A Clinical-oriented Lightweight Network for High-resolution Medical Image Enhancement

Yaqi Wang[1,2,†], Leqi Chen[1,2,†], Qingshan Hou[1,2,†], Peng Cao[1,2,3,(✉)], Jinzhu Yang[1,2,3,(✉)], Xiaoli Liu[1], and Osmar R. Zaiane[4]

[1] Computer Science and Engineering, Northeastern University, Shenyang, China
[2] Key Laboratory of Intelligent Computing in Medical Image of Ministry of Education, Northeastern University, Shenyang, China
[3] National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Shenyang, China
caopeng@mail.neu.edu.cn
yangjinzhu@cse.neu.edu.cn
[4] Alberta Machine Intelligence Institute, University of Alberta, Edmonton, Canada

Abstract. Medical images captured in less-than-optimal conditions may suffer from quality degradation, such as blur, artifacts, and low lighting, which potentially leads to misdiagnosis. Unfortunately, state-of-the-art medical image enhancement methods face challenges in both high-resolution image quality enhancement and local distinct anatomical structure preservation. To address these issues, we propose a Clinical-oriented High-resolution Lightweight Medical Image Enhancement Network, called CHLNet, which proficiently addresses high-resolution medical image enhancement, detailed pathological characteristics, and lightweight network design simultaneously. More specifically, CHLNet comprises two main components: 1) High-resolution Assisted Quality Enhancement Network for removing global low-quality factors in high-resolution images thus enhancing overall image quality; 2) High-quality-semantic Guided Quality Enhancement Network for capturing semantic knowledge from high-quality images such that detailed structure preservation is enforced. Moreover, thanks to its lightweight design, CHLNet can be easily deployed on medical edge devices. Extensive experiments on three public medical image datasets demonstrate the effectiveness and superiority of CHLNet over the state-of-the-art.

Keywords: Medical Image Enhancement · High-resolution Image · Light weight Network.

1 Introduction

Medical images play a vital role in the clinical analysis of various diseases [7]. Image quality is critical for accurate diagnosis and disease screening. However,

† Yaqi Wang, Leqi Chen and Qingshan Hou contribute equally to this work.

diverse imaging devices and non-standard manual processes contribute to the presence of numerous low-quality images in clinical datasets, potentially leading to inaccurate clinical diagnoses. Therefore, it is imperative to develop reliable methods for medical image enhancement.

Initially, traditional image enhancement methods, including LIME [9] and Fu et al. [8] have been attempted to enhance medical images. Nevertheless, these methods lack sensitivity to pathology details are not generalizable in clinical scenarios. More recently, medical image enhancement has made significant progress with the advancement of deep learning techniques. For instance, Cheng et al. [2] present a semi-supervised enhancement framework using pixel-wise importance maps for guided learning. Additionally, a coarse-to-fine enhancement framework [16] based on a diffusion model is proposed and achieves state-of-the-art performance. However, these methods suffer from downsizing images to a fixed low resolution, primarily due to hardware constraints, which hinder enhancement networks from capturing global low-quality factors in high-resolution images. Although some methods [5,11] have employed local image patches for enhancing and restoring the original image, aiming to preserve more image details. They potentially lead to discontinuities and distortions at patch boundaries, thus impacting the recognition of anatomical landmarks and disease-related lesions. Alternatively, other methods [19,5] use GANs to learn suitable mappings from a low-quality domain to a high-quality domain, but they tend to generate undesirable distorted images with pseudo-pathological structures, which is unacceptable for clinical medical images. Taking these into consideration, two questions naturally arise: 1) how to break the traditional paradigm of enhancing low-resolution images and pioneer high-resolution image enhancement to improve global image quality, and 2) how to model rich semantic information to restore local pathological details in reconstructed images?

