Underwater Light Field Retention : Neural Rendering for Underwater Imaging

Tian Ye†, Sixiang Chen†, Yun Liu‡, Yi Ye†, Erkang Chen†, Yuche Li‡
†School of Ocean Information Engineering, Jimei University, Xiamen, China
‡College of Artificial Intelligence, Southwest University, Chonqing, China
§College of Geosciences, China University of Petroleum, Beijing, China
{201921114031, 201921114013} @jmu.edu.cn
yunliu@swu.edu.cn, 201921114003@jmu.edu.cn, ekchen@jmu.edu.cn, liyuche_cq@163.com

Abstract

Underwater Image Rendering aims to generate a true-to-life underwater image from a given clean one, which could be applied to various practical applications such as underwater image enhancement, camera filter, and virtual gaming. We explore two less-touched but challenging problems in underwater image rendering, namely, i) how to render diverse underwater scenes by a single neural network? ii) how to adaptively learn the underwater light fields from natural exemplars, i.e., realistic underwater images? To this end, we propose a neural rendering method for underwater imaging, dubbed UWNR (Underwater Neural Rendering). Specifically, UWNR is a data-driven neural network that implicitly learns the natural degenerated model from authentic underwater images, avoiding introducing erroneous biases by hand-craft imaging models.

Compared with existing underwater image generation methods, UWNR utilizes the natural light field to simulate the main characteristics of the underwater scene. Thus, it is able to synthesize a wide variety of underwater images from one clean image with various realistic underwater images.

Extensive experiments demonstrate that our approach achieves better visual effects and quantitative metrics over previous methods. Moreover, we adopt UWNR to build an open Large Neural Rendering Underwater Dataset containing various types of water quality, dubbed LNRUD. The source code and LNRUD are available at https://github.com/Ephemeral182/UWNR.

Fig. 1. Visual comparison of our method and previous underwater rendering methods. (a) and (c) are respectively obtained by UWGAN [45] and UISA [18]. (b) and (d) are rendered by our method. The underwater rendering results by our method have more natural visual effects closing to real underwater scenes.

1. Introduction

The underwater environment and resource has been gradually developed and explored in the last decade [6, 12, 14, 21, 24]. Unlike the terrestrial environment, it is hard for us to obtain satisfactory paired underwater datasets, limiting the development of many underwater vision tasks. For example, underwater image enhancement is a practical computer vision task that requires paired underwater images with ideal ground-truth for training. However, manual collection often requires a lot of workforce and material resources [19, 20, 28]. Furthermore, the visual effect of the ground-truth generated by various algorithms in current underwater image datasets is not satisfactory for implementation. And the synthetic underwater datasets based on the traditional underwater imaging model hardly model the complex degeneration of the underwater environment well. In particular, underwater image rendering that effectively generates true-to-life underwater images becomes a novel and valuable technique compared to the high cost of acquiring...
paired underwater datasets with real underwater images.

Previous underwater image restoration works [3, 11] usually utilize the following formula as the underwater imaging model:

\[ I(x) = J(x)t(x) + B(x)(1 - t(x)). \]  

where \( B(x) \) is the background light and \( J(x) \) is the image without degrading by underwater particle scattering. \( t(x) = e^{-\beta d(x)} \) is the transmission map, \( \beta \) is the scattering coefficient and \( d(x) \) is the depth of scene. The \( I(x) \) is the underwater image.

Image generation like image rendering etc. usually involves adversarial generative networks (GAN) [13, 22, 23, 31, 47]. Zhu et al. [47] utilized a cycle-consistent GAN to implement image style transfer. Proposing mapping networks and adaptive Instance Normalization techniques were proposed in a large generative network Stylegan [22, 23]. Li et al. [31] presented Weather GAN to render different weather scenes and migrated weather conditions from one category to another. The above GAN-based approaches generate impressive results. However, there are well-known shortcomings for GAN methods. For example, GAN-based methods are prone to mode collapse and create fake features.

For previous underwater image generation methods, commonly focus on physical model-based and GAN generating techniques [18, 45–47]. It can be observed from Eq.(1) that generating an underwater image requires estimating the ambient light \( B(x) \) and the underwater transmission map \( t(x) \). Moreover, physical model-based methods for underwater image generation still have the following limitations: (1) A simple physical model cannot cover all complex underwater scenes, especially the current physical model-based approach is biased by limited estimation of key parameters. (2) Previous methods most ignore the effect of the scattering coefficient in underwater imaging, because of the difficulty of capturing this coefficient accurately, which leads to unfavorable transmission maps can cause the visual effect of generated underwater images to be unsatisfactory. (3) The underwater transmission map \( t(x) \) of the physical model is a coefficient map with depth variation characteristics. Obtaining RGB-D images is relatively expensive and limited in resources.

