International Journal of Information Technology and Applications
Volume 1, Issue 1, March 2024 ISSN 3007-7397

LNet-GIE: A Lightweight Network for Gastroscopic Image Enhancement

Jiayan Huang[1,†], Xinxin Xu[1,†], Pengfei Cao[2], Tingting Li[1], Shaoye Luo[1], Junwu Lin[1], Fuquan Zhang[3,*]

[1]*New Engineering Industry College, Putian University, China*
[2]*Department of Automatic Control, Polytechnic University of Catalonia, Spain.*
[3]*School of Computer and Data Science, Minjiang University, China*

[†]*The first two authors contribute equally to this work*

jyan_huang@163.com, XxinXu04@163.com, pengfei95271@gmail.com, a1399098943@163.com, lsy123@163.com, 68210448@qq.com, zfq@mju.edu.cn

Abstract. Gastroscopic image enhancement is an important research direction in the field of medical image processing, which aims at improving the image quality, and enhancing the accuracy of clinical diagnosis. However, existing algorithms are generally designed with large-scale deep learning network structures, which require high computing resources and storage space, and thus it is difficult to achieve their practical application and scalability. In this paper, we proposed a lightweight network for gastroscopic image enhancement (LNet-GIE). The LNet-GIE first loads the pre-trained weights for initialization, and then passes them into the constructed lightweight network, only designed with common convolutional layers. Quantitative and qualitative experiments on public gastroscopic image enhancement dataset show that the size of LNet-GIE is significantly smaller, while the performances are better than the state-of-the-art methods. The code of the proposed LGNet-GIE is public and available at https://github.com/20618xx/LNet-GIE.

Keywords: Gastroscopic image enhancement, Lightweight network, Model deployment

1. Introduction

Gastroscopic images are obtained by inserting an endoscope into the patient's stomach to capture images for observation, diagnosis, and evaluation of gastric diseases. However, unprocessed gastroscopic images may be affected by factors such as smoke, which will reduce the accuracy of clinical diagnosis and the feasibility of medical research. Gastroscopic image enhancement plays a crucial role in improving the diagnosis, treatment, and research of gastric diseases. It provides clinicians with more accurate information, leading to improved medical outcomes.

Existing image enhancement methods can be briefly divided into traditional methods and deep learning-based methods. Common traditional methods include histogram equalization [1, 2], spatial filtering [3], [4], and sharpening enhancement [5]. Histogram equalization adjusts the pixel distribution of an image to enhance its contrast and brightness, but it may amplify image noise and lead to over-enhancement and distortion. Spatial filtering improves the quality and visual effects of an image by performing pixel-level operations in the original space of the image, but it has limitations in processing complex image features. Sharpening enhancement improves the clarity of an image by enhancing its edges and details, but excessive sharpening can introduce artifacts and noise amplification.

To overcome the limitations of traditional methods, deep learning-based image enhancement methods have been proposed. These methods leverage advantages, such as automatic feature learning [6, 7], context awareness [8], and end to-end optimization [9, 10], to provide more accurate and efficient image enhancement. Automatic feature learning can automatically capture the complex features and structural information from an image, but it may overfit the training data and is not suitable for small-scale datasets and resource constrained platforms. Context-aware models can leverage the contextual information in an image to model its global structure and content, but introducing context information may also introduce erroneous information leading to inaccurate results. End-to-end optimization employs direct mapping to minimize manually designed intermediate steps and parameter settings for simplifying the usage and tuning process. However, such model's decision-making process and working principles is challenging to be interpreted. Moreover, end-to-end training requires significant computational resources and running times.

To address these existing issues, we propose a lightweight network for gastroscopic image enhancement (LNet-GIE). The LNet-GIE is with small model size, which greatly lowers the requirements for computational resources and storage space, and thus it facilitates deployment in resource-constrained platforms. Specifically, the LNet-GIE first loads the weights of a pre-trained model, which was trained on large-scale datasets, to better extract the features from gastroscopic images. Then, two convolutional layers are used to further extract higher-level feature representations for subtle details and structural information enhancement. Next, the ReLU activation function are used to enhance the edges and texture information and improve the clarity and contrast of gastroscopic images. Finally, a haze removal operation is conducted to process the extracted features and generate the enhanced clean images. The main contributions of this paper are:

(1) We propose a lightweight network for gastroscopic image enhancement tasks (LNet-GIE). The LNet-GIE is with small model scale and low computational resource requirements, which can be easily deployed and implemented on various hardware platforms.
(2) We initialize the LNet-GIE by loading the weights of a pre-trained model, which can accelerate the training process and reduce the model size.
(3) Qualitative and quantitative experiments demonstrate that our LNet-GIE outperforms several comparative methods in terms of performance. Moreover, the LNet-GIE exhibits significantly smaller model memory compared to other models.

