

# MetaUE: Model-based Meta-learning for Underwater Image Enhancement

Zhenwei Zhang, Haorui Yan, Ke Tang, and Yuping Duan\*

**Abstract**—The challenges in recovering underwater images are the presence of diverse degradation factors and the lack of ground truth images. Although synthetic underwater image pairs can be used to overcome the problem of inadequately observing data, it may result in over-fitting and enhancement degradation. This paper proposes a model-based deep learning method for restoring clean images under various underwater scenarios, which exhibits good interpretability and generalization ability. More specifically, we build up a multi-variable convolutional neural network model to estimate the clean image, background light and transmission map, respectively. An efficient loss function is also designed to closely integrate the variables based on the underwater image model. The meta-learning strategy is used to obtain a pre-trained model on the synthetic underwater dataset, which contains different types of degradation to cover the various underwater environments. The pre-trained model is then fine-tuned on real underwater datasets to obtain a reliable underwater image enhancement model, called MetaUE. Numerical experiments demonstrate that the pre-trained model has good generalization ability, allowing it to remove the color degradation for various underwater attenuation images such as blue, green and yellow, etc. The fine-tuning makes the model able to adapt to different underwater datasets, the enhancement results of which outperform the state-of-the-art underwater image restoration methods. All our codes and data are available at <https://github.com/Duanlab123/MetaUE>.

**Index Terms**—Meta-learning, model-based learning, image enhancement, underwater image

## I. INTRODUCTION

UNDERWATER imaging has played an important role in deep-sea exploration, underwater robot technology and marine ecological monitoring. However, underwater optical images are vulnerable to underwater turbulence, diffusion, severe absorption, scattering of water bodies, various noises, low contrast, uniform illumination, monotonous color, and complex underwater scenes [1]. Four typical underwater images are displayed in Fig. 1, which makes the high-level image processing tasks based on underwater images become very challenging.

The work was partially supported by the National Natural Science Foundation of China (NSFC 12071345, 11701418). Asterisk indicates the corresponding author.

Z. Zhang and H. Yan are with Center for Applied Mathematics, Tianjin University, Tianjin 300072, China.

K. Tang is with the Guangdong Key Laboratory of Brain-Inspired Intelligent Computation, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China, and also with the Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology, Shenzhen 518055, China. E-mail: tangk3@sustech.edu.cn.

Y. Duan is with Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. E-mail: doveduan@gmail.com.

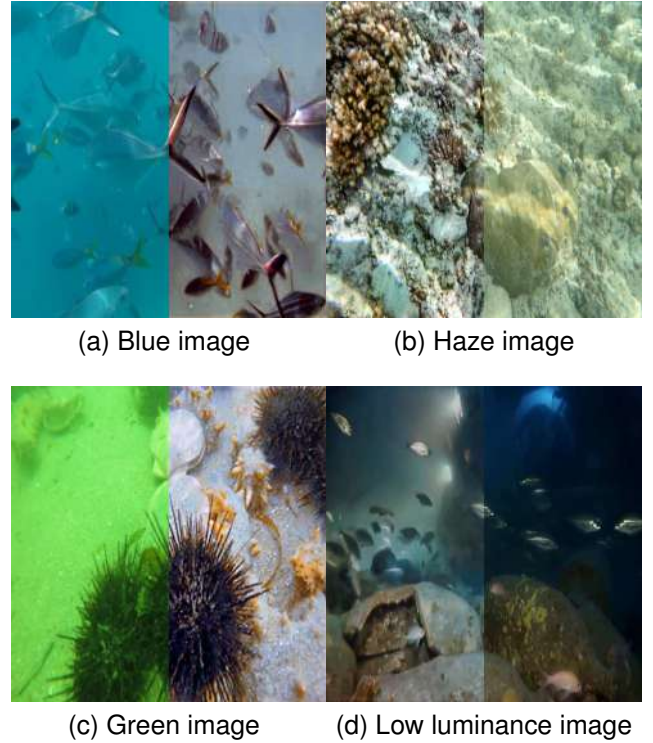


Fig. 1. Illustration of typical underwater images and our enhancement results.

Underwater image degradation is a complex process, mainly influenced by light absorption and scattering. Various methods have been developed to recover clean images from underwater images including traditional methods and learning-based methods. Akkaynak *et al.* [2], [3] proposed an accurate underwater image formation model and used mathematical methods to estimate the coefficients, which can consistently remove water from underwater images. By a rank-one transmission prior, Liu *et al.* [4], [5] presented a real-time light correction method to recover the degraded scenes of underwater, sandstorms and haze, etc. The regularization methods were also used to recover underwater images, such as depth-aware total variation regularization [6], curvature variation regularization [7], and adaptive non-convex non-smooth variation regularization [8], etc. In recent years, the deep learning methods have provided new solutions for underwater image enhancement, which can be roughly divided into Generative Adversarial Network (GAN)-based methods and Convolutional Neural Network (CNN)-based methods. Since the discriminator of GAN can eliminate unreliable data, Li *et al.* [9] generated realistic

underwater images from the in-air images for color correction of underwater images. Later, GAN-based methods have been intensively studied for underwater image enhancement [10]–[15]. On the other hand, Li *et al.* [16] presented a CNN model using the scene prior for underwater image enhancement on the synthesized underwater datasets. Since then, various CNN-based models have been developed such as the Water-Net [17], JLCL-Net [18], Ucolor [19], etc. However, the existing deep learning-based underwater image enhancement methods may lose their effects on images of unknown water, which limits their usages in real applications. Indeed, model-driven deep learning methods are well-known for their high interpretability and abilities in dealing with different image processing tasks. The plug-and-play strategy as a typical model-driven method has been used for image super resolution [20], image denoising [21], image deblurring [22], etc. Most recently, Jonathan, Jain and Abbeel [23] presented the diffusion probabilistic models inspired by considerations from nonequilibrium thermodynamics, which have emerged as a powerful new family of deep generative models with a record-breaking performance in many applications, including image synthesis [24], [25], denoising [26], [27], and segmentation [28], etc.

Deep meta-learning [29]–[31] is a knowledge-driven machine learning framework that attempts to solve the problem of how to learn. Meta-learning has achieved remarkable results on no or fewer reference image processing problems. For instance, Zhu *et al.* [32] presented a no-reference image quality assessment metric based on deep meta-learning to learn the meta-knowledge shared by human when evaluating the quality of images with various distortions. Zhao *et al.* [33] proposed a novel federated meta-learning enhanced acoustic radio cooperative framework to do the transfer. Wang *et al.* [34] proposed a meta memory bank to improve the generalization of segmentation networks by bridging the domain gap between source and target domains. The model-agnostic nature of meta-learning can facilitate underwater image enhancement tasks, to produce good generalization performance in diverse underwater environments.

In this paper, we propose a task-tailored deep learning method based on the physical model of underwater images to recover clean underwater images, which can accurately describe the underwater image distortions. Our network consists of three sub-networks to estimate the clean image, background light, and transmission map, respectively, which are combined according to the underwater image model to obtain the supervision loss. Meta-learning is used to obtain a pre-trained model based on synthetic images, which is then fine-tuned on real underwater images. The pre-trained model is made to capture the shared prior knowledge of the degraded underwater images, while the fine-tuned model can make the pre-trained model easy to adapt to unknown distortions. To sum up, our major contributions are concluded as follows

- 1) We propose a model-driven multi-variable convolution neural network based on the underwater image degradation model, which improves the interpretability by estimating different variables through sub-networks.
- 2) We implement the meta-learning strategy to learn the shared prior knowledge among different types of dis-

tortions on a sophisticated underwater image dataset generated from in-air images to cover a wide range of degradation images, which greatly promotes the generalization ability on diversified distortions.

