination which were essential features for medical image interpretation, so his method proposed a luminance loss function and a structure loss function as extra constraints. The luminance loss measures illumination smoothness by calculating the variance of the illumination map in enhanced image patches, and the structure loss measures the dissimilarity between the low-quality image and its enhanced version by calculating the correlation coefficient in image patches. After analyzing the ophthalmoscope image system, Shen et al. [1] proposed a network named cofe-Net which synthesized low and high color fundus image pairs by modeling the degradation process of uneven illumination, blur, and artifacts. With supervision of synthesized low- and high-quality fundus image pairs and a synthesized degradation map, Shen's cofe-Net can suppress global degradation factors while preserving anatomical retinal structures and pathological characteristics. However, cofe-Net can only ideally model a limited type of degradation factors—it may fail when there are more complicated degradations. Cheng et al. [29] proposed EPC-GAN by training both GAN loss and contrastive loss to make use of high-level features in the fundus domain. A fundus prior loss based on a pretrained diabetic retinopathy classification network was introduced to avoid information modification and over-enhancement. Wang et al. [30] proposed a fundus image enhancement method which firstly decomposes low quality

image into three different layers, then the denoising operation, illumination and detail enhancement are applied to these three layers, respectively. This method yielded enhanced images with strong and sharp feature edges, but this was accompanied with obvious color distortion. Zhang et al. [31] proposed a double-pass fundus reflection (DPFR) model based on intraocular scattering, aimed at improving the clarity of the fundus image. It was found that DPFR improved the visibility of retinal vessels significantly, but color distortion continued to occur in the enhanced images. Raj et al. [32] proposed a residual dense connection based U-Net (RDC-UNet) to enhance five typical degradations in low quality fundus images with synthetic image pairs individually; the five trained model was then ensemble together to enhance low quality images—the enhanced results showed a promising naturalness. However, for different types of degradation, new synthetic algorithms should be designed to fit the model, which will lead to enormous complexity in calculation. Stimulated by this prior loss, this paper considers the design of a non-reference loss method called CPL, based on a pretrained fundus image quality classification network to measure enhancement quality.

3. Method

As illustrated in Figure 1, our network was adopted from [15], which is designed and trained in the GAN manner. The generator component adopts a modified U-Net as the main frame, which takes the concatenation of a low-quality fundus image and its illumination map as input and produces a high-quality fundus image. There is a skip connection between the input and output of the U-Net to preserve most of the important features in the original fundus image; the symmetrically expanding path in the U-Net not only preserves precise details [5], but also incorporates an illumination attention mechanism [15]. A global discriminator is used to make the high-quality image look more realistic, and a local discriminator is used to enhance local areas. In addition, a pretrained image quality classification network is introduced as CPL to drive the generator to remove undesired degradation factors.