2. Related Works

Most existing dehazing methods are proposed for image enhancement tasks, especial for foggy images [22-25]. However, there are few researches focuses on medical images, and most of the existing few methods for medical image enhancement usually use traditional deep learning with large models[11-13]. With the rise of lightweight networks, it has become particularly important to apply them in the field of medical image enhancement. In this section, we will introduce traditional deep learning-based large models and lightweight network models.

2.1. Traditional Deep Learning-Based Large Models

Traditional deep learning-based large models typically increase model capacity by designing deeper and wider network architectures. Generally, it can increase the learning ability and improve the performance of the model by using more convolutional layers and fully connected layers. For instance, DehazeNet [9] adopts a deep architecture based on convolutional neural networks to learn patterns and features for haze removal from a large amount of data. However, large models often contain a large number of parameters, which requires more computing resources and memory for training and inference, resulting in high computational and storage costs. Moreover, the complexity of such network making it difficult to understand and interpret. At the same time, large models are prone to overfitting data and have a high demand for data. In the case of limited equipment, deployment and inference also become difficult. These issues usually be addressed with new methods and technologies, such as designing lightweight network structures, compressing and pruning models.

2.2. Lightweight Networks

With the continuous improvement of computing speed and memory requirements, lightweight networks have become an unstoppable trend in mainstream neural networks. Lightweight networks aim to further reduce the number of model parameters and complexity while maintaining model accuracy. For instance, the MobileNet and SqueezeNet have much smaller in memory than traditional deep learning-based large models, and they are no significant difference in performance, and even better than traditional large models. Unfortunately, lightweight network is an ideal choice for efficient deep learning in resource limited scenarios, there are few research focuses on degining lightweight networks for gastroscopic image enhancement.

3. Proposed Method

In this section, we will introduce the overall structure and designed details of our LNet-GIE for gastroscopic image enhancement.

3.1. Overall Structure

In this paper, we proposed a lightweight network for gastroscopic image enhancement, named LNet-GIE. The structure of the LNet-GIE as shown in Figure 1. Specifically, given hazy gastroscopic images with three channels, the LNet-GIE first loads the weights and bias from the pre-trained AODNet [7] for model initialization. This process aims to adjust the network to adapt to new tasks while retaining some information of the pre-trained model and accelerate training speed, especially when limited training data is available. Then, a 1 × 1 convolutaional layer is used to reduce 3 channels to 1 channel, and followed by a 3 × 3 convolutional layer for further feature processing. Next, the ReLU activation function is applied to introduce nonlinearity for extracting complex features. Finally, nonlinear operation is performed to dehaze for the obtained feature maps and generate the enhanced clean image. Given the hazy gastroscopic image x, the the output enhanced result y of the proposed LNet-GIE can be represented as follows,

$$y = Relu((f^{(2)} \times x) - f^{(2)} + 1) \tag{1}$$

where $f^{(2)}$ denotes the output feature maps after the 2nd convolutional layer (3×3).

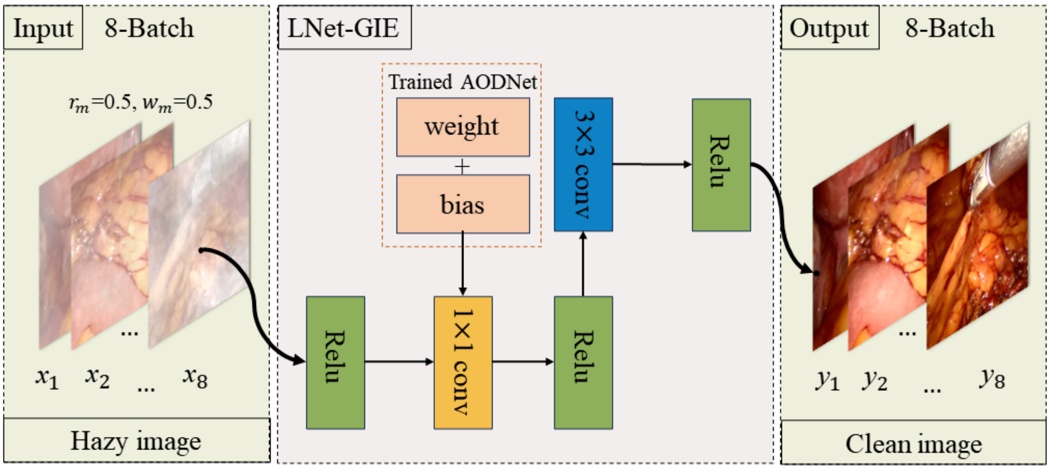

Figure 1. The structure of LNet-GIE.