- 3) We evaluate our meta-learning underwater image enhancement model (shorted by MetaUE) on different underwater image datasets, which outperforms the state-of-the-art methods in terms of enhancement qualities and generalization ability, particularly for images with strong light scattering and insufficient lighting.

The rest of the paper is organized as follows. Section II reviews the most important supervised and unsupervised learning-based methods for underwater image enhancement tasks. In Section III, we propose the underwater image enhancement method based on the underwater image physical model. The meta-learning strategy is presented in Section IV, where a synthetic underwater image dataset is developed for training the pre-trained model. Numerical experiments are provided in Section V by comparing with the state-of-the-art underwater enhancement methods. We conclude the work in Section VI and discuss its possible applications for other low-quality image enhancement problems.

## II. RELATED WORKS

### A. Supervised learning-based methods

With the development of deep learning methods, it is possible to restore clean images from complex underwater environments. Supervised learning-based methods can effectively extract feature expressions from degraded underwater images to high-quality images. As a pioneering work, Li *et al.* [10] proposed a weakly supervised color transfer (WSCT) method by learning a cross-domain mapping function between ground-truth images and underwater images. Hong *et al.* [35] proposed a weakly supervised underwater enhancement model, where a GAN-based architecture is designed to enhance underwater images by unpaired image-to-image transformation. Similarly, weakly supervised methods are proposed to deal with underwater tasks, such as underwater fish segmentation [36], [37], underwater object detection [38].

However, these weakly supervised methods still have much room for improvement to obtain satisfactory underwater images. Using the paired underwater/clean images to train the supervised methods is a good strategy to improve the performance of underwater problems. Li *et al.* [16] presented an underwater image enhancement convolutional neural network (UWCNN) based on the synthesized underwater dataset, which was built up based on the physical underwater image model and Jelov water type [34]. The training dataset in [16] was also used in [39], where the all-in-one network adversatively learned the domain-agnostic features to generate enhanced underwater images from the degraded images of ten different water types. Liu *et al.* [40] derived a new physical underwater image model to generate a wide range of degradation images from the clean underwater images, where the proposed simulation images were selected by color deviations and the Sobel edge map to generate reliable dataset. Wang *et al.* [41] noticed that only training on simulation data makes many

learning methods fail in real underwater images. Thus, they proposed a novel two-phase underwater domain adaptation network to simultaneously minimize the synthetic data and real underwater data gap.

Since there exist significant differences between synthetic and real underwater datasets, which may cause poor generalization in real underwater environments. Li *et al.* [17] constructed an underwater enhancement benchmark (UIEB) for real underwater images, which obtained real underwater images by using the existing enhancement methods. Based on the UIEB [17], the following effective deep-learning networks are proposed to recover the underwater images. Xue *et al.* [18] proposed the joint luminance and chrominance learning network (JLCL-Net). Wang *et al.* [42] developed the UIEC<sup>2</sup>-Net using two color spaces to improve luminance and saturation, respectively. Li *et al.* [19] proposed a network containing the multi-color spaces guided by the medium transmission to cope with color casts and low contrast problems. Liu *et al.* [43] developed an adaptive learning attention network for underwater image enhancement based on supervised learning.

### B. Unsupervised learning-based methods

On the other hand, unsupervised methods were also intensively studied for underwater image enhancement. Fabbri *et al.* [14] used the cycle-consistent adversarial network (CycleGAN) to directly learn to generate degraded underwater images based on two separate groups of clean and degraded real-world images, which are then used to train a UNet for underwater image enhancement. Islam *et al.* [15] utilized the same CycleGAN-based method to build up a large dataset called EUVP and proposed a lightweight conditional GAN to recover the underwater images (FUnIE-GAN). Yang *et al.* [12] proposed a conditional generative adversarial network to improve the quality of the underwater images. Zhou *et al.* [44] proposed an unsupervised underwater loop enhancement network (ULENet) to improve the turbid underwater images. Based on the underwater image formation model, Chai *et al.* [45] proposed an unsupervised method to estimate the components of the physical image model, such as scene, backscatter, etc. Guo *et al.* [46] used the transformer module to capture the global information to train the unsupervised network. Liu *et al.* [13] proposed an unsupervised method called twin adversarial contrastive learning-based underwater enhancement method (TACL) to enhance the quality of the underwater images. Jiang *et al.* [47] proposed a novel domain adaptation framework for real-world underwater image enhancement inspired by transfer learning, which transfers in-air image dehazing to real-world underwater image enhancement. Alejandro *et al.* [48] presented a novel state-of-the-art end-to-end deep learning architecture for underwater image enhancement focused on solving key image degradations related to blur, haze, and color casts and inference efficiency.

## III. OUR APPROACH

### A. Underwater image formation

One crucial challenge of underwater image enhancement is the endless diversity versus the limited diversity of the training

samples, which harms the generalization ability of the conventional underwater methods trained on the synthesized dataset. As shown in Fig. 2, the radiance perceived by the camera  $I_\lambda$  is the sum of direct signal and background scattering, which is caused by light reflected by particles suspended in the water column. By ignoring artificial light, the underwater images can be modeled as follows

$$I_\lambda(x) = J_\lambda(x)t_\lambda(x) + B_\lambda(1 - t_\lambda(x)), \quad \lambda \in \{R, G, B\}, \quad (1)$$

where  $J_\lambda(x)$  is the clean image,  $B_\lambda$  is the background light and  $t_\lambda(x)$  is the transmission map defined by  $t_\lambda(x) = e^{-c_\lambda d(x)}$  with  $c_\lambda$  being the attenuation coefficients and  $d(x)$  being the transmission distance. We propose a multi-variable CNN model to simultaneously estimate the clean image, background light, and transmission map according to the underwater image model (1), which can ideally adapt to complex and diversified underwater environments.

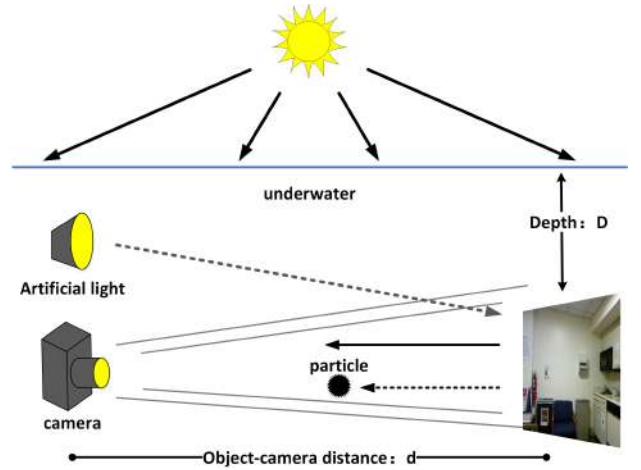


Fig. 2. The radiance perceived by the camera  $I_\lambda$  is the sum of direct signal and background scattering.