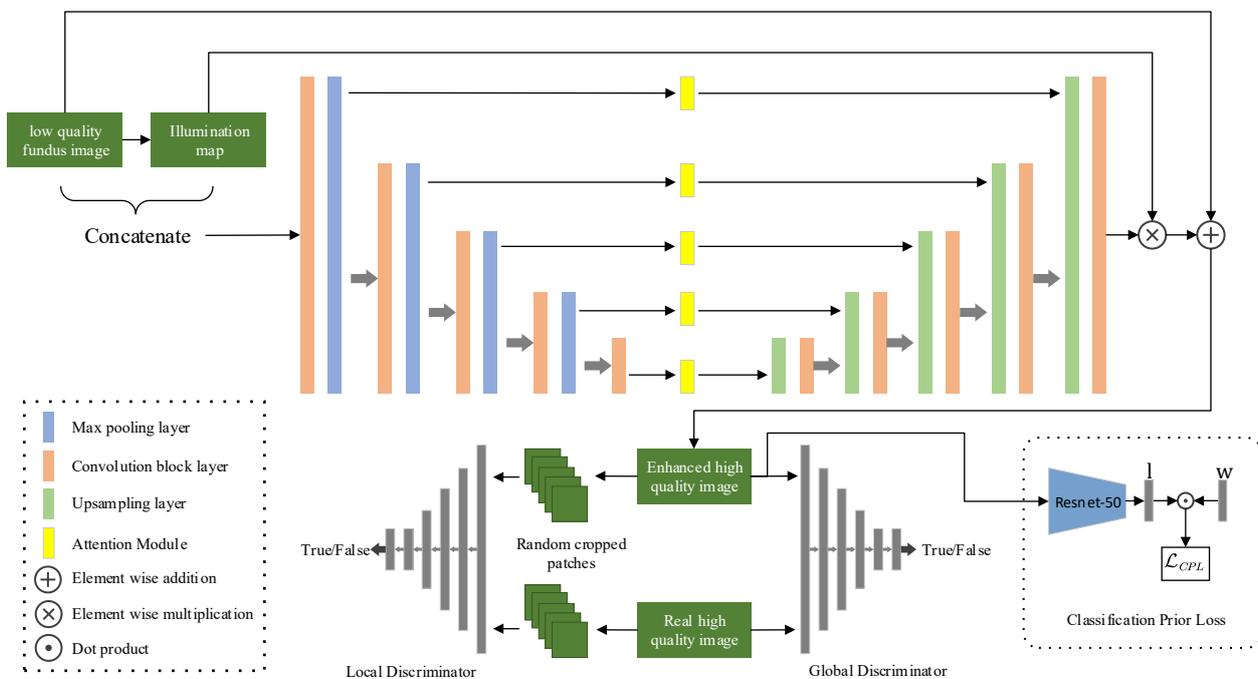


Figure 1. The overall architecture of our network.

3.1. U-Net Generator with Skip Connection

Degradations in color fundus images, such as uneven illumination, artifacts, blurriness, etc., are usually caused by human factors, equipment factors or environmental factors [1].

In a similar way to Retinex theory applied to low light enhancement tasks, a degraded color fundus image is treated as a composite of a clean image and degradation factors, so that a degraded color fundus image can be simply formulated as:

$$x_L = y + d_{x_L} \quad (1)$$

where x_L is the degraded low-quality image, y is the corresponding undegraded clean image, and d_{x_L} represents the mixed degradation factors related to x_L . When a degraded low-quality image x_L was obtained, the restoration process for a high-quality image y' becomes:

$$y' = x_L - d_{x_L} \quad (2)$$

Then the key problem is to find d_{x_L} hidden in x_L . Inspired by the great success of U-Net in medical image segmentation, this paper treats this problem as a segmentation-like task. Instead of directly putting the low-quality image into a network to eliminate the degradation factors and obtain the enhanced image, our network explicitly extracts the degradation factors by a separate network branch design. Given a low image x_L , the opposite degradation factors $-d_{x_L}$ can be extracted by $-d_{x_L} = U(x_L)$, where U represents a modified version of U-Net in Figure 1. Then the full restoration network G can be formulated as:

$$y' = G(x_L) = x_L + U(x_L) \quad (3)$$

This can be achieved by simply adding a skip connection between the input x_L and the output of $U(x_L)$. This kind of operation has the following advantages:

- Most of the important features, e.g., lesions, vessels, macula and optic disc can be well preserved after enhancement and have no obvious spatial shift;
- The generator network is easy to train with skip connections and saves a lot of training time;
- With the condition of inputting a low quality image, the generator will not produce unexpected features as other vanilla GANs do without this condition.

3.2. Adversarial Training for Unpaired Image Enhancement

Since there is no ground truth high-quality image for enhanced fundus images to calculate referenced loss, and supervised learning methods are not applicable to this kind of FIQE task, adversarial training is utilized to train a generator for the purpose of enhancement. Given a low-quality image domain X and a high-quality image domain Y , our goal is to convert a degraded fundus image $x \in X$ into a clear fundus image $y' \in Y'$ by a generator G , both Y' and Y should follow the same distribution. The global-local discriminator loss functions from [15] are adopted as GAN loss functions. The loss functions for the global discriminator D and the generator G are:

$$\mathcal{L}_D^{Global} = \mathbb{E}_{y \sim Y} [(D_{Ra}(y, y') - 1)^2] + \mathbb{E}_{y' \sim Y'} [D_{Ra}(y', y)^2] \quad (4)$$

$$\mathcal{L}_G^{Global} = \mathbb{E}_{y' \sim Y'} [(D_{Ra}(y', y) - 1)^2] + \mathbb{E}_{y \sim Y} [D_{Ra}(y, y')^2] \quad (5)$$

