FMSNet：Underwater Image Restoration by Learning from a Synthesized Dataset

Xiangyu Yin[1[0000-0002-8363-648X]], Xiaohong Liu[2[0000-0001-6377-4730]], Huan Liu[2]

[1] Guilin University of Electronic Technology, Guilin, China
yinx36@mcmaster.ca
[2] McMaster University, Hamilton, Canada

Abstract. Underwater images suffer from various degradation, which can significantly lower the visual quality and the accuracy of subsequent applications. Moreover, the artificial light source tends to invalidate many image restoration algorithms. In this paper, an underwater image restoration (UIR) method using a novel Convolutional Neural Network (CNN) architecture and a synthesized underwater dataset is proposed. We discuss the reason for the over enhancement that exists in current UIR methods and revise the underwater image formation model (IFM) to alleviate the problem. With the revised IFM, we proposed an underwater image synthesizing method that can create a realistic underwater dataset. In order to effectively conduct end-to-end supervised learning, we design a network based on the characteristics of image restoration tasks, namely FMSNet. Different from existing networks, the decomposition and fusion operation in FMSNet can process the feature maps more efficiently and improve the contrast more prominently. The UIR method built by FMSNet can directly recover the degraded underwater images without the need of any preprocessing and post-processing. The experimental results indicate that FMSNet performs favorably against the widely used network architectures and our UIR method can outperform the state-of-the-art methods on both qualitative and quantitative evaluations. Comparing with the original underwater images, the experiments carried out by subsequent mission shows that 285% more feature points can be detected in the restored images by using our method.

Keywords: Neural network · Image restoration · CNN architecture · Image formation model

1   Introduction

Underwater Image restoration (UIR) can significantly improve image quality and the performance of computer vision algorithms in the underwater environment. Developed UIR algorithms can also contribute to the development of marine robotics, marine geology, marine biology, and many marine industries. Although countless UIR algorithms are proposed to achieve the restoration, many of them tend to give rise to partly over enhancement, because of the influence of artificial illumination sources. Thus, Lu et al. [1] consider absorption, scattering, and artificial lighting as three major

distortion issues in the underwater environment. Through experimental observations, we discover that many cases can cause the over-enhancement issue despite no extra artificial luminance used. We attribute the over-enhancement issue to the nonhomogeneous background light, which particularly can be induced by artificial illuminance sources. Nonhomogeneous background light may also appear in other circumstances, such as the underwater landforms where can shelter the light and the wide-angle photographs where the certain area of the water body is brighter than the seabed.

The underwater image formation model (IFM) developed in [2] and [3] is a prevalently used mathematical model demonstrating the underwater imaging process:

$$I_c(x) = J_c(x)T_c(x) + A_c\big(1 - T_c(x)\big) \tag{1}$$

where x is the coordinate of a pixel and $c \in \{R, G, B\}$ is a color channel. $J_c(x)$ and $I_c(x)$ denote the clear object scene radiance and the degraded underwater image. A_c denotes the global ambient light. $T_c(x)$ is the transmission map that represents the residual energy ratio of the scene radiance after the transmission. In this IFM (1), the ambient light A_c is a global value where people assume that the background light is uniform in the whole scene. The neglection of the nonhomogeneous light is one of the reasons that model-based UIR methods tend to suffer from over enhancement.

In addition, most of the deep learning networks are mainly designed by inheriting the classic architectures proposed for the task of image classification or target detection, such as the VGG and the Inception network. There are relatively few CNN architectures particularly designed for end-to-end image transformation. Hence, it is meaningful to design a network architecture that is suitable for the characteristics of image transformation tasks, like image restoration based on supervised learning.

In this paper, we aim to propose a novel CNN network architecture, namely FMSNet, by which we can create a systematic method to recover underwater images with the ability to circumvent the problem of over enhancement and improving the performance of UIR. We revised the traditional underwater IFM by adding an extra item to denote the nonhomogeneous light. And an underwater image synthesizing method is designed to simulate various degraded underwater environments and create an underwater dataset. With the synthesized underwater dataset, we establish a supervised learning framework to train the FMSNet. The FMSNet with frequency-based feature separation and fusion structure is designed based on the characteristics of image restoration tasks and the human visual optimization behaviors. By the convolutional operation with frozen Gaussian convolution kernel, the feature maps can be decomposed into low and high-frequency components and processed efficiently and pertinently.