### B. Network architecture

The overall framework of our model-based meta-learning model is illustrated in Fig. 3, which contains three sub-networks, i.e., the clean image estimation sub-network  $N_J$ , background light estimation sub-network  $N_B$ , and the transmission map estimation sub-network  $N_T$ . More specifically, the **clean image estimation sub-network** is used to estimate the clean image  $J_\lambda$  from the observed underwater image as follows

$$J_\lambda = N_J(I_\lambda; \theta_J),$$

where  $I_\lambda$  is the input image and  $\theta_J$  represents the parameters of network  $N_J$ . The **background light estimation sub-network** is used to evaluate the background light  $B_\lambda$

$$B_\lambda = N_B(I_\lambda; \theta_B),$$

where  $\theta_B$  represents the parameters of network  $N_B$ . Finally, the **transmission map estimation sub-network** aims to estimate the transmission map  $t_\lambda(x)$  from an underwater image as below

$$t_\lambda = N_T(I_\lambda; \theta_T),$$

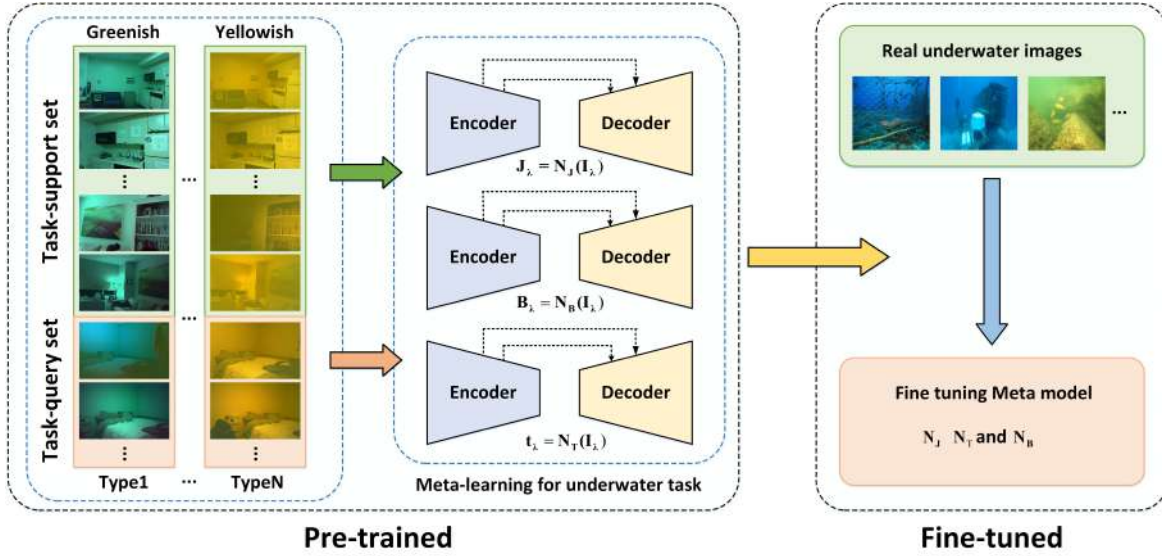


Fig. 3. The overview framework of our deep meta-learning approach for underwater images enhancement, where three variables are jointly pre-trained and fine-tuned on synthetic and real underwater image dataset, respectively.

where  $\theta_T$  is the parameters of network  $N_T$ . All the sub-networks are developed based on an encoder-decoder architecture with skip connections and shortcut connections [49] to perform the translations from degraded images to target images. The detailed network architectures are provided in Fig. 4. As can be seen, the network consists of four encoder blocks and four decoder blocks with the skip connections from each encoder block to the corresponding decoder block. To ensure that useful information is well preserved in the block outputs, we also build an open path by shortcut connection in each block. In the encoder, the image is finally downsampled into 512 feature maps, while we add a bilinear interpolation operation to up-sample the input feature maps in the decoder part.

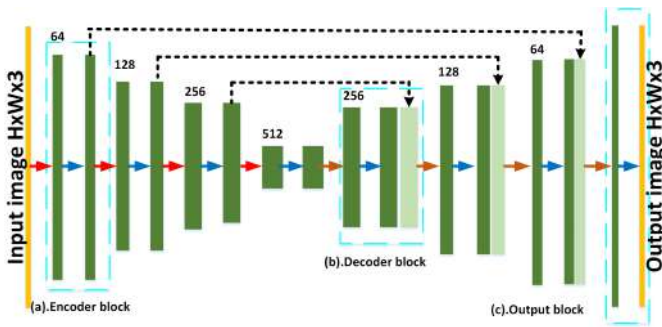


Fig. 4. The architecture of the proposed underwater image enhancement network, where (a) is an encoder block, (b) is the decoder block and (c) is the output block.

#### IV. META-TRAINING STRATEGY

##### A. Meta-learning method

As shown in Fig. 3, the meta-learning strategy is used to build up our underwater image enhancement model. More specifically, we generate the meta training set  $D_{meta}^{p(\tau)} =$

$\{\mathcal{D}_s^{\tau_n}, \mathcal{D}_q^{\tau_n}\}_{n=1}^N$ , where  $\mathcal{D}_s^{\tau_n}$  and  $\mathcal{D}_q^{\tau_n}$  are the support set and query set of each task, and  $N$  is the total number of tasks. To learn a generalized model from different tasks, the sample  $k$  tasks are randomly chosen as a mini-batch from the meta training set. When the model is applied to a new task  $\tau_i$ , the parameters of our model are updated according to the task requirements, with the loss function of the  $i$ -th support set  $\mathcal{D}_s^{\tau_i}$  denoted by  $L_{\tau_i}$  for  $i \in \{1, 2, \dots, k\}$ . To quickly obtain such ability, we use a two-level gradient descent method to update the model parameters on the support set and query set accordingly as follows

$$\begin{cases} \theta'_{J,i} \leftarrow \theta_J - \alpha \nabla L_{\tau_i}(N_J(\theta_J)); & \text{on the support set} \\ \theta_{J,i} \leftarrow \theta'_{J,i} - \alpha \nabla L_{\tau_i}(N_J(\theta'_{J,i})); & \text{on the query set} \end{cases}$$

where  $\alpha$  is the step size of meta-learning. The gradients of all tasks are integrated to obtain the final update for the model parameters such as

$$\theta_J \leftarrow \theta_J - \beta \frac{1}{k} \sum_{i=1}^k (\theta_J - \theta_{J,i}),$$

where  $\beta$  is the outer learning rate. Similarly, the parameters  $\theta_B$  and  $\theta_T$  are also updated by two-level gradient descent method.

To adapt to real underwater environments, we fine-tune the pre-trained model on real underwater images with unknown distortions. Then, we leverage the Adam optimizer to update the enhancement network  $N_J$  on real underwater images as follows

$$\theta_J \leftarrow \theta_J - \alpha_f \nabla L(N_J(\theta_J)),$$

where  $\alpha_f$  is the learning rate of fine-tuning and the other network parameters  $\theta_B$  and  $\theta_T$  are updated in the similar way. Then, the fine-tuned model can be obtained for enhancing the underwater images with unknown distortions. The implementation details of MetaUE are summarized in Algorithm 1.



TABLE I  
ATTENUATION COEFFICIENTS  $e^{-c_\lambda}$  ARE USED FOR SYNTHESIZING NINE KINDS OF UNDERWATER IMAGES.