To address these problems, we propose an underwater rendering framework to achieve the efficient rendering generation of true-to-life underwater images. The mechanism of Light Field Retention (LFR) in our framework can effectively transfer the diverse underwater style from natural underwater images to the objective generated images, which effectively guarantees the diversity of rendering results. It’s worth to note that there is no physical model constraints and GAN-based method in our framework so that it can easily avoid following problems: (1) **Mode Collapse.** [45] (2) **Fake Features.** [45] (3) **Limited performance by incomplete underwater imaging physical model.** [18, 33] And it is also worth mentioning that we utilize the well trained depth estimation network for collaborative work, which significantly alleviates the obstruction of our method in practical applications by avoiding the high cost of RGB-D images as the input of network [29, 45]. Specifically, our method only needs a clean image and its depth map estimated by Li et al. [32]’s way to generate a great quantity of underwater images from real world terrestrial image. To the best of our knowledge, this is the first method that gets rid of the physical model and GAN-based methods to generate underwater images. We summarize our novelties and contributions as follows:

- We develop a natural light field retention module that renders the characteristics of the terrestrial image as close to the underwater situation as possible with the help of underwater dark channel loss and light field consistency loss we proposed.
- To the best of our knowledge, this is the first work to render underwater image without physical models and GAN methods, which can easily render realistic underwater scenes with diverse styles.
- We conduct extensive experiments to demonstrate that our method achieves state-of-the-art rendering performance in terms of objective evaluation.
- We synthesize a large neural rendering underwater datasets (LNRUD), which contains a large number of underwater images synthesized from land images by our method.

2. Related Work

**Image Generation.** Generative adversarial networks were first proposed by Goodfellow et al. [13]. In the field of image generation, many designs with adversarial theory as the base have produced unforgettable impression. CycleGAN [47] utilized cycle consistency principle to enhance GAN for unsupervised image style transfer, which also was extended to some other image generation methods and achieved amazing results [7, 39, 41]. Motivated by the high quality of StyleGAN [22, 23], which performs exceptionally well on image editing and processing tasks. Based on the StyleGAN, Elad et al. [12] presented an end-to-end training method for learning a mapping from images to StyleGAN latent codes to generate high-quality images. Chong et al. [9] allowed StyleGAN to directly start various tasks through pre-training and a little operation on the latent space. In addition to these, there are many interesting image generation tasks using the GAN-based method [8, 31, 35, 42]. For example, Li et al. [31] developed Weather GAN to realize the
transformation of different seasons and weather, Sarkar et al. [42] designed HumanGAN which was the first way to address every aspect human image generation, e.g. global appearance sampling, pose transfer.

There are also inevitable problems behind the excellent effect of generative adversarial networks. The first is that the GAN architecture is usually difficult to train and it is not easy to objectively measure the generation effect [2, 4, 5, 15, 36]. Another difficulty is that GAN will experience mode collapse due to improper punishment during training [30, 34, 44].

Underwater Image Generation. Deep learning has proven to perform well in various underwater tasks, but it is difficult to obtain large datasets in deep-sea environments. How to generate underwater images from existing resources is an important and challenging task [3, 10, 18, 29, 45, 46]. In [3, 10], authors constructed a large number of synthetic datasets utilizing previously calculated ocean attenuation coefficients combined with underwater attenuation models. Hou et al. [18] developed the quadtree to select the background light area and obtained the transmission map according to the DCP principle [16]. Different types of underwater degradation images were designed to imitate different underwater scenes. Li et al. [29] proposed WaterGAN, which takes in-air images, depth maps, and noise vectors as input and outputs synthetic images using a camera model and a generative adversarial network. On this basis, Wang et al. [45] presented UWGAN which only input aerial images and depth maps into the model, adopting the imaging model [1] and GAN to design a simple underwater image generation model.

The previous underwater image generation methods are based on the physical model as the vital core to render underwater images. Still, the biased estimation characteristics of the physical model make it difficult to guarantee the underwater image effect. The GAN-based approaches will lead to train unstably and prone to mode collapse. Our framework retains light field information to avoid the above problems and achieve better performance.