where D_{Ra} is a modified version of relativistic GAN [33] loss function which estimates the probability that real data is more realistic than fake data and directs the generator to synthesize images more realistic than real images. D_{Ra} can be formulated as:

$$D_{Ra}(y, y') = C(y) - \mathbb{E}_{y' \sim Y'} [C(y')] \quad (6)$$

$$D_{Ra}(y', y) = C(y') - \mathbb{E}_{y \sim Y} [C(y)] \quad (7)$$

where C denotes the network of the global discriminator. For the local discriminator, five patches are randomly cropped from y and y' for each image. The original LSGAN [34] is adopted as the adversarial loss for the local discriminator as follows:

$$\mathcal{L}_D^{Local} = \mathbb{E}_{y \sim Y_{patches}} [(D(y) - 1)^2] + \mathbb{E}_{y' \sim Y_{patches}} [(D(y') - 0)^2] \quad (8)$$

$$\mathcal{L}_G^{Local} = \mathbb{E}_{y' \sim Y'_{patches}} [(D(y') - 1)^2] \quad (9)$$

where D denotes the local discriminator. Given the fact that most low quality fundus images are degraded by uneven illumination [35], the self-feature-preserving loss in Enlighten GAN is also adopted in our model to preserve the fundus image content features, which can be adopted to preserve features that are irrelevant to intensity changes. The self-feature-preserving loss can be described as:

$$\mathcal{L}_{SFP}(x) = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(x) - \phi_{i,j}(G(x)))^2 \quad (10)$$

where x denotes the input low quality fundus image, G denotes the generator, $\phi_{i,j}$ denotes the feature map extracted from a VGG-16 [36] model pretrained on ImageNet, i, j represents its j -th convolutional layer after i -th max pooling layer, $W_{i,j}$ and $H_{i,j}$ are the dimensions of the extracted feature maps, and this paper chooses $i = 5, j = 1$.

3.3. Classification Prior Loss Guided Generator

Although the generator can produce realistic images which match the distribution of real high-quality images well under the constraint of GAN loss, it may still introduce some unpleasant artifacts due to the limitation of the discriminator and a lack of real labeled data. Inspired by Guo et al. [23], a carefully designed unreferenced loss can be adopted to constrain the image enhancement network, even though there is no target domain data. This is realized by designing loss functions that constrain certain target domain features. Yifan J et al. [15] proposed an unreferenced self-feature-preserving loss to preserve the image content features based on the observation that classification results are not very sensitive when pixel intensity range changes, which means that features extracted from both low light images and corresponding normal light images share the same feature space. Pujin C et al. [29] designed a fundus prior loss by pre-training a DR classification model to keep the fundus semantic information stable before and after enhancement, because deep features related to pathological areas should be preserved after enhancement.

An unreferenced classification prior loss (CPL) is proposed to constrain the image quality produced by the generator. A Resnet-50 [37] is pre-trained as a fundus image quality classification network P to generate a 3-class label vector l , each component of l represent the probability of each class (e.g., Good, Usable and Reject in the EyeQ [38] dataset), and the image quality loss for y' can be defined as:

$$\mathcal{L}_{CPL} = \sum_{x=1}^N w_i l_i, l = P(y') \quad (11)$$

where w_i is the penalty weight of each corresponding component of l . N represents the number of classes produced by P , which was set to $N = 3$ because l is a 3-class label. Our goal was to generate more images labeled with Good; when images labeled with Usable and Reject were generated, more penalty should be applied, so our method empirically set $w = \{0, 1, 2\}$.

3.4. Objective Function

With the help of CPL, the overall loss function for the generator is formulated as:

$$Loss = \mathcal{L}_{SFP}^{Global} + \mathcal{L}_{SFP}^{Local} + \mathcal{L}_G^{Global} + \mathcal{L}_G^{Local} + \lambda_{CPL} \mathcal{L}_{CPL} \quad (12)$$

where λ_{CPL} is the corresponding trade-off weight for CPL, which was set as $\lambda_{CPL} = 0.1$ in our project.