Types	I	II	III	B	3	G	5	7	Y
blue	0.982	0.94	0.89	0.88	0.8	0.75	0.67	0.5	0.40
green	0.961	0.925	0.885	0.80	0.82	0.79	0.73	0.61	0.60
red	0.805	0.8	0.75	0.70	0.71	0.69	0.67	0.62	0.61

---

**Algorithm 1:** Model-based Meta-learning for Underwater Enhancement (MetaUE)

---

**Input:** Meta-trained set  $\mathcal{D}_{meta}^{p(\tau)} = \{\mathcal{D}_s^{\tau_n}, \mathcal{D}_q^{\tau_n}\}_{n=1}^N$   
**Output:**  $J_\lambda$ ,  $B_\lambda$  and  $t_\lambda$   
 /\* Pre-trained on synthetic image dataset \*/

```

1 for  $epoch = 1, 2, \dots, do$ 
2   Sample a mini-batch of  $k$  tasks in  $\mathcal{D}_{meta}^{\tau_i}$  with
    $i \in \{1, 2, \dots, k\};$ 
3   for  $i = 1, 2, \dots, k$  do
4     Compute  $\theta_{J,i} \leftarrow \theta_J - \alpha \nabla L_{\tau_i}(N_J(\theta_J));$ 
5     Compute  $\theta_{B,i} \leftarrow \theta_B - \alpha \nabla L_{\tau_i}(N_B(\theta_B));$ 
6     Compute  $\theta_{T,i} \leftarrow \theta_T - \alpha \nabla L_{\tau_i}(N_T(\theta_T));$ 
7   end
8   update  $\theta_J \leftarrow \theta_J - \beta \frac{1}{k} \sum_{i=1}^k (\theta_J - \theta_{J,i});$ 
9   update  $\theta_B \leftarrow \theta_B - \beta \frac{1}{k} \sum_{i=1}^k (\theta_B - \theta_{B,i});$ 
10  update  $\theta_T \leftarrow \theta_T - \beta \frac{1}{k} \sum_{i=1}^k (\theta_T - \theta_{T,i});$ 
11 end
  /* Fine-tuned on real image dataset */
12 Fine-tuning the model  $\theta_J$ ,  $\theta_B$ , and  $\theta_T$  as
    
$$\begin{cases} \theta_J \leftarrow \theta_J - \alpha_f \nabla L(N_J(\theta_J)); \\ \theta_B \leftarrow \theta_B - \alpha_f \nabla L(N_B(\theta_B)); \\ \theta_T \leftarrow \theta_T - \alpha_f \nabla L(N_T(\theta_T)). \end{cases}$$


```

---

### B. Synthetic underwater image dataset

In the following, we build up an synthetic underwater image dataset to cover different underwater degradation images for implementing Algorithm 1. As shown in Fig.2, the incident light passing through the surface of water is the major source of illumination in an underwater environment. Here, we use the NYU-V1 dataset [50] as a benchmark dataset to generate various underwater images. Due to beam attenuation, the incident light passing through the water may make the underwater environment appearing blue, green or yellow. On the other hand, the influence of artificial light should be also considered to overcome the disadvantage of insufficient underwater illumination. Therefore, the underwater illumination can be simplified as

$$E_\lambda(x) = \omega_a E_\lambda^S e^{-c_\lambda D} + \omega_b E_\lambda^A e^{-c^\lambda d(x)}, \quad (2)$$

where  $\omega_a$  and  $\omega_b$  are two weights,  $E_\lambda^S$  is light on the water surface,  $E_\lambda^A$  is artificial light,  $d(x)$  is the distance from object to the camera and  $D$  is the water depth. The backscattering component mainly contributes to the hazy look of underwater

images. An efficient model was proposed in [51] to calculate the background light  $B_\lambda$  by

$$B_\lambda(x) = \kappa E_\lambda(x) / c_\lambda,$$

where  $\kappa$  is a scalar defined by the camera system. Then the simulation formula from the scene  $J^{gt}$  without attenuation to the underwater image  $I_\lambda$  can be expressed as follows

$$I_\lambda(x) = t_\lambda(x) J_\lambda^{gt}(x) \frac{E_\lambda(x)}{E_\lambda^S} + \frac{\kappa E_\lambda(x)}{c_\lambda} (1 - t_\lambda(x)). \quad (3)$$

Here, depending on the camera principle,  $J_\lambda^{gt}(x) \frac{E_\lambda(x)}{E_\lambda^S}$  is to remove the light  $E_\lambda^S$  at  $x$  of clean image  $J_\lambda^{gt}$  and use  $E_\lambda(x)$  as the new lighting condition.

**Simulation of lighting parameters  $E_\lambda(x)$ .** To effectively simulate various illumination conditions, the range of air light  $E_\lambda^S$  is set to  $[0.7, 1]$  and the depth variable  $D$  is randomly selected from  $[5m, 20m]$ . The artificial illumination is the only light source to provide uneven brightness and limited visibility. Here, we use two-dimensional Gaussian distribution to simulate the artificial light in the water, where beam pattern  $E_\lambda^A = \mathcal{P}(\tilde{x} | E_\lambda^{art}, \sigma)$  with a peak value  $E_\lambda^{art} \in [0.7, 1]$  and standard deviation  $\sigma$  proportional to the width of the image by a random rate in  $[0.2, 1.1]$ . The light source  $\tilde{x}$  is randomly picked from the image.

**Simulation of channel-wise attenuation coefficients  $c_\lambda$ .** The attenuation coefficients of Jerlov water types [16], [52] are used to generate different water types. To better adapt to complex underwater situations, we remove the clearest water quality Type 1 and the most turbid Type 9 and introduce three kinds of water quality for the three common underwater colors (i.e., Blue, Green, and Yellow), named Type B, G, Y. The corresponding values of  $c_\lambda$  for all nine types of underwater images are presented in Table I.

**Simulation of transmission distance  $d(x)$ .** We use the NYU-V1 dataset [50] to generate the synthetic underwater image dataset, which contains a total of 3733 RGB images and corresponding depth maps.

To sum up, we use the parameters in Table II to build up a synthetic underwater image dataset based on the RGB-D NYU-V1 dataset. In particular, we randomly selected parameters three times for each water type to obtain a total of 27 specific underwater distortion types and 100791 images.

### C. Loss function

To effectively train our multi-variable CNN model, we introduce a novel task-specific loss function as below.

**Clean image supervision loss.** To remove water, we compensate the disparities of wavelength attenuation for traversing

TABLE II  
THE PARAMETERS USED TO GENERATE THE SYNTHETIC UNDERWATER  
IMAGE DATASET, WHERE  $\lambda \in \{R, G, B\}$ .

Notation	Description	Range
$D$	Water depth	$[5m, 20m]$
$d$	Transmission distance	NYU-V1 dataset [50]
$c_\lambda$	Attenuation coefficients	Table I
$E_\lambda^S$	Air light	$[0.7, 1]$
$E_\lambda^{art}$	Peak value of artificial light	$[0.7, 1]$
$\tilde{x}$	Location of $E_\lambda^{art}$	A random point in image
$\sigma$	Coverage of artificial light	Random rate $[0.2, 1.1]$
$\omega_a, \omega_b$	Weights of lighting	$\omega_a \in [0, 1]$ and $\omega_a + \omega_b = 1$
$\kappa$	Camera system parameter	$[0.7, 1.1]$
$E_\lambda^A$	Artificial light	$E_\lambda^A = \mathcal{P}(\tilde{x} E_\lambda^{art}, \sigma)$
$E_\lambda$	Underwater illumination	estimated by (2)
$t_\lambda$	Transmission map	$t_\lambda = e^{-c_\lambda d}$
$B_\lambda$	Background light	$B_\lambda = \kappa E_\lambda / c_\lambda$

the water depth to the surface of water. To be specific, we use  $L_1$  loss function to measure supervision loss of clean image

$$L_J = \|J_\lambda - J_\lambda^{gt}\|_1,$$

where  $J_\lambda^{gt}$  denotes the clean image without attenuation.

**Background light supervision loss.** Similarly, the background light is supervised by the  $L_1$  loss function as follows

$$L_B = \|B_\lambda - B_\lambda^{gt}\|_1,$$

where  $B_\lambda^{gt} = \kappa E_\lambda / c_\lambda$  can be regarded as the ground truth of the background light.

**Transmission map supervision loss.** We penalize the difference between  $t_\lambda$  and the reference transmission map  $t_\lambda^{gt}$  by

$$L_T = \|t_\lambda - t_\lambda^{gt}\|_1,$$

where the reference transmission map is given by  $t_\lambda^{gt} = e^{-c_\lambda d}$ .