2   Related Work

Model-based Underwater Image Restoration. Most model-based methods [4-7] follow the scheme of estimating the variables, like ambient light A_c and transmission map $T_c(x)$, then calculate the restored images by the underwater IFM (1). As the estimations from a single degraded image is considered as an ill-post inverse problem,

effective prior information tends to be necessary for these methods. The underwater light attenuation prior (ULAP) [4] is proposed for the estimation of the depth map and A_c. The underwater dark channel prior (UDCP) [5] modifies the Dark channel prior (DCP) [6] by the consideration of the divergency of underwater light attenuation to estimate the $T_c(x)$. Li et al. [7] estimate the variables by minimizing the information loss in the red channel. Model-based methods can markedly improve the contrast and reduce the scattering effect. However, specific priors will inevitably fail in certain cases and lead to incorrect estimation. Moreover, the reliance on the traditional IFM often results in over-enhancement or over compensatory failure for these methods.

Learning-based Underwater Image Restoration. Learning-based methods [8-12] tend to train a neural network to learn the map between the distorted images and its corresponding clear version from a large amount of data. Hence, the main issue to resolve is the requirement of the paired datasets. Fabbri et al. [8] use real underwater images to train a CycleGAN [9], which can generate the degraded version of good quality underwater images. Furthermore, Islam et al. [10] apply a dataset generated by the method of [8] to their proposed FUnIE-GAN network architecture to obtain the recovered results. However, generating datasets using the style-transform ability of GAN is extremely inefficient, because one well-trained GAN can only generate one type of distortion. On the contrary, Li et al. [11] synthesize underwater images directly from the simplified underwater IFM, and the UWCNN network is designed to achieve better restoration performance with end-to-end training. Wang et al. [12] fuse the prior information into the process of the image synthesizing and two CNN frameworks are designed to learn the estimation of the A_c and $T_c(x)$ respectively. Learning-based methods are relatively robust, but the performance of the results is restrained by the quality of datasets and the performance of the CNN architectures.

3 Proposed Method

3.1 Synthesizing Underwater Dataset

Our dataset synthesizing method roughly follows the process of [11], by which we can efficiently simulate underwater data from a clear depth map dataset (RGBD dataset). In our method, the randomly generated variables for IFM can ensure the diversity of synthesized data. To synthesize underwater datasets that take nonhomogeneous light conditions into account, we revise the underwater IFM by adding an extra item to simulate the nonhomogeneous background light. Once the synthesized underwater dataset can simulate the appearance of the nonhomogeneous light, the neural network trained by the dataset will be able to learn the way to eliminate the effect:

$$I_c(x) = J_c(x)T_c(x) + A_c\big(1 - T_c(x)\big) + \eta N(x) \tag{2}$$

where $N(x)$ denotes the nonhomogeneous light component, η is the weight of the $N(x)$. Considering that the nonhomogeneous light should be a smoothed signal with a

gentle changed gradient, we use Perlin noise to simulate it. The Perlin noise is configured as 2-dimension noise with only one noise cycle. We randomly generate 1000 Perlin noises as the set for $N(x)$. Besides, we randomly generate η in the range of (0.2, 0.4) with uniform distribution. In some cases, the nonhomogeneous background light does not exist, so we set 30% of the η as 0. A_c is the ambient light scattered into the sight, which is considered as the homogeneous component of background light with three global values for RGB channels. For the best simulation of A_c, we propose to use the ambient light values from real-world underwater images. Specifically, we use the ULAP [4] to calculate the ambient light in a real-world underwater dataset that can comprehensively cover various water types. Therefore, nearly 1000 groups of ambient light values, A_r, A_g, A_b, are obtained. Before using these values in the synthesizing process, we multiply them by a random jittering generated from a Gaussian distribution. The transmission map $T_c(x)$ is important, which can be expressed as:

$$T_c(x) = e^{-p_\lambda^c d(x)} \tag{3}$$

where p_λ^c is the attenuation coefficient and the $d(x)$ is the distance between the object and the camera. We followed the approach in [7] to generate $T_c(x)$, in which we generated the p_λ^c of the blue channel in the uniform distribution between 1 and 3.

