Underwater Image Enhancement Using Improved Shallow-UWnet

Toyoki Yasukawa(✉), Keisuke Hamamoto, Yuchao Zheng, and Huimin Lu

Kyusyu Institute of Technology, Kitakyushu, Japan
yasukawa.toyoki296@mail.kyutech.jp

Abstract. Currently, with the development of industrialization, ocean exploration is actively carried out to investigate energy resources and undersea ecosystems. Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs), which are fully automated and not operated by humans, are used for these surveys. However, it is difficult to obtain clear images underwater due to low contrast and blur caused by the unique optical characteristics of the underwater environment. In recent years, many methods have been proposed for underwater image enhancement, with the most common methods using deep learning, such as generative adversarial networks (GANs) and convolutional neural networks (CNNs). Most of these methods are computationally and memory intensive, making real-time underwater image correction difficult. The Shallow-UWnet method is developed to solve this problem, and it enables a significant reduction in computational complexity compared to conventional methods.

In this study, we improve Shallow-UWnet using Deformable Convolution, propose a new method with higher accuracy, compare its accuracy with that of the previous method, and verify its usefulness.

Keywords: Underwater Image Enhancement · Deformable Convolution · Deep Learning

1 Introduction

1.1 Background

The oceans are home to many living organisms and resources that play an important role in sustaining life on Earth. In recent years, the consumption of oil, coal and minor metals, which are energy resources, has increased year by year due to industrialization. In particular, rare metals are essential materials for the production of lithium-ion batteries, semiconductors, motors, and other components that are indispensable in industries [1, 2] such as automobiles, AI-equipped devices, and IoT devices that are powered by electricity. However, Japan imports almost all of its rare metals, making their supply unstable. To solve this problem, Japan is actively exploring and developing mineral resources within its exclusive economic zone (EEZ). In addition, since the mid-20th century, not only Japan but also the world has been actively engaged in ocean exploration

H. Lu and J. Cai (Eds.): ISAIR 2023, CCIS 1998, pp. 175–180, 2024.
https://doi.org/10.1007/978-981-99-9109-9_17

using high technology to develop energy resources. Among these, underwater robots that can explore even in harsh environments inaccessible to humans have attracted much attention. Currently, underwater robots are playing an important role in a variety of fields, including marine geological exploration, resource development, ecosystem research, and fisheries.

Several papers have been published on underwater image enhancement and restoration, and Yan [3] et al. The two types are described below.

- IFM-based: The IFM-based method analyzes the image formation and light propagation in water, constructs an effective underwater degradation model, infers the parameters of the physical model, and finally reconstructs the image through a compensation process.
- IFM-free: IFM-free corrects images mainly based on pixel intensity redistribution, so it does not need to consider underwater-specific optical properties as IFM-based methods do.

Recently, it has been used for image enhancement based on the idea of learning CNN hidden features to improve image quality.

2 Method

2.1 Shallow-UWnet [4]

Shallow-UWnet is an IFM-free CNN-based model with significantly lower computational complexity and memory requirements than conventional CNN-based models and adversarial generative networks. The architecture of Shallow-UWnet is a convolutional network consisting of three densely connected ConvBlocks connected in series, as shown in Fig. 1. The input image is connected to the output of each ConvBlock by skip connections.

Although the shallow-UWnet is a lightweight model, we believe it can improve the quality of image correction.

2.2 Proposed Method

In this research, the proposed method is to change the second convolution layer in each ConvBlock of the Shallow-UWnet to Deformable Convolution, as shown in Fig. 2. The reason for this is that Deformable Convolution can extract features more efficiently than the convolution layer, and thus can provide more accurate image correction than conventional methods.