3.2. Parameter Setting

To maintain model lightweight while preserving image details and quality as possible, we introduce the resolution multiples rm and width multiples wm for the input images. In particular, lower rm value can reduce the size of the input, thereby reducing computational complexity and the impact of haze. And lower wm can reduce the parameter number and computational load, thereby improving the running speed and memory efficiency of the LNet-GIE. Notably, the rm and wm are all set to 0.5 in our specific experiments. Although the selection of the two parameters were achieved good results in our task, it should be noted that the optimal values of them may vary depending on different datasets and tasks. In practical applications, adjustments and optimizations need to be made according to specific circumstances to meet performance requirements while minimizing computational costs to achieve optimal results.

3.3. Pre-trained Weights Loading

To effective extract the features and improve the training speed of our LNet-GIE, we used the pre-trained weights from AODNet to initialize it. Specifically, the pseudo code of the pre-trained weights loading is shown in Procedure 1. By this process, it can ensure the LNet-GIE has the feature extraction ability of the pre-trained model for effective dehazing, while reducing model storage requirements and accelerating training speed.

Procedure 1. Pre-trained weights loading for the LNet-GIE.

Input: the pre-trained weights from AODNet
1. Creating the dehazing network object by instantiating the *dehaze_net* class
2. Loading the weights from file (AODNet.pth) and storing on disk in the form of a state dictionary (*state_dict*)
3. Re-loading the pre-trained weights and bias of *e_uconv1* layer on the state dictionary for *self.e_uconv1*
4. Copying part of *e_uconv1.weght* and *e_uconv1.bias* of the pre-trained AODNet to the corresponding parameters of the LNet-GIE

Output: the initialized LNet-GIE with pre-trained weights

4. Experimental Results

In this section, we will introduce the experiment settings, dataset and evaluation metrics, and the gastroscopic image enhanced result comparisons.

4.1. Experimental Settings

The proposed LNet-GIE is coded under the PyTorch framework with PyThon 3.7 version and trained on NVIDIA GeForce RTX 3050 GPU. The initial learning rate is set to 0.01, and the Adam optimizer adaptively adjusts it during the training processing. The total training times are 30 epochs.

4.2. Dataset and Evaluation Metrics

We used the public dataset [14], which is specifically designed for image enhancement. The dataset contains 2700 clear-foggy image pairs. To fully utilize different samples in the dataset, we used 2160 pairs as training set and the trained 540 as testing set. We used the common full-reference metrics PSNR and SSIM to quantitatively evaluate the enhanced results.

4.3. Enhanced Result Comparisons

To evaluate the superiority of the proposed LNet-GIE, we compared it with nine compared methods in qualitatively and quantitatively, including AODNet [7], C2PNet [14], CAP [15], DA_dehazing [16], DCP [17], FFANet [18], GCANet [19], GridDehazeNet [20], and MSBDN-DFF [21]. It is noted that all of them are proposed for image dehazing task.