4. Experiments

4.1. Datasets

The EyeQ dataset [38] was used to train and validate our proposed method. EyeQ dataset is a re-annotation subset from the EyePACS [39] dataset for fundus image quality assessment, which has 28,792 retinal images with a three-level quality grading (i.e., ‘Good’, ‘Usable’ and ‘Reject’). The whole dataset was split into a training dataset consisting of 12,543 images and a testing dataset consisting of 12,649 images. The Resnet-50 for CPL was pretrained on the training dataset of EyeQ. Images graded as ‘Good’ and ‘Usable’ from the training dataset were used as real high-quality images and low-quality images respectively to train our model. Images graded as ‘Usable’ from the testing dataset were used for qualitative comparison. Since there were no paired images for calculating the full-reference metric, the open-sourced code from [1] was followed to degrade the ‘Good’ images in EyeQ dataset into synthesized low-quality images, which contained the permutation of three kinds of degradation factor (illumination, artifacts and blur), finally yielding seven kinds of mixed degradation. This paired image dataset was called the synthetic dataset.

4.2. Implementation

Our model was implemented with PyTorch (Version 1.7) and trained on a PC with two NVIDIA TITAN RTX GPUs. All the images from EyeQ were first preprocessed into a size of 512×512 with code released by [1].

The Resnet-50 for CPL was trained on the training dataset of EyeQ, which takes images of size 512×512 and produces image quality labels (e.g., 0: Good, 1: Usable, 2: Reject) under the constraint of cross entropy loss. An Adam optimizer with learning rate of 0.0002 was used for optimization. Finally, the weights chosen for CPL reached a highest classification accuracy of 0.8623.

Images for training enhancement GAN were resized to 256×256 to save training time. Images for testing were retained at a size of 512×512 for fine grained visual comparison. The Adam optimizer with a learning rate of 0.0002 was used as the optimizer for GAN training.

4.3. Quantitative and Qualitative Evaluation

In order to demonstrate the advantages of our proposed method, this paper conducted both quantitative and qualitative evaluations for comparison with other GAN-based methods, including cGAN [40], CycleGAN [26], CutGAN [41] and StillGAN [28]. The cGAN was trained with synthetic datasets, while CycleGAN, CutGAN and StillGAN were trained with the dataset on which our model was trained.

4.3.1. Quantitative Results

PSNR (peak signal-to-noise ratio) and SSIM (structural similarity index measure) were employed as quantitative metrics to measure enhancement quality. All the scores were calculated from the test part of the synthetic dataset. Table 1 shows the PSNR and SSIM score of different methods for the synthetic dataset. The PSNR score of our proposed method outperformed the other methods. Although our SSIM score was lower than that for cGAN which was trained on the synthetic dataset, our method still obtained the highest score compared with methods such as CycleGAN, CutGAN and StillGAN, which require no paired images. The ablation test data shows that our CPL loss functions improved

both PSNR and SSIM scores compared with our method without CPL constraint. The quantitative results demonstrated the effectiveness of our proposed method on removing synthetic degradations in color fundus images.

Table 1. Average PSNR (dB) and SSIM results on synthetic dataset, higher score means better performance. The best results are marked in bold.

Method	Synthetic Dataset	
	PSNR (dB)	SSIM
cGAN	23.2	0.8946
CycleGAN	22.84	0.843
CutGAN	21.89	0.8534
StillGAN	23.44	0.8693
Proposed w.o. CPL	23.82	0.8753
Proposed	23.93	0.8859

4.3.2. Qualitative Results

To demonstrate the improvements in enhanced images in downstream tasks, such as retinal vessel segmentation, the model and pre-trained weight from Iter-Net [42] were utilized to perform vessel segmentation operations on fundus images before and after enhancement.

As shown in Figure 2, our enhanced images showed more vessel branches than the original one. In the meantime, our method and StillGAN did not modify the structure of retinal vessels while other methods, such as CycleGAN and CutGAN may generate fake retinal vessel content. Case 1 shows that retinal vessels segmented from our enhanced image were clearer and morphologically continuous. Case 2 shows that our enhanced image did not introduce vessel branches which did not exist in the original low-quality images.

Figure 3 Shows the visual results of our methods and other methods. Pathological features in fundus images are critical to clinical diagnosis. Case 1 shows that pathological features like exudation and micro-angioma in our enhanced image were even more sharp and clear than the original image. In contrast, some of these features were blurred or erased by other methods. Case 2 shows that vessels over pathological areas were well preserved after our enhancement while methods such as CycleGAN and CutGAN failed to recover this area. Case 3 shows the results of optic disc regions which were brighter than other areas. Vessels were still clear in our results, but this area suffered from severe degradation after enhancement with CycleGAN and CutGAN.