To circumvent the above issues, we propose a novel **C**linical-oriented **H**igh-resolution **L**ightweight Medical Image Enhancement **Net**work, termed CHLNet, consisting of a High-resolution Assisted Quality Enhancement Network (HAQE) and a High-quality-semantic Guidance Quality Enhancement Network (HGQE). Specifically, HAQE adopts low-quality, low-resolution images as input and leverages Parallel Dilated Convolutional Attention Blocks (PAB) to construct a high-quality field instead of directly enhancing the image. The high-quality field is then combined with the input to generate a low-resolution enhanced image, and it can also be super-resolved to match low-quality, high-resolution images, leading to dual outputs that enhance both low and high-resolution spaces, ultimately improving global image quality. Furthermore, HGQE utilizes knowledge distillation [10] to guide the intermediate semantic feature distributions of HAQE. This ensures the capture of high-quality semantic knowledge to preserve local anatomical landmarks and disease-related lesion information. Additionally, thanks to the lightweight design, CHLNet can be easily deployed on medical edge devices.

To the best of our knowledge, this is the first work to design an enhancement network tailored for high-resolution medical images. In summary, the main contributions of this work are three-fold: 1) We propose a novel clinical-oriented

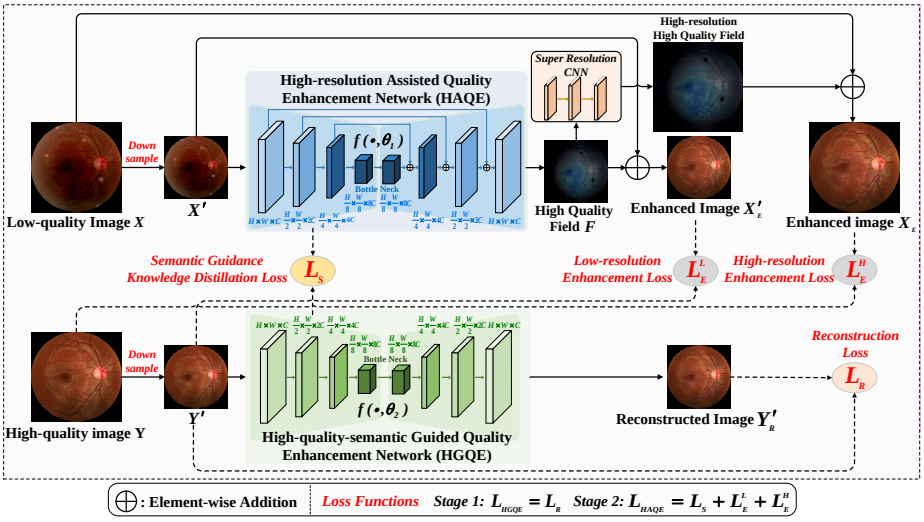

Fig. 1. Overview of CHLNet. **Stage 1:** Train the High-quality-semantic Guided Quality Enhancement Network $f(\cdot, \theta_2)$ constrained by the reconstruction loss $\mathcal{L}_R$. **Stage 2:** Train the High-resolution Assisted Quality Enhancement Network $f(\cdot, \theta_1)$ constrained by the low-resolution and high-resolution enhancement loss $\mathcal{L}_E^L$ and $\mathcal{L}_E^H$ to remove low-quality factors from the global high-resolution image. Additionally, the semantic guidance knowledge distillation loss $\mathcal{L}_E^L$ enforces semantic consistency between $f(\cdot, \theta_1)$ and $f(\cdot, \theta_2)$, thereby ensuring detailed pathological and lesion information.

high-resolution lightweight framework with PAB blocks capable of processing high-resolution inputs, capturing global pathological structures, and removing various low-quality factors. 2) We design two enhancement networks from global and local perspectives, namely HAQE and HGQE. The former aims to capture sufficient high-resolution global structural information, while the latter facilitates the former's learning by efficiently transferring local high-quality semantic knowledge. 3) Extensive quantitatively and qualitatively comparison experiments are conducted on three medical image datasets. The results prove that CHLNet significantly outperforms the previous best methods.

2 Methodology

An overview of our proposed framework, CHLNet, is shown in Fig. 1. We first elaborate on the overall enhancement process (Sect. 2.1) and then concentrate on the details of HAQE (Sect. 2.2) and HGQE (Sect. 2.3), respectively.