We randomly draw an $N(x)$ and A_c in the set of nonhomogeneous background light and the set of ambient light. Based on equation (2) and the variables mentioned above, we can synthesize the underwater images from any RGBD images. To avoid the overexposure phenomenon brought by the additional $N(x)$, an adaptive rule is applied to the value of η. To be specific, we calculate the proportion of white pixels before and after adding the $\eta N(x)$, and mark them as p_o and p_N. If $p_N - p_o \geq 0.1$, we reset the $\eta = \eta * 0.6$ until the $p_N - p_o < 0.1$ or $\eta < 0.1$. We visually illustrate the images synthesized by our method in Fig. 1. The synthesized images can simulate both the color characteristics and the light conditions of realistic underwater scenes.

Fig. 1. The synthesized underwater images of our method

3.2   Proposed FMSNet CNN Framework

The advanced visual system of humans can easily distinguish objects from distorted scenes. Goffaux et al. [18] found that the effect of a low-pass filtered object on the holistic face perception for humans is significantly larger than the high-pass filtered object. On the other hand, most of the hazy and turbid distortions are appeared in the low-frequency components of an underwater photograph, while most information expected to be preserved, like edges and textures, exists in the high-frequency components. Hence, if we hope a neural network can do the restoration job as well as

humans, we should guide low spatial frequencies and high spatial frequencies to different paths. We propose a novel CNN structure, namely FMSNet, which embodies the frequency-based decomposition and fusion operations for feature maps. The FMSNet consists of three parts: downsampling, backbone, and upsampling.
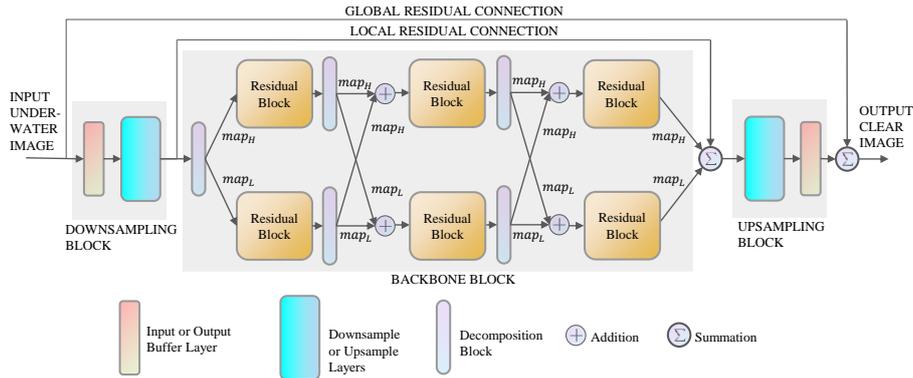

Fig. 2. Architecture of the proposed network

Downsampling and Upsampling Block. The downsampling and upsampling blocks roughly follow the architecture of [14], where there is a convolutional buffer layer with a large receptive field of the inputs for performing the trainable pre-processing and post-processing. In the sampling blocks, we use two stride-2 convolutional layers for downsampling and two stride-1/2 transposed convolutional layers for upsampling of the feature maps. With these blocks, we can save the computational cost and lay the foundation for the subsequent processing. All the convolutional layers use 3*3 kernels, except for the buffer layers who use 9*9 kernels.

Decomposition Block. In the backbone of the FMSNet, we introduce the operation of Gaussian filtering to partition the feature maps into low and high-frequency components. The filtering is performed by a convolutional operation with a pre-set Gaussian kernel. The configuration of the Gaussian kernel is treated as a hyperparameter. In our implementation, we set the kernel size as $15 * 15$ pixels and the gaussian standard deviation as 2.75 for both x and y-directions. The parameters in the kernel are frozen, which will not affect the gradient descent process. With the reflected padding, the filtering will not change the shape of the feature maps. The results of the filtering are the low-frequency components of the feature maps, denoted by map_L. To refrain from the information loss, map_L is subtracted by the original feature maps synchronously to receive the high-frequency feature maps, denoted by map_H. The operation above is merged as a decomposition block in the backbone block.