In summary, it can be observed that real low-quality images enhanced by cGAN still suffered from uneven illumination degradations, as shown in both Figures 2 and 3. Though cGAN achieved the highest SSIM score on the synthetic test dataset, unsatisfying, or worse, visual results appeared when it was adapted to real degraded fundus images. CycleGAN and CutGAN appeared to have better overall visual quality, but low scale features and structures were modified to a certain extent, which means they may introduce interference to ophthalmologists' diagnosis. StillGAN achieved similar vessel-structure-preserving performance to our method, but the pathological features were not sharp enough. It can be concluded that our method not only improved the overall visual quality of degraded fundus images, but also preserved low-scale pathological features and structures.

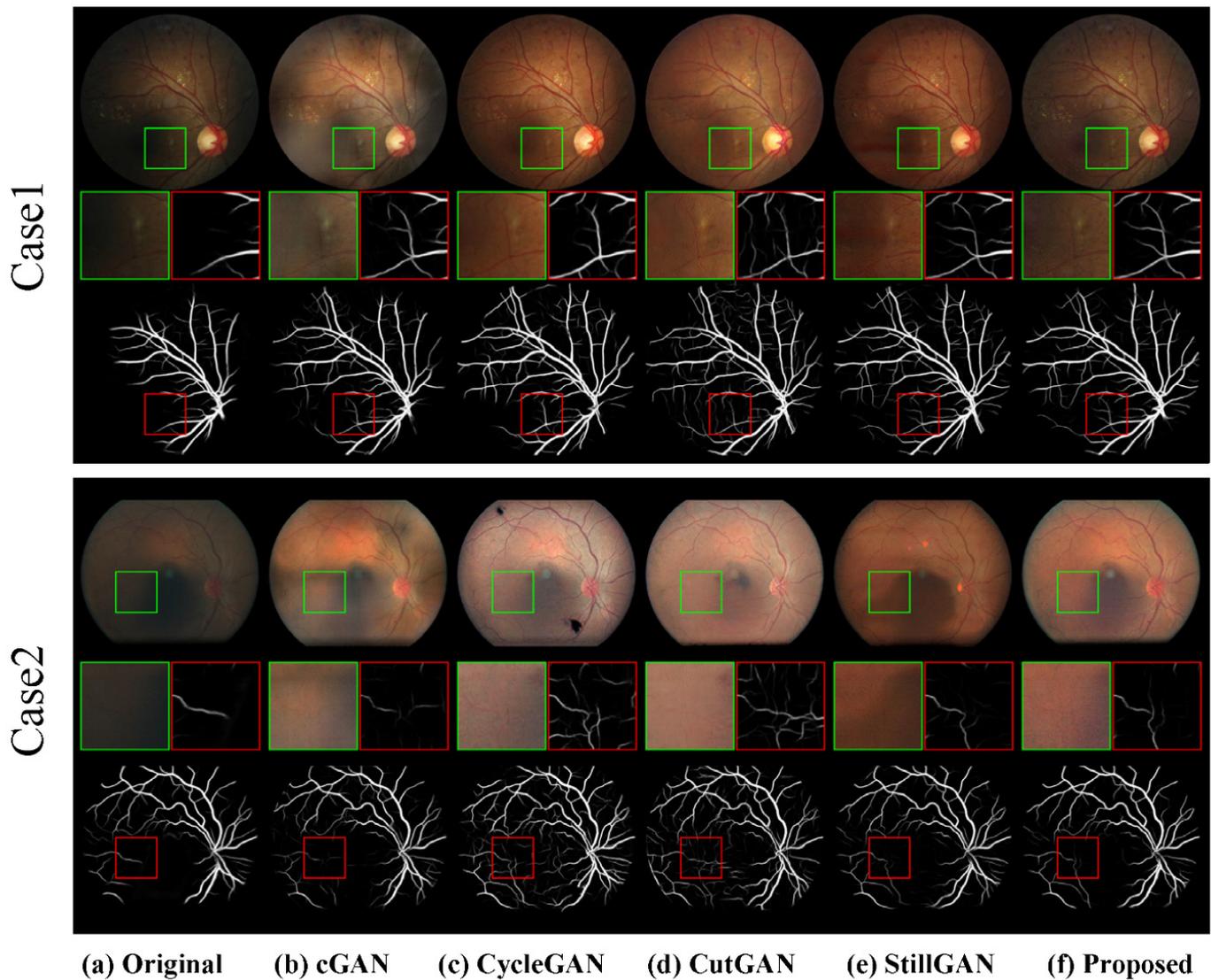


Figure 2. Visual comparisons on the vessel segmentation task. In each case, the first row contains color fundus images enhanced by different methods, and the last row contains corresponding segmentation results. In the middle row, the green box is cropped from color fundus while the red box is cropped from the corresponding area of the segmentation results.