2.1 Overall Enhancement Process

Given a training dataset $\mathcal{D} = \{(X_i, Y_i)_{i=1}^N\}$, where X_i denotes the i_{th} high-resolution, low-quality image, Y_i corresponding to its high-quality image, and

N is the number of samples, respectively. We first down-sample the dataset to $X_i' \in \mathbb{R}^{H\times W\times C}$ and $Y_i' \in \mathbb{R}^{H\times W\times C}$ with scaling factor c. The overall framework consists of two stages. In the first stage, we exclusively utilize the high-quality image Y_i' as the input of the HGQE network $f(\cdot,\theta_2)$ to generate the reconstructed image Y_{R_i}' and train the network self-supervised. Thus, it can be formulated as $min_{\theta_2}\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_R(f(X_i,\theta_2),Y_i)$, where $\mathcal{L}_R=\mathbb{E}[||f(X_i,\theta_2),Y_i||_1]$ presents the L_1 reconstruction loss. In the second stage, the HAQE network $f(\cdot,\theta_1)$ processes the input image X_i' to generate the high-quality fields F_i, which is then combined with the input to produce a low-resolution enhanced image X_{E_i}'. Additionally, F_i can be super-resolved to match the size of low-quality, high-resolution image X_i through pre-trained Super Resolution CNN [6], and combined to produce high-quality, high-resolution enhanced image X_{E_i}. Note that the Super Resolution operation can preserve more image details compared to naive up-sampling. Thus, the procedure can be expressed as:

$$F_i=f(X_i',\theta_1); X_{E_i}'=X_i'+F_i; X_{E_i}=X_i+SuperResolutionCNN(F_i). \quad (1)$$

In this way, the dual enhancement results X_{E_i}' and X_{E_i} in both low and high-resolution spaces facilitate HAQE $f(\cdot,\theta_1)$ in capturing global structural information, ultimately improving overall high-resolution image quality. Therefore, the losses of $\mathcal{L}_E^L$ and $\mathcal{L}_E^H$ in both resolution spaces are defined as:

$$\mathcal{L}_E^L=\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_{En}(X_{E_i}',Y_i'); \mathcal{L}_E^H=\frac{1}{N}\sum_{i=1}^{N}\mathcal{L}_{En}(X_{E_i},Y_i),$$

$$\begin{aligned}\mathcal{L}_{En}(x,y)&=\mathcal{L}_1(x,y)+\mathcal{L}_{SSIM}(x,y)+\mathcal{L}_{HF}(x,y)\\&=\mathbb{E}[||x,y||_1]+\mathbb{E}[1-SSIM(x,y)]+\mathbb{E}[||\psi_{HF}(x),\psi_{HF}(y)||_1]\end{aligned} \quad (2)$$

where $\mathcal{L}_{En}$ denotes the enhancement loss. It is composed of three parts: First, the $\mathcal{L}_1$ loss is the basic pixel-level loss, preserving the consistency between the enhanced and high-quality images. Second, the $\mathcal{L}_{SSIM}$ loss reflects the human visual perception of image quality, aiming to enforce consistency in brightness, contrast, and structure between the enhanced and high-quality images. Finally, the $\mathcal{L}_{HF}$ loss extracts the high-frequency information (such as anatomical structures and lesion details) of medical images using a high-pass filter $\psi_{HF}(\cdot)$ and suppresses blurry low-frequency information. Moreover, to better preserve local pathological details, we ensure semantic consistency between pre-trained HAQE $f(\cdot,\theta_1)$ and HGQE $f(\cdot,\theta_2)$ through feature map knowledge distillation loss. Specifically, the encoder feature maps of HAQE and HGQE are denoted as $A=\{A_1,A_2,..,A_M\}$ and $G=\{G_1,G_2,..,G_M\}$, respectively, where M represents the total number of feature map layers. Hence, the semantic guidance knowledge distillation loss is defined as $\mathcal{L}_S=\frac{1}{N\times M}\sum_{i=1}^{N}\sum_{j=1}^{M}\mathcal{L}_{KD}(A_{ij},G_{ij})$, where $\mathcal{L}_{KD}$ represents the KL divergence loss.