3. Proposed Method

This section presents our underwater image rendering framework with natural light field retention. First, we discuss our natural light field retention module. Then we elaborate on the underwater image generation module of our method. Finally, we present the loss functions we used in training. The overview of our underwater rendering method is shown in Fig. 3.

3.1. Natural Light Field Retention

In theory, the transferring procedure consists of the estimation of three key parameters (i.e., \( d(x), \beta \) and \( B(x) \)) based on the underwater imaging model. Due to particle scattering, the light characteristics of underwater scenes are extremely different from terrestrial images because they have strong randomness, which causes difficulty for the traditional physical method to model these accurately.

To address the above problem, we propose the light field retention scheme to transfer the diverse underwater light field information to the objective image possibly. For an underwater image \( x_u \) in dataset \( X_u \{x_u, y_u\} \), according to Retinex [25, 26] theory, the image can be disassembled into

\[
x_u = x_l \cdot x_r.
\]

where \( x_l \) represents the illumination component of ambient light, \( x_r \) is the reflection component of the target object carrying the image detail information. We perform a multiscale Gaussian low-pass filter on the \( x_u \) to obtain underwater light field map:

\[
x_g = \frac{1}{3} \sum_{\sigma} Gauss_\sigma(x_u), \sigma \in \{15, 60, 90\}.
\]

similar to MSR [40], we empirically set \( \sigma \) to 15, 60, 90, respectively. The appropriate convolution kernel is selected by \( \sigma \) adaptive control in the formula. Considering that Gaussian filter may still contain object details, we transform it to logarithmic domain and scale it to get the final underwater light field map:

\[
x_l = \text{Normalization}(\log x_g).
\]
Fig. 3. Underwater image rendering framework we proposed. In the training phase, with the pair of real underwater image and its clean ground-truth, the light field map and the scene depth are estimated using a natural light field retention module and a pre-trained depth estimation module respectfully. Then a MHB-Unet is trained to generate the synthetic underwater image. In the generating stage, a real underwater image can be used to render any unrelated clean image into an underwater image.

Using the above operations, the natural light field maps of the real underwater images we obtained are shown in Fig. 2, from which we can found that our estimated light field does not contain object information and it well represents the underwater natural light field information. The reserved features in the underwater light field map focus on the natural style information of diverse underwater scenes without detailed and structured information from original underwater images. In theory, the underwater light field map contains two significant pieces of information for underwater characteristic transferring: $B(x)$ and $\beta$. It’s worth noting that previous methods [18, 45] ignored the importance of $\beta$ for underwater imaging; recent work [27] about hazy image generation calls our attention to focus on implicit estimation of the above two coefficients with entangled way. Our experiments demonstrate the obvious advantages of our transferring procedure.

3.2. Underwater Image Generation Module

**Depth Estimation Network.** After the light field preservation module, we get the information of the light field. In addition, we also need to obtain the depth map that replaces the transmission map in the clean image to preserve the depth information of the generated image. Therefore we design a depth estimation network whose network architecture we utilize Li et al.’s pre-trained model [32]. It is worth saying that our estimated depth map is a biased image. We feed this together into our underwater generative model (MHB-Unet) and let the network learn this biased property. The benefit is that our method does not require paired depth map datasets and can synthesize a large number of underwater images for practicality.

**Multi-branch Hybrid Unet (MHB-Unet).** We design the Unet style architecture network as our underwater image generation model, called MHB-Unet. For further improving the performance of MHB-Unet, we develop a multi-scale hybrid convolutional attention module as shown in Fig. 4(a). Considering that the local features of underwater scenes are complex and diverse, we first obtain different receptive fields through 1x1 and 3x3 convolutions to per-
form multiple feature fusion. At the same time, we also use residual connections, which can solve the problem of vanishing gradients and take into account that the spatial structure and color of certain regions in underwater images are not affected by scene degradation. After multi-branch fusion, we apply a combination of spatial attention Fig. 4(b) and channel attention module Fig. 4(c). The spatial attention mechanism improves the network’s ability to pay attention to complex areas such as light field distribution and depth information in underwater images, while the channel attention pays attention to the network’s expression of important channels in features, thereby improving the overall expression performance of the model.

It is worth noting that our model has excellent performance and even high-resolution images can be rendered well. Please refer to the more discussion section of experiments for more details of actual performance.