Backbone Block. Since the output image should share structure with the input for the image restoration task, the residual learning is employed in the backbone block. As

the certified performance, we adopt the residual block designed in [15] for feature processing, where there are two convolutional layers with residual connection. All the convolutional layers use $3 * 3$ kernels. We adjust the batch normalization to instance normalization which can significantly improve the quality of feedforward. The body of FMSNet thus consists of the alternate decomposition blocks and residual blocks. The map_L and map_H from decomposition blocks will be processed by different residual blocks and crosswise fused to reduce the number of maps. The output of the last pair of residual blocks are summed and propagated to the upsampling block.

In addition, due to the effectiveness of the multi-scale residual learning presented in [16], we design two residual connection for the feature maps of different levels to enforce the network to learn the residual information between the input side and output side. The exact architecture of our FMSNet network is shown in Fig. 2.

4 Experimental Evaluation and Discussion

4.1 Datasets and Training Strategy

Training and Testing Dataset. Synthesizing realistic underwater images to compose a training and testing dataset with high generalization capability is a crucial step. We use the RGBD dataset from [17] with 600 outdoor in-street scenes to synthesize the underwater dataset. On the other hand, the network will learn the explicit/implicit statistical relationship between the degraded images and their ground truth [18]. If there are too many scenes with different features in the ground-truth targets, the network may confuse about the features of its output, which is not conducive to the one-way training of the network. Hence, we only select 200 images that have consistent features of high contrast and appropriate luminance among the RGBD dataset. We synthesize 8 underwater images with different parameters for each RGBD image. Thus, there are only 1600 underwater images in our training and testing dataset.

Training Strategy. We use the summation of L1 loss and the weighted multi-scale structural similarity (MS-SSIM) loss in [19] as the loss functions to achieve supervised learning for the FMSNet. We randomly set 80% of the synthesized data as the training set, and the others as the testing set. For the augmentation of the training set, we crop a patch of $400 * 400$ pixels at a random position of the training samples, and we resize the images who smaller than the size to $400 * 400$. Moreover, we set three types of transforms to augment the training set: horizontal flipping, vertical flipping, and rotate 90 degree. Each of them may happen in a probability of 30% on the training set.

The network is initialized by the normal distribution. We set the batch size as 10 and train the network for 100 epochs. The Adam optimizer with β_1 and β_2 of default values is used to accelerate the training. The learning rate is set as 0.002 initially, and we decrease it by half when the epoch number reaches the milestones of 25, 45, 60, 70, 80, 90. The process of the training is performed on a PC with two NVIDIA GeForce GTX 1080Ti GPU and Xeon(R) Gold 5218 CPU.

4.2 Network Performance

To demonstrate the performance of the FMSNet, we reimplement the widely used network architectures, the ResNet [14] and the GridDehazeNet [18], as baselines. All the models are trained by our synthesized training set for 100 epochs. Since the loss function is the summation of L1 loss and MS-SSIM loss, we compare the PSNR and SSIM metrics on both the training set and the testing set to show the loss function minimizing process of the three networks. Fig. 3(left) depicts the PSNR curves, where the PSNR of both the training set and testing set for FMSNet are significantly higher than the other models all over the epochs. In Fig. 3(right), the GridDehazeNet achieves highest SSIM curve because of its multi-scale structure. The SSIM curves of the other networks are extremely similar, where FMSNet's SSIM curves are slightly higher. The experimental results demonstrate that the proposed FMSNet performs favorably against the widely used baseline networks for end-to-end training. With the help of the decomposition and fusion structure, the FMSNet is particularly sensitive to the pixel-level loss function, like L1, which leads to the rapid convergence and the best PSNR of 26.8 dB on the training set and 26.2dB on the testing set.
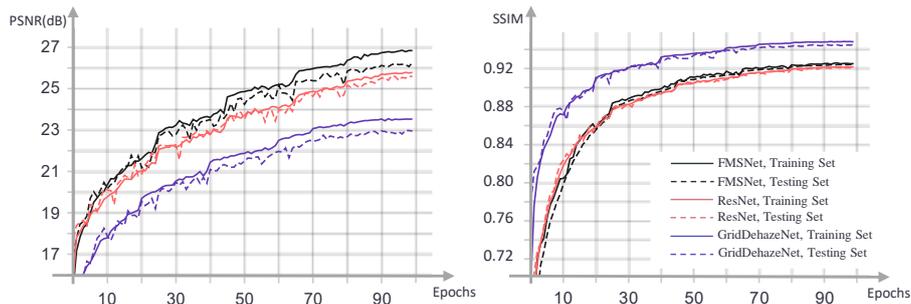