Overall, the total loss function for the HAQE network can be represented as $\mathcal{L}_{HAQE}=\mathcal{L}_E^L+\mathcal{L}_E^H+\mathcal{L}_S$. Notably, Although CHLNet includes two networks, only the HAQE network is required in the testing stage.

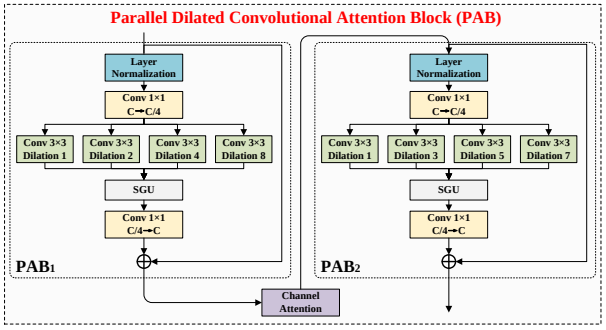

Fig. 2. The procedure of Parallel Dilated Convolutional Attention Block.

2.2 High-resolution Assisted Quality Enhancement Network

The low-quality factors in medical images comprise defocus, motion blur, etc., with uneven spatial distribution. To estimate the low-quality distribution for image enhancement, the model requires capturing a sufficiently large image region to obtain global illumination and structural information. To this end, we propose the High-quality-semantic Guided Quality Enhancement Network (HAQE) to balance computational efficiency and global information capture. As shown in Fig. 1, the HAQE network adopts a U-shaped architecture (encoder, bottleneck, and decoder) with well-designed Parallel Dilated Convolutional Attention Blocks (PABs). To ensure the preservation of fine structures and texture details in the enhanced images, residual connections are established between each decoder layer and its corresponding encoder layer. Additionally, the PixelShuffle method for upsampling in the decoder also facilitates retaining detailed information and reduces computational complexity. Notably, HAQE utilizes an asymmetric encoder-decoder architecture with fewer PABs in the decoder. This prioritizes the encoder for enhancement tasks and applies semantic-guided methods more effectively.

Figure 2 illustrates the procedure of PAB, consisting of two sub-blocks: PAB_1 and PAB_2. These sub-blocks include layer normalization, 1×1 convolution, multi-scale dilated convolutions, spatial gating unit (SGU) [4], channel attention, and residual connection. PAB_1 utilizes dilation rates $d_1 = [1, 2, 4, 8]$ to integrate receptive fields of different scales, while PAB_2 uses non-overlapping dilation rates $d_2 = [1, 3, 5, 7]$ to avoid grid artifacts. The dilated convolutions expand the receptive field without increasing model parameters, thus contributing to the lightweight character of the FAQE network. Subsequently, SGU aggregation dynamically adjusts weights to fuse feature maps for emphasizing pathological details and lesion information while eliminating low-quality factors. Moreover, a channel attention mechanism [12] is incorporated between the two PAC sub-blocks to automatically highlight important channels, thereby enhancing essential feature extraction from medical images.

2.3 High-quality-semantic Guided Quality Enhancement Network

Due to the difficulty of maintaining structural consistency with real medical images and the sensitivity to noise, enhancement models often generate undesired distorted images with pseudo-pathological structures. To preserve real anatomical landmarks and disease-related lesion information, we propose the High-quality-semantic Guided Quality Enhancement Network (HGQE) which is trained on high-quality images in a self-supervised manner (as described in Section 2.1). Inspired by knowledge distillation [10], the pre-trained HGQE can act as a teacher network, while HAQE serves as a student network. By aligning their feature map distributions through KL divergence, the knowledge from HGQE is transferred to HAQE, thus improving the preservation of intricate pathological details. Additionally, HGQE shares a similar architecture with HAQE but omits the residual connections between the encoder and decoder, which compels the encoder to preserve more high-quality information for better semantic knowledge transfer.