Deformable Convolution [5]: Convolutional layers are not always able to extract features properly because the shape and scale are fixed at the network design stage, and the scale and shape of objects in an image differ from each other. There are obvious problems in dealing with object viewpoints, distortions, etc. in visual recognition tasks, especially for non-rigid objects. Moreover, convolution has the disadvantage of processing fixed positions, which makes it vulnerable to geometric deformations. Deformable Convolution was introduced to solve this problem. It adds a two-dimensional offset to

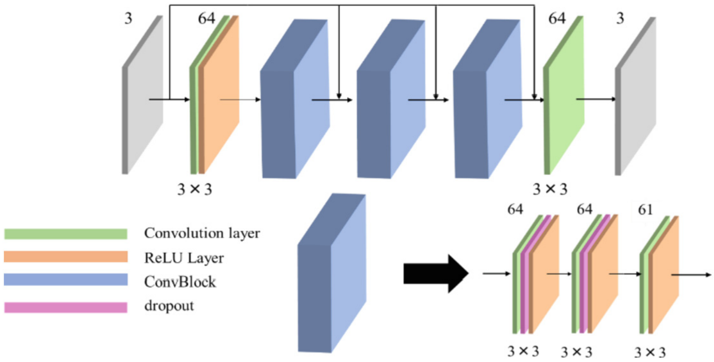

Fig. 1. Shallow-UWnet network architecture

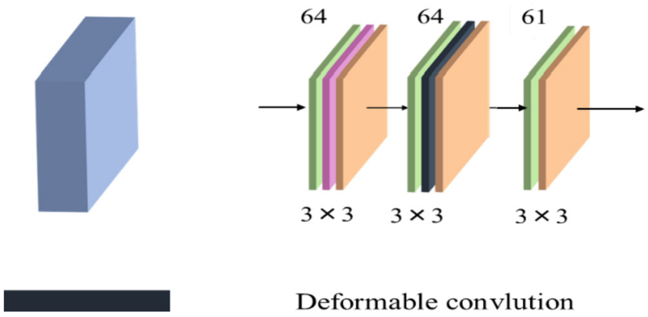

Fig. 2. Image of the proposed method

the sampling position of the regular grid in the standard convolution. This allows free deformation of the sampling grid. The offset is learned from the previous feature map via an additional convolution layer. In this way, deformations can be applied to the input features in a localized, dense, and adaptive manner.

Deformable Convolution consists of two steps: (1) Sampling using a regular grid R for the input feature x (2) Summing the sampled values weighted by ω.

$$R = \{(-1, -1), (-1, 0), \ldots, (0, 1), (1, 1)\} \tag{1}$$

defines a kernel of 3×3 with a dilation 1. For each position p_0 on the output feature map y

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n) \tag{2}$$

The following is an example of the grid R. where p_n enumerates the positions of the grid R. The regular grid R is extended by the offset $\{ \Delta p_0 | n = 1, \ldots, N\}$ where $N = |R|$.

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n + \Delta p_0) \tag{3}$$

The sampling position is undefined and is on the offset position $p_n + \Delta p_0$. Since the offset Δp_0 is usually not an integer, Eq. (3) can be expressed by bilinear interpolation as

$$x(p) = \sum_q G(q, p) \cdot x(q) \tag{4}$$

where p is any fractional position ($p = p_0 + p_n + \Delta p_n$ in Eq. (3)), q is all integer spatial positions of the feature map x, and $G(\cdot, \cdot)$ is the bilinear interpolation kernel.

2.3 Training

We used 1500 images of UFO-120 [6] as training data. For the test data, we used 120 images from the UFO-120 dataset and 890 images from the UIEB dataset [7]. We also use PSNR, SSIM, and UIQM [8] as our evaluation metrics.

3 Results and Discussion

Experimental environments are Intel® core™ i9-10900X and NVIDIA GeForce RTX3090. Table 1 shows a comparison of the metrics between the original model and the proposed method. The comparison image is also shown in Fig. 3.

Table 1. Comparison of evaluation metrics

	UFO-120			UIEB		
	PSNR	SSIM	UIQM	PSNR	SSIM	UIQM
original	26.211	0.777	2.806	19.319	0.720	**2.785**
proposed	**26.602**	**0.785**	**2.810**	**19.382**	**0.721**	**2.764**

Figure 1 shows that the proposed method has a higher evaluation index for both UFO-120 and UEIB. Furthermore, from the image in Fig. 3, the proposed method removes slightly more blue and green colors, which is closer to the correct image. However, there are still some green and blue colors in the image, and we believe this is because Shallow-UWnet and the proposed method are not IFM-based methods. The IFM-based method, which considers an optical model, can correct images by taking into account the distance decay of light, so if the distance of the subject from the camera is known, the original color can be approximately identified, and thus more appropriate image correction can be performed. However, it has the disadvantage that it requires the use of images that include depth data in the dataset [9–11], which limits the selection of underwater image datasets, which are already limited.