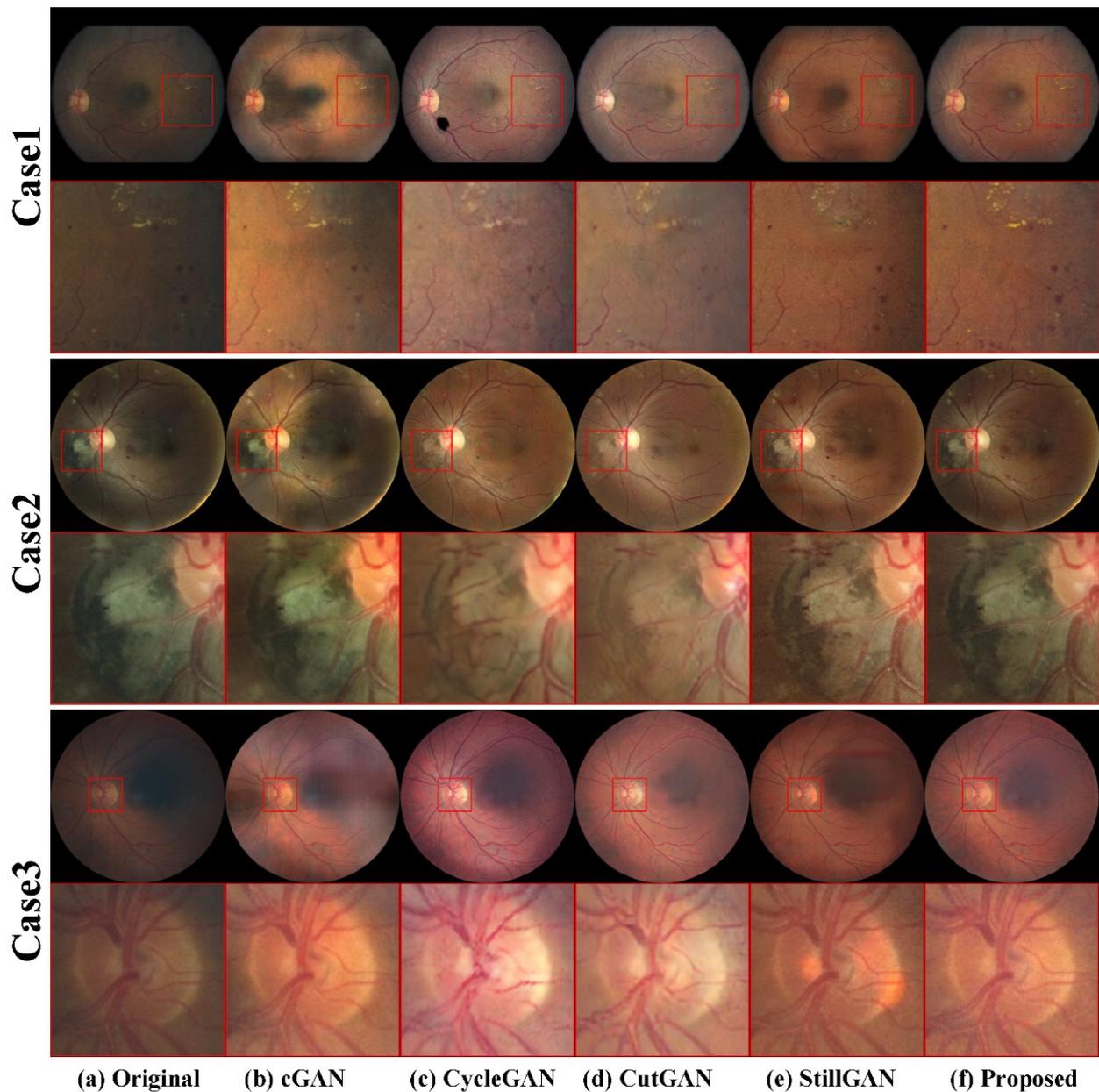


Figure 3. Visual comparisons on image enhancement task. In each case, the first row contains color fundus images enhanced by different methods, and the second row are patches cropped from red box areas in the first row.