**Underwater supervision loss.** Based on the physical model (1), the underwater image can be calculated by

$$\tilde{I}_\lambda(x) = J_\lambda t_\lambda + B_\lambda(1 - t_\lambda).$$

Thus, the loss based on the physical process is defined to minimize the difference between the underwater  $I_\lambda$  and the physics-based estimation solution as follows

$$L_I = \|I_\lambda - \tilde{I}_\lambda\|_1.$$

Finally, our multi-variable CNN model is jointly optimized by a weighted combination of  $L_J, L_B, L_T$ , and  $L_I$  as follows

$$L = c_J L_J + c_B L_B + c_T L_T + c_I L_I, \quad (4)$$

where  $c_J, c_B, c_T, c_I$  are the weights used to balance different terms. During the meta-learning implementation, the loss function  $L_{\tau_i}$  of  $i$ -th support set is defined as (4). The weights in our loss function (4) are empirically set as  $c_J = c_B = c_T = 1$  and  $c_I = 0.5$  for the pre-trained model. Due to the lack of ground truth of transmission map and background lights for underwater images, we set the weights as  $c_J = 1$ ,  $c_B = c_T = 0$  and  $c_I = 1$  when we fine-tune the model.

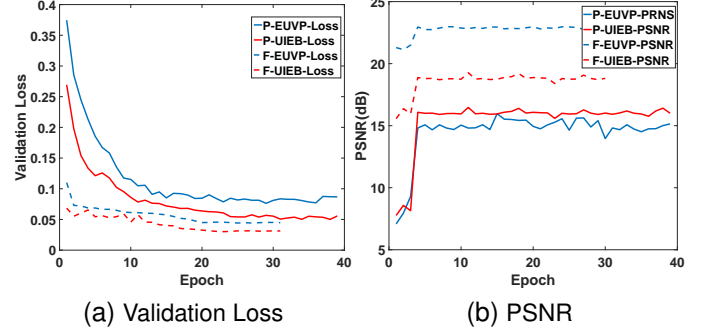


Fig. 5. There are the performances of validation loss and averaged PSNR for the pre-trained and fine-tuned model on the UIEB dataset and EUVP dataset during the training process.

## V. EXPERIMENTAL RESULTS

### A. Implementation details

In the proposed model, all training images are cropped to  $256 \times 256$  pixel patches to feed into the proposed model. We train our model using bi-level gradient optimization with an inner learning rate  $\alpha$  of  $1e-4$  and outer learning rate  $\beta$  of  $5e-5$ , which is implemented based on Pytorch [53]. The total epoch is set to 40, and the batch size of training data is fixed to 8. We set the fine-tuned learning rate  $\alpha_f$  to  $1e-5$  and the total epoch is 30. These learning rates drop to a factor of 0.8 after every five epochs. The mini-batch size is set as  $k = 5$ . The learning rates drop to a factor of 0.8 after every two epochs.

### B. Comparison methods

The performance of our MetaUE is compared against eight state-of-the-art underwater image enhancement methods, including three model-based methods, i.e., ACDC [54], Rank1 [4] and MMLE [55], and five learning-based methods, i.e., UWCNN [16], Water-Net [17], FUnIE-GAN [15], Ucolor [19], and TACL [13]. Note that both Water-Net [17] and Ucolor [19] were trained on the UIEB dataset, while FUnIE-GAN [15] was trained on the EUVP dataset. The UWCNN [16] was trained in simulation of multiple water types, while TACL [13] was trained on UIEB and BSD500 to train the twin adversarial enhancement network.

**Training samples.** The pre-trained model is fine-tuned on two underwater image datasets with reference images, i.e., UIEB [17] and EUVP [15], respectively. For a fair comparison, we also retrained the Water-Net and Ucolor on the EUVP dataset and retrained FUnIE-GAN on the UIEB dataset. Since the trainable codes of both UWCNN and TACL are not available, we use the public models directly in comparison.

**Testing samples.** For testing the learning ability and generalization ability of our MetaUE, six underwater image datasets are used to evaluate the performance, which are UIEB [17], EUVP [15], U45 [56], UIEB-60 [17], EUVPUN [15], UCCS [57]. In particular, it contains 90 paired images of UIEB and 200 paired images of EUVP, respectively. We randomly selected 150 unpaired images from EUVPUN, 20 images from

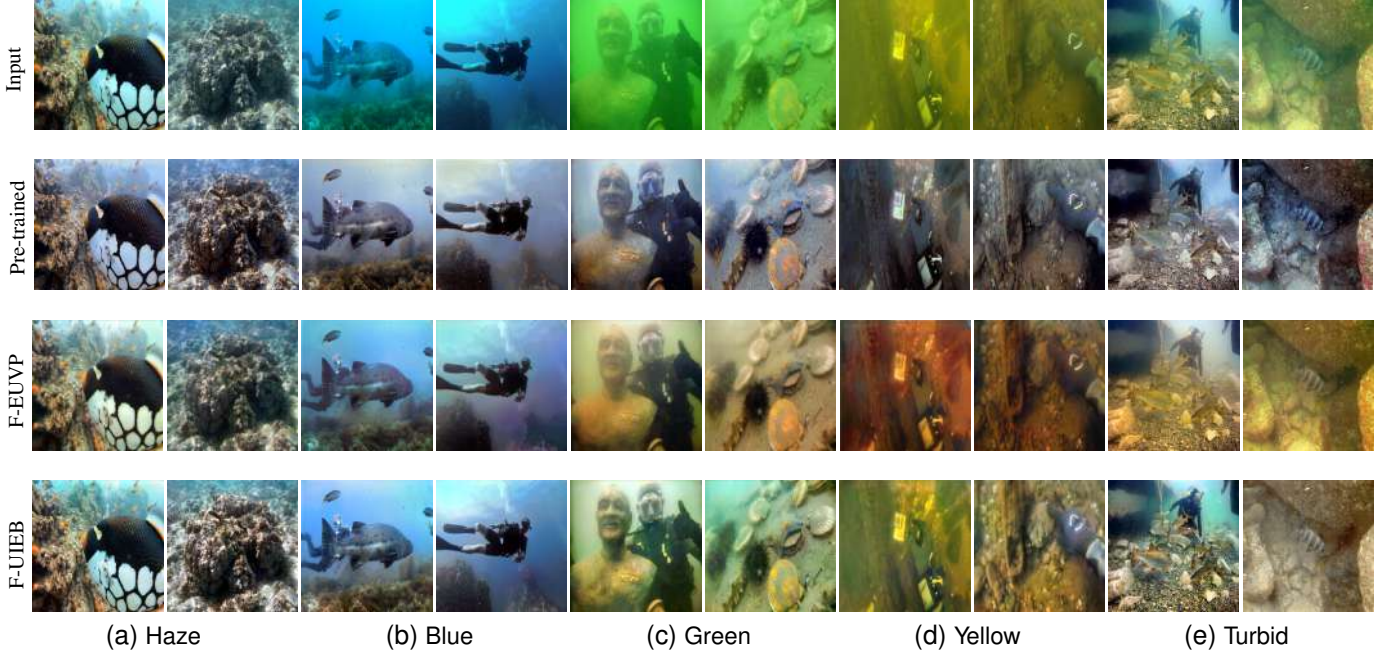


Fig. 6. The enhanced results of pre-trained model for common underwater degradation types, where F-EUVP and F-UEIB are the fine-tuned models on EUVP and UIEB, respectively.

UCCS, and all real underwater images of U45 and UIEB-60 for comparison.