![Diagram](image)

**Fig. 4.** From left to right are multi-branch hybrid modules, spatial attention and channel attention, respectively.

### 3.3. Training Losses

We propose following loss functions to ensure the rendering effect of the generated underwater images.

**Reconstruction Loss.** To ensure that our predicted images are enough close to the real underwater images, we introduce the $L_1$ loss as our basic reconstruction loss:

$$
\mathcal{L}_{\text{rec}} = \| \mathcal{I}_u - x_u \|_1.
$$

where $\| \cdot \|$ represents L1 loss, $\mathcal{I}_u$ is the output underwater image by our architecture.

**Perceptual Loss.** Perceptual loss which utilizes the feature layers extracted from pre-trained VGG19 model [43] as the loss network aims to maintain perceptual structure consistency. It is defined as follows:

$$
\mathcal{L}_{\text{per}} = \sum_{j=1}^{5} \frac{1}{C_j H_j W_j} \| \phi_j(\mathcal{I}_u) - \phi_j(x_u) \|_2^2.
$$

where $C_j$, $H_j$, $W_j$ represents the number of channels, length and width of the feature map extracted in the j-th hidden features respectively. $\phi_j$ is the specified j-th layer of the loss network.

**Underwater Dark Channel Loss.** UDCP [11] applied the dark channel prior principle to underwater. We define an underwater dark channel loss to make the generated underwater image consistent with the clean image at the dark channel level.

$$
UDC(x) = \min_{y \in N(x)} \left[ \min_{c \in \{g, b\}} x_i^c(y) \right].
$$

where $x$ and $y$ are pixel coordinates of image $x_i$, $x_i^c$ represents c-th color channel of $x_i$, and $N(x)$ denotes the local neighborhood centered at $x_i$.

The formula for the underwater dark channel loss is as follows:

$$
\mathcal{L}_{\text{udc}} = \| UDC(\mathcal{I}_u) - UDC(x_u) \|_1.\tag{8}
$$

**Light Field Consistency Loss.** In order to effectively maintain the light field characteristic of real underwater images, we introduce the light field consistency loss based on the natural light field map for better rendering performance. We utilize multi-scale Gaussian filter to capture the light field map:

$$
\mathcal{L}_{\text{lfc}}(\mathcal{J}) = \frac{1}{3} \sum_{\sigma} \text{Gauss}_\sigma(\mathcal{J}), \sigma \in \{15, 60, 90\}.\tag{9}
$$

where the $\mathcal{J}$ denotes the image and $\mathcal{L}_f(\cdot)$ denotes the capturing operation of light filed. And the light field consistency loss function is defined as follows:

$$
\mathcal{L}_{\text{lfc}} = \| \mathcal{L}_f(\mathcal{I}_u) - \mathcal{L}_f(x_t) \|_1.\tag{10}
$$

**Overall Loss Function.** The overall loss function is expressed as follows:

$$
\mathcal{L} = \lambda_{\text{rec}} \mathcal{L}_{\text{rec}} + \lambda_{\text{per}} \mathcal{L}_{\text{per}} + \lambda_{\text{udc}} \mathcal{L}_{\text{udc}} + \lambda_{\text{lfc}} \mathcal{L}_{\text{lfc}}.\tag{11}
$$

where $\lambda_{\text{rec}}$, $\lambda_{\text{per}}$, $\lambda_{\text{udc}}$ and $\lambda_{\text{lfc}}$ are trade-off weights.

### 4. Experiments

In this section we conduct extensive experiments to evaluate the effectiveness of our method. First we describe the specific implementation details of our training part. Second, we adopt the FID [17] evaluation metric to objectively evaluate the effect of our generated images, then we adopt PSNR, SSIM and UIQM metrics to measure the effect of our generated underwater image dataset compared with other underwater image generation methods on the underwater enhancement network. Finally we perform ablation experiments to verify the necessity of the components in our framework.
Table 1. Comparative metric analysis of various methods on different datasets. Best results marked in bold.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>NYU Dataset(890)</th>
<th>SUID Dataset(890)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FID</td>
<td>239.36</td>
<td>236.23</td>
</tr>
<tr>
<td>PSNR</td>
<td>-</td>
<td>18.41</td>
</tr>
<tr>
<td>SSIM</td>
<td>-</td>
<td>0.70</td>
</tr>
<tr>
<td>UIQM</td>
<td>-</td>
<td>2.28</td>
</tr>
</tbody>
</table>

Fig. 5. Underwater images synthesized by UWGAN [45] and our method (UWNR) on the NYU [38] dataset.