3 Experiments

3.1 Dataset and Implementation Details

We conduct experiments on the following three benchmark datasets: (1) the Real Fundus (RF) dataset [5] consists of 120 low-quality and high-quality clinical fundus image pairs; (2) the endoscopy dataset is derived from the publicly available CVC-EndoSceneStill dataset [18] for endoluminal scene enhancement, consisting of 624 high-quality images and 288 low-quality images; and (3) the skin lesion dataset is curated from the ISIC Challenge Dataset 2017 [3] provided by the International Skin Imaging Collaboration (ISIC), comprising 1635 high-quality images and 1115 low-quality images. Due to paired high-quality and low-quality images not existing in EndoSceneStill and ISIC datasets, the pipeline proposed by [2] is adopted to synthesize paired images. Additionally, all datasets are split into a 1:3 ratio of training and testing sets. For evaluation, we employ peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) as quantitative metrics to assess the performance of the enhancement methods. Higher PSNR and SSIM values indicate that the enhanced images closely resemble the reference images. **More training details are in the supplementary material.**

3.2 Comparison with the State-of-the-Art Methods

To evaluate the effectiveness of CHLNet, comprehensive quantitative and qualitative comparisons with state-of-the-art medical image enhancement methods are conducted. We compare CHLNet with traditional image enhancement methods, including LIME [9] and Fu et al. [8], as well as several deep learning methods: supervised methods (e.g., PCENet [15], ScrNet [13], DeepRFT [17]), unsupervised methods (e.g., ArcNet [14], StillGAN [16]), semi-supervised method

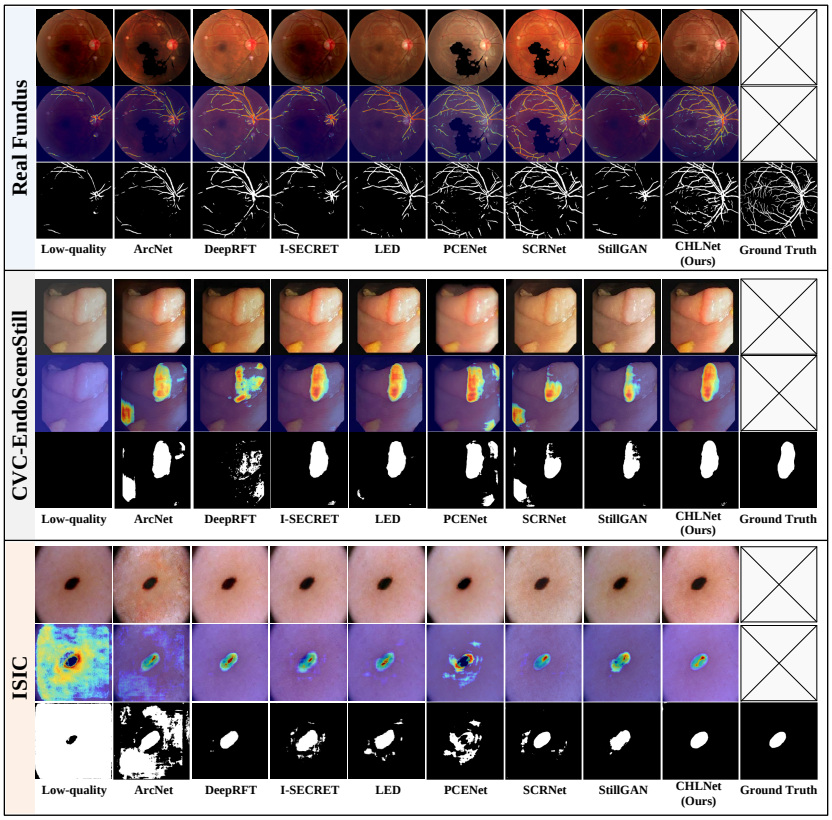

Fig. 3. Visual comparisons between comparable methods and CHLNet for image quality enhancement with corresponding heatmap and segmentation results.