### C. Quantitative evaluation

**Full Reference Evaluation:** For the test datasets with reference images, we conduct a full reference evaluation using three commonly-used metrics, PSNR, SSIM, and MSE. These results of full-reference image quality evaluation by using the reference images can provide realistic feedback on the performance of different methods to some extent. A higher PSNR and SSIM or lower MSE indicates that the result is close to the reference image.

**Non-reference Evaluation:** For the test datasets without references, we employ two non-reference metrics, i.e., UCIQE [58] and UIQM [59]. A higher UCIQE or UIQM score indicates a better human visual perception. The UCIQE is a linear combination of chroma  $\sigma_c$ , contrast  $con_l$ , and saturation  $\mu_s$  defined as follows

$$UCIQE = c_1 \times \sigma_c + c_2 \times con_l + c_3 \times \mu_s$$

with  $c_1 = 0.4680$ ,  $c_2 = 0.2745$  and  $c_3 = 0.2575$  as in [58]. A higher UCIQE gives a better tradeoff among chroma, contrast and saturation. The UIQM is a linear composition of UICM, UISM, and UIConM, which represent the metrics of colorfulness, sharpness and contrast, respectively. The formula of UIQM is given as follows

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UIConM$$

with  $c_1 = 0.0282$ ,  $c_2 = 0.2953$ ,  $c_3 = 3.5753$  as suggest in [59]. A higher UIQM indicates a better tradeoff among colorfulness, sharpness and contrast.

TABLE III  
THE AVERAGED UNSUPERVISED SCORES FOR THE TEST IMAGES IN FIG. 6.

Index	UCIQE	$\sigma_c$	$con_l$	$\mu_s$
Pre-trained	<b>0.50</b>	<b>0.5483</b>	<b>0.8732</b>	0.001
F-UIEB	0.37	0.2819	0.8581	0.0014
F-EUVP	0.37	0.3302	0.782	<b>0.0015</b>
Index	UIQM	UICM	UISM	UIConM
Pre-trained	<b>1.66</b>	4.5596	<b>1.7938</b>	0.2763
F-UIEB	1.59	<b>4.9446</b>	1.3277	0.299
F-EUVP	1.55	4.7811	0.953	<b>0.3114</b>

### D. Comparison between the pre-trained and fine-tuned models

We use meta-learning to obtain a pre-trained model by learning the shared knowledge among different types of distortions on the synthetic underwater image dataset. Based on the pre-trained model, we then fine-tune two image enhancement models on the real underwater image dataset, i.e., UIEB and EUVP, where F-UIEB and F-EUVP are used to denote the fine-tuned models on UIEB and EUVP, respectively. For simplicity, we use P-UIEB and P-EUVP to represent the results of the pre-trained model on the UIEB and EUVP datasets, respectively. As demonstrated in Fig. 5, we monitor the training process by tracking the validation loss and averaged PSNR on the test datasets of UIEB and EUVP. We can observe that both pre-trained models converge in around 30 epochs, while the fine-tuned models require about 20 epochs to reach convergence. More importantly, the values of loss function and PSNR have been significantly improved by the fine-tuning process on real underwater images, which convinces the merit of meta-learning strategy in dealing with such image restoration problems.

Generalization is an important capability for the underwater image enhancement model, which can equip it to handle



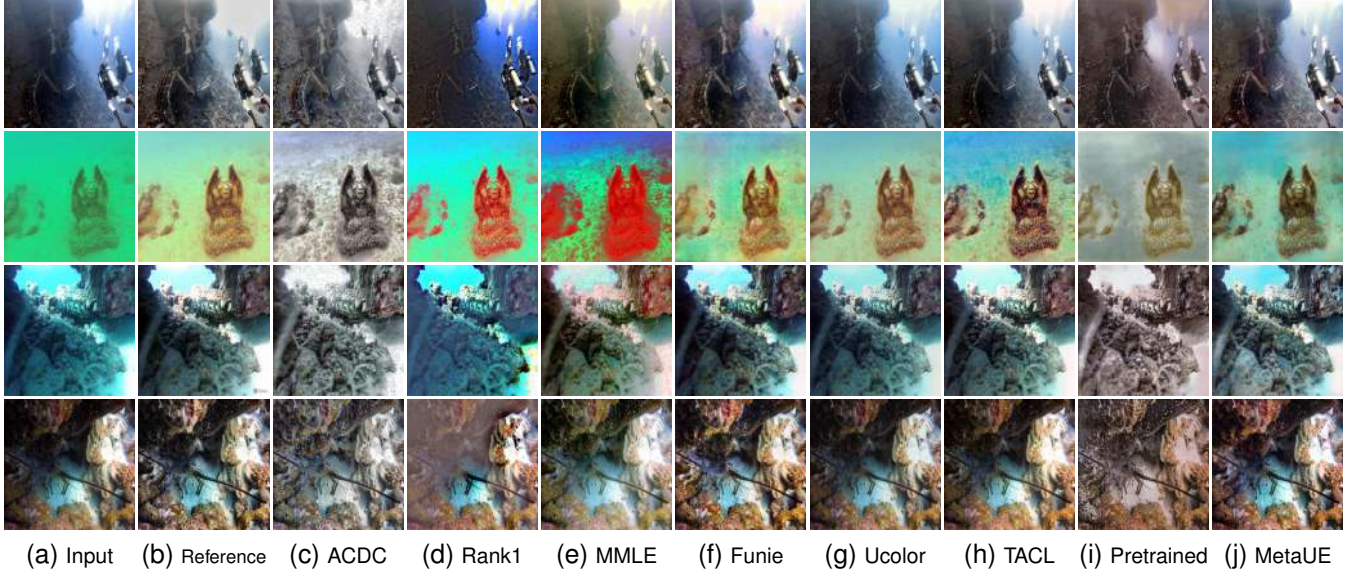


Fig. 7. The visual comparison among the enhancement methods on the paired underwater images, where our MetaUE is fine-tuned on the UIEB dataset.

TABLE IV  
THE PRE-TRAINED MODEL IS FINE-TUNED ON THE UIEB AND COMPARED WITH OTHER METHODS.

Datasets Methods	UIEB				EUVP				U45		UIQS		UIEB-60		EUVPUN	
	PSNR	SSIM	MSE	UIQM	PSNR	SSIM	MSE	UIQM	UIQM	UCIQE	UIQM	UCIQE	UIQM	UCIQE	UIQM	UCIQE
ACDC [54]	18.04	0.8212	1.21	1.26	14.74	0.659	2.53	1.22	1.03	1.02	0.73	0.34	1.03	0.38	1.14	0.41
Rank1 [4]	18.16	0.8443	1.21	1.63	15.16	0.7121	2.56	1.45	1.57	0.48	1.51	0.48	1.36	0.4	1.48	0.42
MMLE [55]	15.48	0.8242	2.34	1.48	14.84	0.6331	2.56	1.54	1.38	<b>0.49</b>	1.58	0.51	1.53	0.38	1.35	0.31
UWCNN [16]	13.43	0.6253	3.88	0.89	14.84	0.7161	1.33	1.28	0.68	0.31	0.88	0.23	0.63	0.29	0.66	0.31
WaterNet [17]	18.35	0.8397	1.34	1.73	12.24	0.5576	1.69	1.69	1.57	0.37	1.28	0.42	1.63	0.41	1.71	0.39
Funie [15]	21.53	0.8512	0.55	1.69	18.31	0.7393	1.15	1.56	1.45	0.39	1.46	0.39	1.61	0.38	1.44	0.38
Ucolor [19]	20.63	0.8495	1.08	1.74	15.54	0.7215	1.55	1.69	1.72	0.39	1.69	0.41	1.63	0.34	1.59	0.39
TACL [13]	20.41	0.8478	0.48	1.77	18.42	0.7501	1.24	2.24	1.85	0.42	1.59	0.39	1.58	0.41	1.91	0.43
Pre-trained	16.42	0.7354	1.54	<b>1.95</b>	15.57	0.7145	1.96	<b>2.15</b>	1.88	0.46	<b>1.72</b>	<b>0.55</b>	1.55	<b>0.45</b>	1.87	<b>0.46</b>
MetaUE	<b>22.01</b>	<b>0.8607</b>	<b>0.54</b>	1.81	<b>19.11</b>	<b>0.7804</b>	<b>1.09</b>	2.12	<b>1.93</b>	0.42	1.61	0.41	<b>1.65</b>	0.42	<b>1.98</b>	0.39

TABLE V  
THE PRE-TRAINED MODEL IS FINE-TUNED ON THE EUVP AND COMPARED WITH OTHER METHODS.