Fig. 6. Underwater images synthesized by UIQA [18] and our method (UWNR) on the SUID [18] dataset.

4.1. Experiment Settings

**Training Details.** In training details, we select 500 images from the underwater real dataset UIEB [28] as our training set. We also add random rotations between 0, 90, 180 and 270 degrees and horizontal flip for data augmentation. The patch of the input for training is set to 256x256 and we train the network for 200 epochs, where we set the initial learning rate to 0.0002, and after 100 epochs we start a linear decay of the learning rate. We employ the Adam optimizer whose first momentum and second momentum are taken to be 0.9 and 0.999, respectively. For perceptual loss, we select layers 1, 3, 5, 9 and 13 in the VGG19 [43] model to extract hidden features.

**Evaluation Metrics.** How to evaluate the data generated by the image generation network is a vital question. The generated dataset should be visually appealing, diverse, and close to the real domain. Based on the above, we choose FID [17] as the evaluation metrics. FID directly considers
the distance between the synthetic data and the real data at the feature level to measure the difference between the generated image and the real image. The smaller the value, the closer it is to the image in the real domain. To measure the performance of our synthetic dataset, we use it with other synthetic underwater datasets to test the recovered results in the underwater image enhancement network. We apply the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) to objectively evaluate the enhanced performance. In addition, we also use the non-reference indicator UIQM, which evaluates colorfulness (UICM), sharpness (UISM), and contrast (UIConM).

### 4.2. Quantitative Comparison

We apply the various real underwater image \(x_r\) in \(\mathcal{X}_r\) \(\{x_r, y_r\}\) as the reference for the FID evaluation metric, and its distance from the generated underwater dataset measures the reality of the data. We compare our method with previous state-of-the-art underwater image synthesis methods. We select another part of the light field map that is different from the training details to generate images so as to avoid the risk of data leakage. Since the two methods adopt NYU [38] and SUID [18] datasets to generate underwater datasets. For the sake of fairness using FID metric, we compare our method with previous two rendering methods on their datasets respectively. Specifically, for a fair and efficient comparison, we randomly select 890 clean images from the NYU and SUID datasets, using UWGAN [45] and UISA [18] methods to generate corresponding underwater images, and we render the clean images into the underwater images by our method, which also are included in our LNIRU dataset. In the visual comparison, as shown in Tab. 1, we can observe that the results of our method are closer to the natural underwater images than other methods.

At the same time, in order to verify the effectiveness of our method, we use our dataset and the datasets of other methods where the contents of the scenes are the same to train the underwater enhancement network Shallow-UWnet [37], in which we train for 100 epochs, the other training details are the same as those in the original paper. For the test part, we adopt 90 small image dataset in the UIEB dataset to measure the effect of underwater enhancement by utilizing PSNR, SSIM and UIQM metrics.

### Model Complexity Analysis

We analyze the parameters of our model is 11.57M. We also study the computation and inferencing runtime of rendering is 276.26GMac/0.0023s when the image size is 1024×1024, which illustrates the efficiency of our model during generating stage and provides a prerequisite for the actual project landing.

### 4.3. Visual Comparison

To verify the superiority and diversity of our proposed method, we compare the state-of-the-art method of underwater image synthesis and present the results in Fig. 5 and Fig. 6. Due to the mode collapse issue of the GAN architecture, the UWGAN [45] method in the NYU dataset results in a single underwater dataset which only has the green underwater feature. Compared with UWGAN, our method can generate more diverse underwater images with better visual effects.

In the SUID dataset, the UISA [18] method based physical model estimation to generate different types of underwater images. But the estimated deviation of its parameters will cause a large visual difference between the synthetic underwater image and the real underwater environment, which makes the generated underwater images susceptible to appear too bright or too dark to match the real underwater scene. Our framework does not have this problem in terms of visual effects and is more in line with underwater characteristics.

### 4.4. Ablation Study

To validate the necessity of the components in our approach, we performed extensive ablation experiments to validate our claims.