5. Conclusions

An unsupervised color fundus image enhancement method based on classification prior loss was proposed in this paper. Synthetic paired-image datasets were no longer required for the CFIE task, and our method could generalize well to real clinical color fundus images. With the help of skip connections between the input and output of the generator in our GAN network and the non-reference classification prior loss, the visual quality of color fundus images was significantly improved with structural details and pathological features well-preserved, which is not only beneficial for downstream tasks, but also makes it straightforward for ophthalmologists to distinguish pathological features. Both quantitative and qualitative results demonstrated that our method achieved better enhancement performance than other GAN-based methods.

Author Contributions: Conceptualization, S.C., H.Z. and Q.Z.; methodology, S.C., H.Z. and Q.Z.; software, S.C. and Q.Z.; validation, Q.Z.; formal analysis, S.C. and H.Z.; writing—original draft preparation, S.C. and H.Z.; writing—review and editing, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Technology Cooperation Project of The Xinjiang Production and Construction Corps under Grants 2019BC008 and 2017DB004.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were used in this study. The EyeQ dataset can be found here: <https://github.com/HzFu/EyeQ/tree/master/data> (accessed on 28 October 2021), The EyePACS dataset can be found here: <https://www.kaggle.com/c/diabetic-retinopathy-detection/data> (accessed on 29 October 2021).

Acknowledgments: The authors acknowledge funding from the Science and Technology Cooperation Project from The Xinjiang Production and Construction Corps (grants no. 2019BC008 and no. 2017DB004).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shen, Z.; Fu, H.; Shen, J.; Shao, L. Modeling and Enhancing Low-Quality Retinal Fundus Images. *IEEE Trans. Med. Imaging* **2020**, *40*, 996–1006. [[CrossRef](#)] [[PubMed](#)]
2. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. SwinIR: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 1833–1844.
3. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353. [[PubMed](#)]
4. Li, R.; Pan, J.; Li, Z.; Tang, J. Single image dehazing via conditional generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8202–8211.
5. Wei, P.; Wang, X.; Wang, L.; Xiang, J. SIDGAN: Single Image Dehazing without Paired Supervision. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 2958–2965.
6. Ren, D.; Zuo, W.; Hu, Q.; Zhu, P.; Meng, D. Progressive image deraining networks: A better and simpler baseline. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3937–3946.
7. Bulat, A.; Yang, J.; Tzimiropoulos, G. To learn image super-resolution, use a gan to learn how to do image degradation first. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 185–200.
8. Chen, C.; Xiong, Z.; Tian, X.; Zha, Z.; Wu, F. Camera Lens Super-Resolution. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 1652–1660.
9. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.
10. Wolf, V.; Lugmayr, A.; Danelljan, M.; Gool, L.; Timofte, R. DeFlow: Learning Complex Image Degradations from Unpaired Data with Conditional Flows. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021.
11. Yuan, Y.; Liu, S.; Zhang, J.; Zhang, Y.; Dong, C.; Lin, L. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 701–710.
12. Zhao, T.; Ren, W.; Zhang, C.; Ren, D.; Hu, Q. Unsupervised degradation learning for single image super-resolution. *arXiv* **2018**, arXiv:1812.04240.
13. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, arXiv:1808.04560.
14. Zhang, Y.; Zhang, J.; Guo, X. Kindling the darkness: A practical low-light image enhancer. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 1632–1640.
15. Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. EnlightenGAN: Deep Light Enhancement Without Paired Supervision. *IEEE Trans. Image Processing* **2021**, *30*, 2340–2349. [[CrossRef](#)] [[PubMed](#)]
16. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
17. Shao, Y.; Li, L.; Ren, W.; Gao, C.; Sang, N. Domain adaptation for image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2808–2817.