**Fig. 3.** The PSNR and SSIM curves trained by our network and the baseline networks

4.3 Evaluation on Synthesized Underwater Images

The FMSNet trained by our synthesized underwater dataset can directly recover any degraded underwater images. By which, we can conveniently construct a UIR method without the need of any pre-processing or post-processing. For the evaluation of the restoring performance, we synthesize extra underwater images from the OTS dataset in RESIDE [20], which is never observed by our model. We use the synthesized underwater images to evaluate the restoring results of our method and several state-of-the-art UIR methods, including ULAP in [4], UDCP in [5], the method in [7], FUnIE-GAN in [10], and UWCNN in [11]. We compare the restoring results in Fig. 4.

According to Fig. 4, although the clarity can be improved, the results of [7] and UDCP prone to over enhancement the light and overcompensate the color. Similarly, the first image of column (e) can be restored by ULAP, but it is nearly invalid for the other images. (f) and (g) are from learning-based methods, thus their performances in different scenarios are relatively stable. The contrast ratio of every result from

UWCNN gets increased, while the original color saturation is completely dropped. The FUnIE-GAN's restoration can correct the color significantly, nevertheless, the hazy effect cannot be eliminated completely. In contrast, the results of our method in (h) can produce the best clarity and free of the major turbid distortion.

We treat the images before appending the underwater effect as the reference to evaluate the restoration ability of the methods. The Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM), and Color-Difference Formula CIEDE2000 are used as the quantitate metrics. Table 1 shows the comparison of the metrics and the best numbers are bolded. Our method performs the best result for all the metrics.
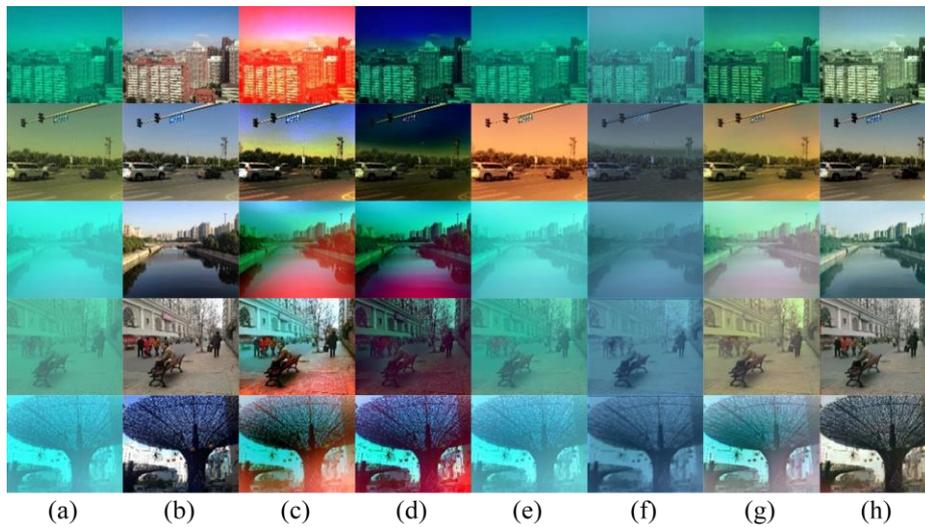

Fig. 4. Comparison on synthesized images: (a) the synthesized images; (b) the corresponding reference images; (c) the method of [7]; (d) the method by UDCP; (e) the method by ULAP; (f) the method by UWCNN; (g) the method by FUnIE-GAN; (h) our proposed method