As shown in Tab. 4, we perform ablation experiments of the light field retention (LFR) module and the depth estimation network (DEN) to show their necessity. We eliminate these two modules separately to carry out the following training: (1) LFR+MHB-Unet. (2) DEN+DCP+MHB-Unet. To demonstrate the superiority of our light field preserving module, we replace our light field module with a dark channel prior to obtain background light.(3) LFR+DEN+MHB-Unet (Ours). From the Tab. 3, we can observe that the lack of depth information and light field information will reduce the quantitative performance of our method. In particular, as shown in Fig. 7, compared to the
To elaborate the effectiveness of our loss function, we also perform extensive ablation experiments to eliminate each loss function separately to train our model and test it. From the Tab. 2, we observe that the absence of reconstruction loss causes a significant drop in metrics because it is critical for pixel-level reconstruction of the image. The loss of underwater dark channel is considered from the statistical characteristics of underwater images, and the light field consistency loss is based on the characteristics of underwater ambient light, which provide conditions for underwater image rendering from different angles. Not using them can also cause model performance to drop. As show in Fig. 8, it is worth noting that removing the perceptual loss will make the indicator FID of underwater synthesis more excellent. But in terms of intuitive visual effects, we will find that the loss of detail information is serious, especially in some contours, which is a fatal issue in image synthesis.

And we perform another ablation experiment to demonstrate the effectiveness of the multi-branch Hybrid Block in our MHB-Unet. We present the following three training methods: (1) Unet+Multi-Branch Hybrid Block w/o spatial attention. (2) Unet+Multi-Branch Hybrid Block w/o channel attention. (3) Unet+Multi-Branch Hybrid Block w/o Multi-branch Hybrid Convolution (MHC). From the results in Tab. 4, we prove that the multi-branch-based channel attention and spatial attention can better allow our model to learn the characteristics of real underwater complex environments and achieve better metrics. We also found that the absence of multi-branch convolutions and residual connections reduces our metric performance. Probably because of the lack of multiple feature fusion, the processing power in diverse and complex regions is reduced.

### 4.5. Large Neural Rendering Underwater (LNRU) Dataset

We create a large neural rendering underwater dataset with 50,000 underwater images, in which each underwater image has the associate groundtruth of a latent clean image. Specifically, we collected 5,000 authentic underwater images and randomly selected them for rendering. We will open-source the LNRU dataset after our submitted paper is received. Please refer to our supplementary material for reviewing the thumbnails and more details of LNRU dataset.

### 5. Conclusion

In this paper, we propose an underwater image rendering framework that avoids the problem of ground truth inaccuracy in underwater paired datasets. We utilize light field information and a multi-scale unet network to generate a large number of diverse images using only unpaired images during rendering. In addition to this, we have open-sourced a large-scale underwater image synthesis dataset. The experimental results demonstrate that our method achieves the best results compared with the state-of-the-art methods in terms of vision and metrics.

---

**Table 3.** Ablation experiments with light field retention and depth estimation modules. Bold presents best result.

<table>
<thead>
<tr>
<th>Method</th>
<th>NYU Dataset(890)</th>
<th>SUID Dataset(890)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FID</td>
<td>PSNR</td>
</tr>
<tr>
<td>Ours w/o LFR+DCP</td>
<td>223.52</td>
<td>18.58</td>
</tr>
<tr>
<td>Ours w/o DEN</td>
<td>221.96</td>
<td>18.82</td>
</tr>
<tr>
<td>Ours</td>
<td>221.93</td>
<td>18.86</td>
</tr>
</tbody>
</table>

**Fig. 8.** Top is the underwater image w/o perceptual loss, bottom is the underwater image with perceptual loss.

**Fig. 7.** Top is the underwater image which applys DCP [16] method to replace our LFR module, bottom is the underwater image with LFR.

---

**Table 4.** The network architecture ablation experiment of MHB-Unet, bold indicates the best result.

<table>
<thead>
<tr>
<th>Method</th>
<th>NYU Dataset(890)</th>
<th>SUID Dataset(890)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FID</td>
<td>PSNR</td>
</tr>
<tr>
<td>Ours w/o CA</td>
<td>221.93</td>
<td>18.86</td>
</tr>
<tr>
<td>Ours w/o MHC</td>
<td>221.93</td>
<td>18.86</td>
</tr>
<tr>
<td>Ours</td>
<td>221.93</td>
<td>18.86</td>
</tr>
</tbody>
</table>

495
References


[38] Pushmeet Kohli, Nathan Silberman, Derek Hoiem, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In ECCV, 2012. 6, 7


[45] Nan Wang, Yabin Zhou, Fenglei Han, Haitao Zhu, and Jingzheng Yao. Uwgan: Underwater gan for real-world underwater color restoration and dehazing, 2021. 1, 2, 3, 4, 6, 7