18. Zhang, K.; Zuo, W.; Gu, S.; Zhang, L. Learning Deep CNN Denoiser Prior for Image Restoration. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2808–2817.
19. McCann, J. Retinex Theory. In *Encyclopedia of Color Science and Technology*; Luo, M.R., Ed.; Springer: New York, NY, USA, 2016; pp. 1118–1125.
20. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Processing* **2007**, *16*, 2080–2095. [[CrossRef](#)] [[PubMed](#)]
21. Wang, W.; Wei, C.; Yang, W.; Liu, J. GLADNet: Low-Light Enhancement Network with Global Awareness. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 751–755.
22. Zhao, Z.; Xiong, B.; Wang, L.; Ou, Q.; Yu, L.; Kuang, F. RetinexDIP: A Unified Deep Framework for Low-light Image Enhancement. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *32*, 1076–1088. [[CrossRef](#)]
23. Guo, C.; Li, C.; Guo, J.; Loy, C.C.; Hou, J.; Kwong, S.; Cong, R. Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 1780–1789.
24. Cheng, J.; Li, Z.; Gu, Z.; Fu, H.; Wong, D.W.K.; Liu, J. Structure-preserving guided retinal image filtering and its application for optic disk analysis. *IEEE Trans. Med. Imaging* **2018**, *37*, 2536–2546. [[CrossRef](#)] [[PubMed](#)]
25. You, Q.; Wan, C.; Sun, J.; Shen, J.; Ye, H.; Yu, Q. Fundus Image Enhancement Method Based on CycleGAN. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 4500–4503.
26. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
27. Woo, S.H.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the 15th European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
28. Ma, Y.; Liu, J.; Liu, Y.; Fu, H.; Hu, Y.; Cheng, J.; Qi, H.; Wu, Y.; Zhang, J.; Zhao, Y. Structure and Illumination Constrained GAN for Medical Image Enhancement. *IEEE Trans. Med. Imaging* **2021**, *40*, 3955–3967. [[CrossRef](#)] [[PubMed](#)]
29. Cheng, P.; Lin, L.; Huang, Y.; Lyu, J.; Tang, X. Prior Guided Fundus Image Quality Enhancement Via Contrastive Learning. In Proceedings of the 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 13–16 April 2021; pp. 521–525.
30. Wang, J.; Li, Y.-J.; Yang, K.-F. Retinal fundus image enhancement with image decomposition and visual adaptation. *Comput. Biol. Med.* **2021**, *128*, 104116. [[CrossRef](#)] [[PubMed](#)]
31. Zhang, S.; Webers, C.A.B.; Berendschot, T.T.J.M. A double-pass fundus reflection model for efficient single retinal image enhancement. *Signal Processing* **2022**, *192*, 108400. [[CrossRef](#)]
32. Raj, A.; Shah, N.A.; Tiwari, A.K. A novel approach for fundus image enhancement. *Biomed. Signal Processing Control.* **2022**, *71*, 103208. [[CrossRef](#)]
33. Jolicoeur-Martineau, A. The relativistic discriminator: A key element missing from standard GAN. *arXiv* **2018**, arXiv:1807.00734.
34. Mao, X.D.; Li, Q.; Xie, H.R.; Lau, R.Y.K.; Wang, Z.; Smolley, S.P. Least Squares Generative Adversarial Networks. In Proceedings of the 16th IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2813–2821.
35. Li, C.; Fu, H.; Cong, R.; Li, Z.; Xu, Q. Nul-Go: Recursive Non-Local Encoder-Decoder Network for Retinal Image Non-Uniform Illumination Removal. In Proceedings of the 28th ACM International Conference on Multimedia, Seattle, WA, USA, 12–16 October 2020; pp. 1478–1487.
36. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
37. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
38. Fu, H.; Wang, B.; Shen, J.; Cui, S.; Xu, Y.; Liu, J.; Shao, L. Evaluation of retinal image quality assessment networks in different color-spaces. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; pp. 48–56.
39. Cuadros, J.; Bresnick, G. EyePACS: An adaptable telemedicine system for diabetic retinopathy screening. *J. Diabetes Sci. Technol.* **2009**, *3*, 509–516. [[CrossRef](#)] [[PubMed](#)]
40. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
41. Park, T.; Efros, A.A.; Zhang, R.; Zhu, J.-Y. Contrastive learning for unpaired image-to-image translation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 319–345.
42. Li, L.; Verma, M.; Nakashima, Y.; Nagahara, H.; Kawasaki, R. Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass, CO, USA, 1–5 March 2020; pp. 3656–3665.