Table 1. Comparison of the methods on quantitate metrics

Method	Method [7]	UDCP	ULAP	FUnIE-GAN	UWCNN	Our method
PSNR	14.44	12.24	11.01	14.99	12.87	**18.70**
SSIM	0.75	0.69	0.66	0.78	0.58	**0.87**
CIEDE 2000	20.63	23.17	28.39	20.80	22.84	**14.14**

4.4 Evaluation on Real-World Underwater Images

Evaluating the algorithms on real-world underwater images is the hinge to validate their performance and generalization ability. We collect 80 various real underwater images carrying serious distortion. 60 of them are from the challenging dataset of [13] and the rest are collected from the internet. Fig. 5 presents the visual comparison of the restoration results on the real-world underwater images.
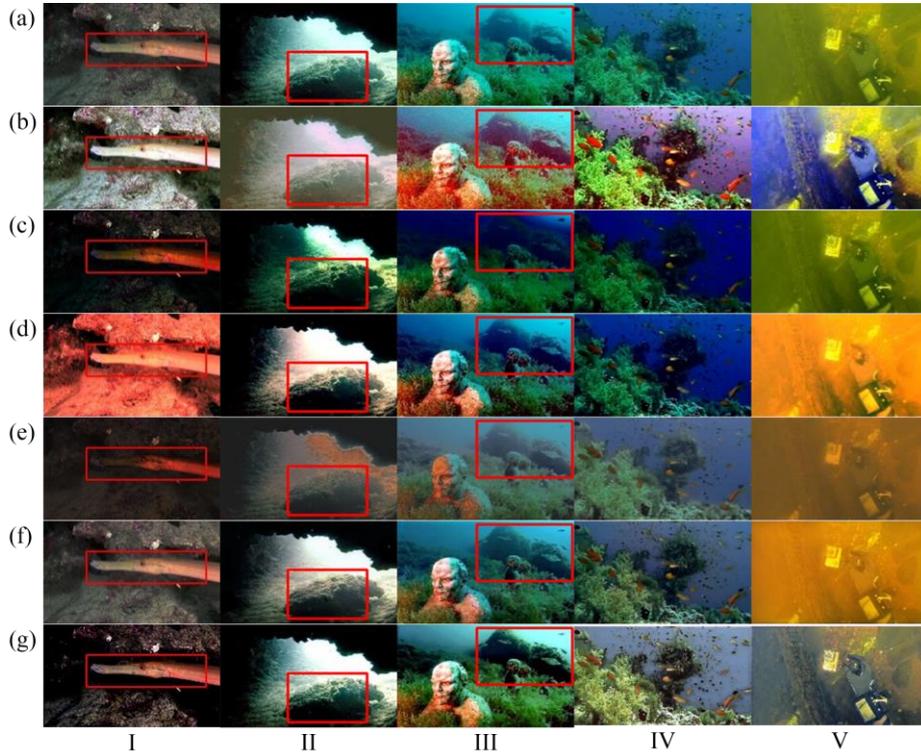
Fig. 5. Evaluation on the real-word images: (a) the real-world underwater images; (b) the method of [7]; (c) the method by UDCP; (d) the method by ULAP; (e) the method by UWCNN; (f) the method by FUnIE-GAN; (g) our proposed method

According to the images in the row (b), the results of [7] can balance the color distribution and the contrast ratio to a large extent. But the manifest over enhancement and oversaturation are also produced in images I, II, and III because of the ignoration of the nonhomogeneous light in their model. The prior of UDCP and ULAP in the row of (c) and (d) succeed in removing the foggy noise utterly but tend to drastically change the original color distribution and darken the brightness. The loss of their effectiveness is due to the fact that their specific priors can hardly adapt to these lighting conditions. Comparing with UDCP, the method by UWCNN in (e) can maximumly eliminate the color distortion, but the beneficial color information is destroyed too. Owing to the using of a real underwater dataset in the method of FUnIE-GAN, the results in (f) look limpid and natural, whereas the degree of removing scattering effect is still not enough. The reason is maybe relative to the architecture of their neural network. By contrast, our method can enhance the best visibility and the relatively pleasing color distribution. Thanks to the consideration of the nonhomogeneous light condition in our method, we have fewer over-enhancement areas. By the comparison of the marked region of I, II, and III, we can figure out much more haze-free details and actual information in (g), which is contributed by the FMSNet architecture. For

the quantitate evaluation of the results, we use two non-reference metrics, Underwater Image Quality Metric (UIQM) and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), on the 80 restoration results. Smaller BRISQUE means better naturalness and larger UIQM means better comprehensive quality. In table 2, we list the average scores of the different methods and bold the best scores. Our method gets the best score for UIQM and the second-best score for BRISQUE.