I-SECRET [2], and diffusion model-based method LED [1]. From Table 1, it is evident that our method exhibits a significant advantage over both traditional and deep learning-based approaches in terms of PSNR and SSIM. Notably, CHLNet retains a relatively small number of parameters and FLOPs compared to other methods, making it preferable for deployment in medical edge devices. In addition, we also provide visualizations for different methods in Figure 3, where the enhancement images with corresponding heatmap and segmentation results are exhibited. It can be observed that our model enhances low-quality images by simultaneously considering both localized anatomical landmarks and overall image quality, resulting in superior visual outcomes compared to other methods. Thanks to CHLNet, the enhanced images exhibit clearer vessel/lesion structures, facilitating their identification by the segmentation model and leading to improved segmentation results, thereby benefiting clinical diagnosis.

Table 1. Quantitative comparisons with the State-of-the-art methods for enhancing low-quality medical images on three datasets, highlighting the best and second-best scores with bold and underlined, respectively.

Methods	Params(M)	Flops(G)	Real Fundus		CVC-EndoSceneStill		ISIC	
			PSNR(dB)	SSIM	PSNR(dB)	SSIM	PSNR(dB)	SSIM
LIME [9]	-	-	15.52	0.766	14.86	0.854	12.93	0.824
Fu et al. [8]	-	-	11.74	0.562	11.75	0.692	10.24	0.633
PCENet [15]	26.65	343.45	19.86	0.832	23.63	0.703	24.70	0.894
ArcNet [14]	54.42	72.81	22.85	0.824	22.05	0.802	19.23	0.851
SCRNet [13]	89.29	137.29	24.05	0.886	23.46	0.854	24.26	0.895
StillGAN [16]	78.64	268.35	24.97	0.845	27.34	0.914	24.89	0.885
DeepRFT [17]	12.65	363.19	28.35	0.902	31.64	<u>0.947</u>	26.44	<u>0.934</u>
I-SECRET [2]	11.76	228.76	28.18	0.896	<u>33.26</u>	0.934	<u>29.04</u>	0.898
LED [1]	113.62	993.58	<u>28.88</u>	<u>0.918</u>	32.68	0.923	27.58	0.821
CHLNet(Ours)	8.57	21.64	**30.41**	**0.942**	**34.54**	**0.958**	**34.51**	**0.935**

Table 2. Ablation study of the loss terms on Real Rundus dataset.

$\mathcal{L}_E^L$ [a]	$\mathcal{L}_E^H$ [b]	$\mathcal{L}_S$ [c]	PSNR	SSIM
✓			27.77	0.902
✓	✓		29.09	0.926
✓		✓	28.04	0.930
✓	✓	✓	**30.41**	**0.942**

[a] Low-resolution enhancement loss
[b] High-resolution enhancement loss
[c] Semantic guidance loss

Table 3. Ablation study of the various high-resolution image sizes on Real Fundus dataset.

HR[a] Image Sizes	LR[b] Image Sizes	PSNR	SSIM
320×320	256×256	28.87	0.941
640×640	256×256	29.83	**0.943**
1280×1280	256×256	**30.41**	0.942
2560×2560	256×256	29.39	0.932

[a] High-resolution
[b] Low-resolution

3.3 Ablation Study

We perform ablation experiments to assess the effectiveness of each component, the results of which are shown in Table 2. It is evident that both the proposed HAQE (corresponding to $\mathcal{L}_E^L$ and $\mathcal{L}_E^H$) and HGQE (corresponding to $\mathcal{L}_S$) boost the enhancement performance significantly. More specifically, the incorporation of the high-resolution enhancement loss $\mathcal{L}_E^H$ (the 2nd row) facilitates the capture of global image structural information, resulting in a significant improvement of 1.32 dB in PSNR and 2.4% in SSIM over the baseline. Subsequently, the introduction of the semantic guidance loss $\mathcal{L}_S$ (the 4th row) preserves pathological structures while enforcing structural consistency with high-quality images, ultimately achieving the optimal PSNR/SSIM values of 30.41 dB/94.2%. Additionally, table 3 illustrates the performance impact of high-resolution image sizes, showcasing a gradual performance increase as the resolution ranges from 320×320 to 1280×1280. This indicates that capturing a broader range of global structures benefits quality enhancement. However, further size increases lead to information loss and performance decline, thus we compromise by choosing 1280×1280 as the high-resolution image size.