Fig. 3. Comparison of original and proposed underwater image enhancement Input, Original, Proposal, Grand Truth

4 Conclusion

In this paper, we propose an improved method of Shallow-UWnet with Deformable Convolution and compare its accuracy using two datasets UFO-120, UIEB. As a result, we confirmed that the PSNR, SSIM, and UIQM indices increased for almost all datasets. This indicates that offset learning, in which the kernel of the Deformable Convolution is dynamically transformed to match the shape and scale of the object, is enhancing the underwater image correction.

Currently, underwater images are used for processing [12–14]. Future work is to perform real-time processing in an actual underwater environment to compare and verify the results [15, 16].

References

1. Lu, H., Tang, Y., Sun, Y.: DRRS-BC: decentralized routing registration system based on Blockchain. IEEE/CAA J. Autom. Sinica **8**(12), 1868–1876 (2021). https://doi.org/10.1109/JAS.2021.1004204
2. Lu, H., Zhang, M., Xing, Xu., et al.: Deep fuzzy hashing network for efficient image retrieval. IEEE Trans. Fuzzy Syst. **29**(1), 166–176 (2021)
3. Yan, W., Wei, S., Giancarlo, F., et al.: Antonio liotta an experimental-based review of image enhancement and image restoration methods for underwater imaging. IEEE Access 140233–140251
4. Naik, A., Swarnakar, K.: Shallow-UWnet: Compressed models for underwater image enhancement
5. Jifeng, D., Haozhi, Q., Yuwen, X., et al.: Deformable convolutional networks. In: IEEE International Conference on Computer Vision (ICCV), pp. 764–773 (2017)
6. The UFO-120 https://irvlab.cs.umn.edu/resources/ufo-120-dataset. Last accessed 26 Apr 2023
7. UIEB dataset https://li-chongyi.github.io/proj_benchmark.html. Last accessed 26 Apr 2023
8. Panetta, K., Gao, C., Agaian, S.: Human-visual-system-inspired underwater image quality measures. IEEE J. Oceanic Eng. **41**(3), 541–551 (2016)

9. Lu, H., Li, Y., Chen, M., et al.: Brain Intelligence: go beyond artificial intelligence. Mobile Netw. Appl. **23**, 368–375 (2018)
10. Lu, H., Yang, R., Deng, Z., et al.: Chinese image captioning via fuzzy attention-based DenseNet-BiLSTM. ACM Trans. Multimedia Comput. Commun. Appl. **17**(1s), 1–18 (2021). https://doi.org/10.1145/3422668
11. Lu, H., Qin, M., Zhang, F., et al.: RSCNN: A CNN-based method to enhance low-light remote-sensing images. Remote Sens. **13**(1), 62 (2020). https://doi.org/10.3390/rs13010062
12. Lu, H., Wang, D., Li, Y., et al.: CONet: a cognitive ocean network. IEEE Wirel. Commun. **26**(3), 90–96 (2019)
13. Lu, H., Zhang, Y., Li, Y., et al.: User-oriented virtual mobile network resource management for vehicle communications. IEEE Trans. Intell. Transport. Syst. **22**(6), 3521–3532 (2021)
14. Lu, H., Li, Y., Nakashima, S., et al.: Underwater image super-resolution by descattering and fusion. IEEE Access **5**, 670–679 (2017)
15. Lu, H., Li, Y., Mu, S., et al.: Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. IEEE Internet Things J. **5**(4), 2315–2322 (2018)
16. Xu, X., Lu, H., Song, J., et al.: Ternary adversarial networks with self-supervision for zero-shot cross-modal retrieval. IEEE Trans. Cybern. **50**(6), 2400–2413 (2020)