Table 2. Comparison on real-world testing set with non-reference metrics and feature points matching

Algorithms	Non-reference Metrics		Matched Feature Points Number				
	UIQM	BRISQUE	Image I	Image II	Image III	AVG	INC (%)
Original	1.24	39.17	64	134	32	76	0%
Method [7]	2.76	34.00	412	226	83	240	213.71%
UDCP	2.15	40.02	187	99	60	115	50.52%
ULAP	1.83	38.17	217	180	53	150	95.82%
UWCNN	2.11	43.78	91	80	24	65	-15.14%
FUnIE-GAN	2.16	**23.59**	209	184	64	152	98.82%
Our method	**3.13**	23.78	**489**	**272**	**119**	**293**	**285.52%**
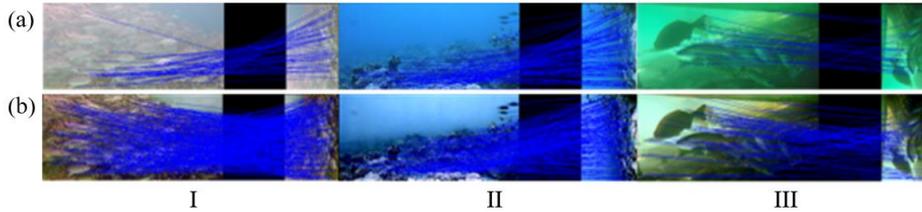
Fig. 6. Experiment of SIFT feature points matching: (a) is for the original underwater images; (b) is for the restoration results of our method

4.5 Evaluation by The Subsequent Application

In this experiment, we indirectly measure and compare the quality of the restorations by observing that how much the restoring results can improve the performance of the subsequent algorithm. The feature point detecting and matching by SIFT algorithm [21] is implemented on the restoration results of section 4.4. We choose three typical underwater scenes with fish groups and rugged seabed as testing objections (I, II, III). These images are rotated and zoomed in to format the paired targets for feature points matching. Fig 6. shows the matching lines on the original underwater images and the restored images of our method. The more feature points that can be detected and matched means the more information is reserved in the image. Therefore, we list the matched number of the feature points of each image in Table 2. The AVG is the average number for the three examples and the INC shows the improvement comparing with the original underwater images. From Table 2, our method gets the highest number of feature points for all the examples, which shows that the results of our method

can grab more detailed information. After the restoration by our method, the average number of the matched feature points increased by 285.52%.

5   Conclusion

In this paper, we presented an underwater image restoration method based on an end-to-end CNN model and a synthesized underwater dataset. The proposed CNN architecture, FMSNet, showed promising performance on image restoration tasks. Comparing with the baseline models, the FMSNet can efficiently perform supervised learning. We also gave an explanation about the over enhancement caused by artificial luminance for the existed methods. To circumvent the effect of nonhomogeneous background light, we revised the IFM and designed a new underwater image synthesizing method, by which we can create realistic underwater datasets for network training. The experimental evaluations indicate that our restoration method can significantly improve the quality of underwater images and refrain from the effect of over enhancement. The experiment by SIFT algorithm shows that our restoration method can increase the matched feature points number by 285%.

In the FMSNet, we designed a decomposition block to separate the feature maps into two components by fixed Gaussian filtering, which sheds some light on future works about more appropriate ways to modify the decomposition block, such as introducing the multi-scale Gaussian filtering or random-scale Gaussian filtering.