4 Conclusion

In this paper, we propose a novel lightweight medical image enhancement network for enhancing high-resolution images, termed CHLNet. Benefiting from the guidance provided by high-quality and high-resolution images, the enhanced images are capable of ensuring global image quality, while preserving local anatomical landmarks and disease-related lesion information. Additionally, its lightweight design and high performance make it adaptable to a variety of clinical scenarios.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No.62076059) and the Science and Technology Joint Project of Liaoning province (2023JH2/101700367, ZX20240193).

Disclosure of Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Cheng, P., Lin, L., Huang, Y., He, H., Luo, W., Tang, X.: Learning enhancement from degradation: A diffusion model for fundus image enhancement. arXiv preprint arXiv:2303.04603 (2023)
2. Cheng, P., Lin, L., Huang, Y., Lyu, J., Tang, X.: I-secret: Importance-guided fundus image enhancement via semi-supervised contrastive constraining. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24. pp. 87–96. Springer (2021)
3. Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al.: Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). pp. 168–172. IEEE (2018)
4. Dauphin, Y.N., Fan, A., Auli, M., Grangier, D.: Language modeling with gated convolutional networks. In: International conference on machine learning. pp. 933–941. PMLR (2017)
5. Deng, Z., Cai, Y., Chen, L., Gong, Z., Bao, Q., Yao, X., Fang, D., Yang, W., Zhang, S., Ma, L.: Rformer: Transformer-based generative adversarial network for real fundus image restoration on a new clinical benchmark. IEEE Journal of Biomedical and Health Informatics **26**(9), 4645–4655 (2022)
6. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence **38**(2), 295–307 (2015)

7. 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: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22. pp. 48–56. Springer (2019)
8. Fu, X., Zhuang, P., Huang, Y., Liao, Y., Zhang, X.P., Ding, X.: A retinex-based enhancing approach for single underwater image. In: 2014 IEEE international conference on image processing (ICIP). pp. 4572–4576. IEEE (2014)
9. Guo, X., Li, Y., Ling, H.: Lime: Low-light image enhancement via illumination map estimation. IEEE Transactions on image processing **26**(2), 982–993 (2016)
10. Hinton, G., Vinyals, O., Dean, J.: Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 (2015)
11. Hou, Q., Cao, P., Wang, J., Liu, X., Yang, J., Zaiane, O.R.: A reference-free self-supervised domain adaptation framework for low-quality fundus image enhancement. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 7383–7393 (2023)
12. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)
13. Li, H., Liu, H., Fu, H., Shu, H., Zhao, Y., Luo, X., Hu, Y., Liu, J.: Structure-consistent restoration network for cataract fundus image enhancement. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 487–496. Springer (2022)
14. Li, H., Liu, H., Hu, Y., Fu, H., Zhao, Y., Miao, H., Liu, J.: An annotation-free restoration network for cataractous fundus images. IEEE Transactions on Medical Imaging **41**(7), 1699–1710 (2022)
15. Liu, H., Li, H., Fu, H., Xiao, R., Gao, Y., Hu, Y., Liu, J.: Degradation-invariant enhancement of fundus images via pyramid constraint network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 507–516. Springer (2022)
16. 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 Transactions on Medical Imaging **40**(12), 3955–3967 (2021)
17. Mao, X., Liu, Y., Liu, F., Li, Q., Shen, W., Wang, Y.: Intriguing findings of frequency selection for image deblurring. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 1905–1913 (2023)
18. Vázquez, D., Bernal, J., Sánchez, F.J., Fernández-Esparrach, G., López, A.M., Romero, A., Drozdzal, M., Courville, A., et al.: A benchmark for endoluminal scene segmentation of colonoscopy images. Journal of healthcare engineering **2017** (2017)
19. 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. pp. 2223–2232 (2017)