4.3.1. Qualitative Comparisons

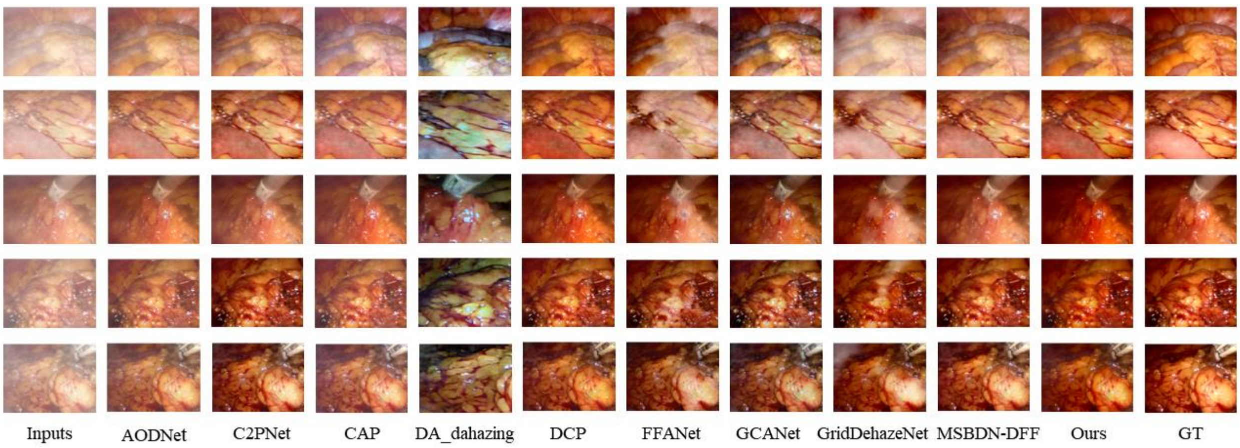

Figure 2. Qualitative comparisons of different methods.

The qualitative compared results are shown in Figure 2, we selected five hazy gastroscopic images for comparison. From the enhanced results, it is evident that AODNet, C2PNet, CAP, FFANet, GridDehazeNet, and MSBDN-DFF still suffer from haze residual. DCP and GCANet are excessive enhancement. DA_dehazing has significant color differences after enhancement. On the contrary, our LNet-GIE effectively eliminates the haze, avoids over-enhancement, and preserves the natural colors of the images. Our method achieves a better balance between enhancement effects and image naturalness, which reflects its effectiveness and practicality. In the other words, the effective dehazing results provides strong support for research and applications in the field of gastroscopy image enhancement.

4.3.2. Quantitative Comparisons

To quantitatively evaluate the performance of the proposed LNet-GIE, we employed PSNR and SSIM to measure the differences between the ground truth clear images and the enhanced images. Higher PSNR and SSIM values indicate a higher similarity between the enhanced images and the ground truth clear images. Table 1 presents the PSNR and SSIM values of our LNet-GIE and the nine compared methods, it can be seen that our LNet-GIE has similar enhancement effects as most compared methods even with a minimalist network structure. This validates the effectiveness and practicality of our method, which can establish a solid foundation for its broader application in practice.

Table 1. Quantitative comparisons of different methods.

Methods	PSNR	SSIM
AODNet	28.0039	0.7155
C2PNet	28.1190	0.7207
CAP	27.9454	0.7169
DA_dehazing	27.9394	0.4151
DCP	28.3117	0.7882
FFANet	27.9982	0.6933
GCANet	28.4199	0.7929
GridDehazeNet	28.2590	0.7467
MSBDN-DFF	28.7328	0.8123
Ours	28.1082	0.7452

Table 2. The parameter number and memory requirements comparisons of different methods.

Methods	Parameter Number	Memory Requirement (KB)
AODNet	1761	11
C2PNet	7169327	40459
CAP	8140000	1228
DA_dehazing	38720	213355
DCP	5343040	21829
FFANet	4455913	4789
GCANet	702818	2758
GridDehazeNet	958051	983
MSBDN-DFF	36845053	122641
Ours	14	2

4.4. Lightweight Validation of the LNet-GIE

To demonstrate the lightweight of our LNet-GIE, we compared the parameter number and memory requirements of different methods. As shown in Table 2, our LNet-GIE has the smallest model size with the minimum parameter number of 14 and memory requirements of 2KB. Compared with the best baseline AODNet, the number of parameters and the memory requirements have been reduced by thousands of times and 5.5 times, respectively. Our LNet-GIE requires minimal computing and storing resources and is more suitable to be deployed in some resource limited platform. In the other words, the lightweight of the LNet-GIE make it have great potential application prospects in real-time gastroscopic image enhancement.

5. Conclusion

In this paper, we proposed a lightweight network for gastroscopic image enhancement (LNet-GIE), which aims to reduce the model size for deploying under limited computing and memory platforms. By loading the weights of a pretrained model and designing easy structure, the LNet-GIE greatly speeds up the computation and reduces the memory usage without significant performance degradation. Experiments show that our LNet-GIE has achieved parameter number and memory requirements of up thousands and five times, respectively. Compared with other methods, our LNet-GIE has lower computational and memory cost while ensuring significant enhancement effects. This means that the LNet-GIE has great potential application significance for achieving real-time gastroscopic image enhancement.

Acknowledgements

This work was jointly supported by the Young and Middle-aged Teacher Education Research Project of Fujian Province (Science and Technology, Project No. JAT220829, JAT210410), and the Fujian Provincial Natural Science Foundation Project (2022J011170).