References

1. Lu, H., Li, Y., Xu, X., Li, J., Liu, Z., Li, X., Yang, J., Serikawa, S.: Underwater image enhancement method using weighted guided trigonometric filtering and artificial light correction. Journal of Visual Communication and Image Representation. 38, 504–516 (2016). https://doi.org/10.1016/j.jvcir.2016.03.029
2. Chiang, J.Y., Ying-Ching Chen: Underwater Image Enhancement by Wavelength Compensation and Dehazing. IEEE Trans. on Image Process. 21, 1756–1769 (2012). https://doi.org/10.1109/TIP.2011.2179666
3. Schechner, Y.Y., Karpel, N.: Recovery of Underwater Visibility and Structure by Polarization Analysis. IEEE J. Oceanic Eng. 30, 570–587 (2005). https://doi.org/10.1109/JOE.2005.850871
4. Song, W., Wang, Y., Huang, D., Tjondronegoro, D.: A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration. In: Hong, R., Cheng, W.-H., Yamasaki, T., Wang, M., and Ngo, C.-W. (eds.) Advances in Multimedia Information Processing – PCM 2018. pp. 678–688. Springer International Publishing, Cham (2018)
5. Jr, P.D., Nascimento, E. do, Moraes, F., Botelho, S., Campos, M.: Transmission Estimation in Underwater Single Images. In: 2013 IEEE International Conference on Computer Vision Workshops. pp. 825–830 (2013)

6. Kaiming He, Jian Sun, Xiaoou Tang: Single Image Haze Removal Using Dark Channel Prior. IEEE Trans. Pattern Anal. Mach. Intell. 33, 2341–2353 (2011). https://doi.org/10.1109/TPAMI.2010.168

7. Li, C.-Y., Guo, J.-C., Cong, R.-M., Pang, Y.-W., Wang, B.: Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior. IEEE Trans. on Image Process. 25, 5664–5677 (2016). https://doi.org/10.1109/TIP.2016.2612882

8. Fabbri, C., Islam, M.J., Sattar, J.: Enhancing Underwater Imagery Using Generative Adversarial Networks. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 7159–7165. IEEE, Brisbane, QLD (2018)

9. Zhu, J.-Y., Park, T., Isola, P., Efros, A.A.: Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2242–2251 (2017)

10. Islam, M.J., Xia, Y., Sattar, J.: Fast Underwater Image Enhancement for Improved Visual Perception. IEEE Robotics and Automation Letters. 5, 3227–3234 (2020). https://doi.org/10.1109/LRA.2020.2974710

11. Li, C., Anwar, S., Porikli, F.: Underwater scene prior inspired deep underwater image and video enhancement. Pattern Recognition. 98, 107038 (2020). https://doi.org/10.1016/j.patcog.2019.107038

12. Wang, K., Hu, Y., Chen, J., Wu, X., Zhao, X., Li, Y.: Underwater Image Restoration Based on a Parallel Convolutional Neural Network. Remote Sensing. 11, 1591 (2019). https://doi.org/10.3390/rs11131591

13. Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D.: An Underwater Image Enhancement Benchmark Dataset and Beyond. arXiv:1901.05495 [cs]. (2019)

14. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In: Leibe, B., Matas, J., Sebe, N., and Welling, M. (eds.) Computer Vision – ECCV 2016. pp. 694–711. Springer International Publishing, Cham (2016)

15. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016)

16. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual Dense Network for Image Super-Resolution. arXiv:1802.08797 [cs]. (2018)

17. Li, R., Cheong, L.-F., Tan, R.T.: Heavy Rain Image Restoration: Integrating Physics Model and Conditional Adversarial Learning. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1633–1642. IEEE, Long Beach, CA, USA (2019)

18. Liu, X., Ma, Y., Shi, Z., Chen, J.: GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 7313–7322. IEEE, Seoul, Korea (South) (2019)

19. Zhao, H., Gallo, O., Frosio, I., Kautz, J.: Loss Functions for Neural Networks for Image Processing. arXiv:1511.08861 [cs]. (2018)

20. Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z.: Benchmarking Single-Image Dehazing and Beyond. IEEE Transactions on Image Processing. 28, 492–505 (2019). https://doi.org/10.1109/TIP.2018.2867951

21. Dai, C., Lin, M., Wang, Z., Zhang, D., Guan, Z.: Color Compensation Based on Bright Channel and Fusion for Underwater Image Enhancement. Acta Optica Sinica. 38, 1110003 (2018)