Datasets Methods	UIEB				EUVP				U45		UIQS		UIEB-60		EUVPUN	
	PSNR	SSIM	MSE	UIQM	PSNR	SSIM	MSE	UIQM	UIQM	UCIQE	UIQM	UCIQE	UIQM	UCIQE	UIQM	UCIQE
ACDC [54]	18.04	0.8212	1.21	1.26	14.74	0.659	2.53	1.22	1.03	1.02	0.73	0.34	1.03	0.38	1.14	0.41
Rank1 [4]	18.16	0.8443	1.21	1.63	15.16	0.7121	2.56	1.45	1.57	0.48	1.51	0.48	1.36	0.4	1.48	0.42
MMLE [55]	15.48	0.8242	2.34	1.48	14.84	0.6331	2.56	1.54	1.38	<b>0.49</b>	1.58	0.51	1.53	0.38	1.35	0.31
WaterNet [17]	15.42	0.6542	3.51	1.15	18.64	0.7215	2.15	1.78	1.39	0.31	1.41	0.33	1.31	0.35	1.45	0.33
Funie [15]	16.91	0.7291	1.70	1.66	21.99	0.8003	0.45	2.05	1.45	0.35	1.31	0.31	1.33	0.37	1.74	0.37
Ucolor [19]	16.54	0.7464	1.82	1.16	20.08	0.7815	1.54	1.98	1.77	0.41	1.61	0.39	0.88	0.32	1.66	0.38
Pre-trained	16.42	0.7354	1.54	<b>1.95</b>	15.57	0.7145	1.96	2.15	<b>1.88</b>	0.46	<b>1.72</b>	<b>0.55</b>	1.55	<b>0.45</b>	<b>1.87</b>	<b>0.46</b>
MetaUE	<b>18.67</b>	<b>0.8494</b>	<b>1.07</b>	1.71	<b>22.72</b>	<b>0.8623</b>	<b>0.39</b>	<b>2.19</b>	1.81	0.43	1.61	0.42	<b>1.68</b>	0.41	1.75	<b>0.46</b>

diversified distortions. As shown in Fig. 6, we evaluate the performance of the pre-trained model and fine-tuned model on typical underwater attenuation images including haze, blue, green, yellow, and turbid. We notice that our model-driven pre-trained model can well remove the background color of the underwater images, while the fine-tuned models can adapt to the characteristics of real underwater datasets. For quantitative comparison, Table III provides the averaged UCIQE and UIQM as well as the component scores of UCIQE and UIQM for the test images in Fig. 6, where the pre-trained model gives higher UCIQE and UIQM than the two fine-tuned models. It

confirms that the pre-trained model can provide enhancement results with better human visual perception. By analyzing the composition of these indicators, the pre-trained model is superior to other models in contrast and sharpness, but the color saturation needs to be further improved. Therefore, the model-based pre-trained model can remove the color degradation from various underwater scenarios. Since the fine-tuned models learned features from a specific real underwater dataset, they may have some performance degradation when the water environment changes.



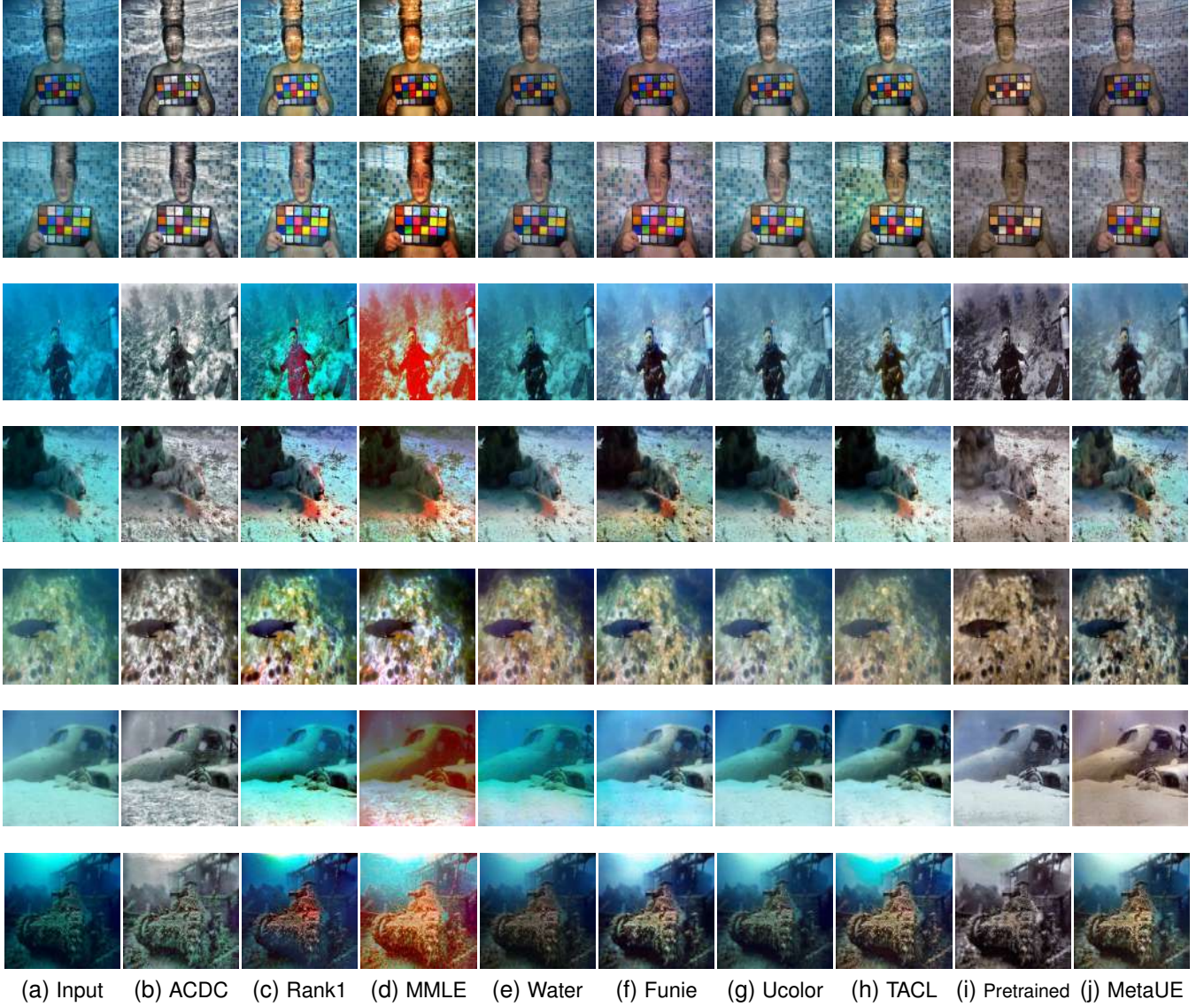


Fig. 8. The visual comparison among the enhancement methods on the unpaired underwater images, where our MetaUE is fine-tuned on the EUVP dataset.