References

[1] Meng X, Huang J, Li Z, et al. DedustGAN: Unpaired learning for image dedusting based on Retinex with GANs[J]. Expert Systems with Applications, 2024, 243: 122844

[2] Yan J, Wang Y, Fan H, et al. LEPF-Net: Light Enhancement Pixel Fusion Network for Underwater Image Enhancement[J]. Journal of Marine Science and Engineering, 2023, 11(6): 1195.

[3] Fu A, Ma J, Wang C, et al. Traditional Chinese Medicine Health Status Identification with Graph Attention Network[C]//International Conference on Machine Learning for Cyber Security. Cham: Springer Nature Switzerland, 2022: 1-14.

[4] Wang C, Grau A, Guerra E, et al. Semi-supervised wildfire smoke detection based on smoke-aware consistency[J]. Frontiers in .

[5] Huang J, Xu H, Liu G, et al. SIDNet: a single image dedusting network with color cast correction[J]. Signal Processing, 2022, 199: 108612.

[6] Wang C, Hu J, Luo X, et al. Color-dense illumination adjustment network for removing haze and smoke from fire scenario images[J]. Sensors, 2022, 22(3): 911.

[7] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-One Dehazing Network. Proceedings of the IEEE International Conference on Computer Vision, pp. 4770–4778, 2017.

[8] Huang J, Li Z, Wang C, et al. FFNet: A simple image dedusting network with feature fusion[J]. Concurrency and Computation: Practice and Experience, 2021, 33(24): e6462.

[9] Feng T, Wang C, Chen X, et al. URNet: A U-Net based residual network for image dehazing[J]. Applied Soft Computing, 2021, 102: 106884.

[10] Huang J, Li Z, Wang C. Image dust storm synthetic method based on optical model[C]//Machine Learning for Cyber Security: Third International Conference, ML4CS 2020, Guangzhou, China, October 8–10, 2020, Proceedings, Part III 3. Springer International Publishing, 2020: 215-226.

[11] Zhang F, Wang C. MSGAN: generative adversarial networks for image seasonal style transfer[J]. IEEE Access, 2020, 8: 104830-104840.

[12] Wang C, Li Z, Wu J, et al. Deep residual haze network for image dehazing and deraining[J]. IEEE Access, 2020, 8: 9488-9500.

[13] Zhou X, Wang C, Li Z, et al. Adaptive histogram thresholding-based leukocyte image segmentation[C]//Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceedings of the 15th International Conference on IIH-MSP in conjunction with the 12th International Conference on FITAT, July 18–20, Jilin, China, Volume 2. Springer Singapore, 2020: 451-459.

[14] Yu Zheng, Jiahui Zhan, Shengfeng He, Junyu Dong, and Yong Du. Curricular Contrastive Regularization for Physics-aware Single Image Dehazing. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5785–5794, 2023

[15] Qingsong Zhu, Jiaming Mai, and Ling Shao. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior. IEEE Transactions on Image Processing, vol. 24, no.11, pp. 3522–3533, 2015.

[16] Yuanjie Shao, Lerenhan Li, Wenqi Ren, Changxin Gao, and Nong Sang. Domain Adaptation for Image Dehazing. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2808– 2817, 2020.

[17] Kaiming He, Jian Sun, and Xiaoou Tang. Single Image Haze Removal Using Dark Channel Prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2341–2353, 2010.

[18] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature Fusion Attention Network for Single Image Dehazing. Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 11908–11915, 2020.

[19] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Liheng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua. Gated Context Aggregation Network for Image Dehazing and Deraining. Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1375– 1383, 2019.

[20] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-Based Multi-Scale Network for Image Dehazing. Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 7314–7323, 2019.

[21] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-Scale Boosted Dehazing Network with Dense Feature Fusion. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2157–2167, 2020.

[22] Wang C, Fan H, Zhang H, et al. Pixel-Level dehazed image quality assessment based on dark channel prior and depth[C]//2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/

SustainCom). IEEE, 2019: 1545-1549.

[23] Wang C, Zhang H, Li Z, et al. White blood cell image segmentation based on color component combination and contour fitting[J]. Current Bioinformatics, 2020, 15(5): 463-471.

[24] Feng T, Li Z, Wang C, et al. Image dehazing network based on dilated convolution feature extraction[C]//2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE, 2019: 1-5.

[25] Zhang H, Wang C, Li X, et al. Design and implementation of an infrared radiant source for humidity testing[J]. Sensors, 2018, 18(9): 3088.