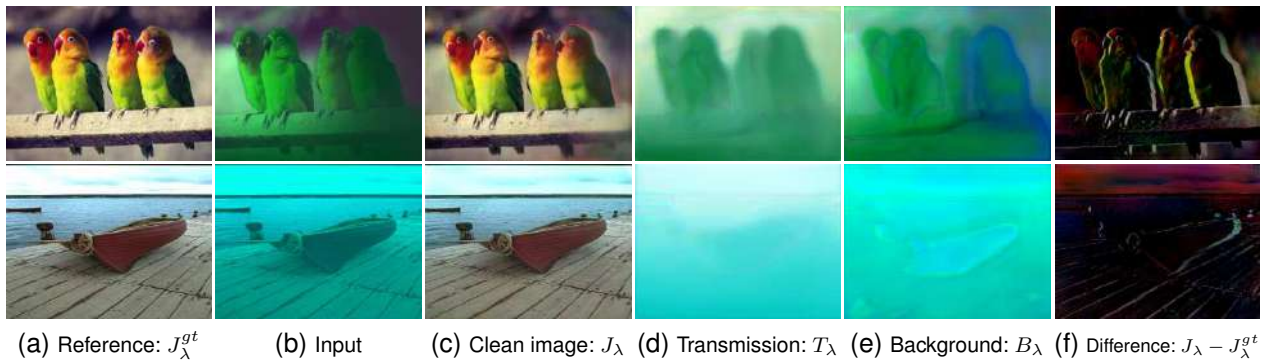


Fig. 9. The visual results of the clean image  $J_{\lambda}$ , transmission map  $T_{\lambda}$ , background light  $B_{\lambda}$  and the difference image, which is estimated as the difference between the reference image and the restored clean image.

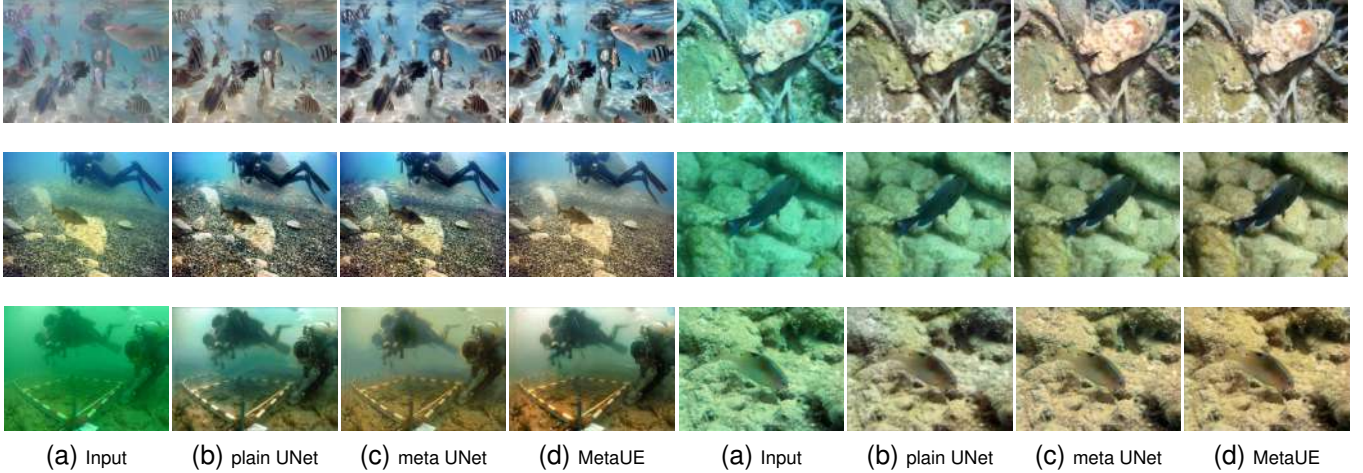


Fig. 10. The visual comparison on ablation study, where the first to fourth columns are the results on the UIEB dataset and the fifth to eighth columns are the results of the EUVP.

#### E. Comparison with other underwater image enhancement methods

We further evaluate our MetaUE model by comparing it with state-of-the-art underwater image enhancement methods. The quantitative comparison is displayed in Table IV, where our pre-trained model is fine-tuned on the UIEB dataset. As shown in Table IV, our MetaUE model outperforms other comparison methods in terms of PSNR, SSIM, and MSE, demonstrating the advantages of our model-based meta-learning strategy. The learning-based methods, Water-Net and Ucolor, perform poorly on other underwater image datasets except for UIEB, which convinces the advantages of building up the deep learning model based on the physical model for underwater images. For the unpaired datasets, the UIQM and UCIQE are the main evaluation metrics. It can be observed that both the pre-trained model and MetaUE provide better results than other comparison methods. Furthermore, it is shown that the pre-trained model presents its generalization ability on real underwater datasets. When we fine-tune the pre-trained model on UIEB, both PSNR and SSIM on the test dataset are significantly improved on UIEB, but it loses generalization on the EUVP dataset. Similar results can be observed in Table V, for which we fine-tuned the learning-based models on the EUVP dataset.

In Fig. 7, we exhibit the visual enhancement results among different methods on the degraded underwater images including scenes such as haze, green, blue, and low light. As can be observed, for the comparison methods, some work well on low-light images or yellow images, but none of them can provide satisfactory results on all applications. Our MetaUE presents good generalization ability on various underwater images, where the visual comparison is consistent with the quantitative results in Table IV. Furthermore, the pre-trained model is shown to be able to remove the water to obtain clear images, which provides excellent results for low-brightness images by restoring the objects in the shadow. And the MetaUE produces similar restoration results as the reference images provided by the UIEB dataset, which reflects the effect

of the fine-tuning process. More specifically, the first two examples in Fig. 8 reveal that the pre-trained model may lose color information due to the limited coverage of the synthetic dataset. On the other hand, the fine-tuned model can learn the color features from real underwater images. Similar results can be observed in Fig. 8, where the pre-trained model shows good generalization ability by removing the color degradation and repairing the insufficient illumination. The fine-tuning can effectively compensate for the underwater characteristic to make the enhanced images more realistic.

Some physical model-based methods [3], [4], [54], [60] assume the background light to be a constant function, which leads to inaccurate estimation. Actually, both background light and transmission map are smooth functions [3]. As shown in Fig. 9, we provide two examples to illustrate the solutions of the background light and transmission map in our model. As can be seen, the clean image, transmission map and background light are correctly estimated by our method. In particular, the residual image between the reference image  $J_{\lambda}^{gt}$  and the estimated clean image  $J_{\lambda}$  indicates that the pre-trained model can accurately remove the background color.

#### F. Ablation study

The ablation studies are presented to explore the effectiveness of the model-driven meta learning method for underwater image enhancement. More specifically, we build up three deep learning models for comparison, namely the Plain UNet, Meta UNet, and MetaUE. Note that the ‘Plain UNet’ is realized by random initialization and trained on UIEB and EUVP datasets, respectively. The ‘Meta UNet’ model is introduced to evaluate the effectiveness of meta-learning, where only the clean image  $N_J(I; \theta_J)$  was trained by meta-learning. As shown in Table VI, the pre-trained strategy plays a major role in improving underwater image quality, while the physical image model can guide the deep-learning model to recover clean images from the observations. Fig. 10 provides the visual comparison of the three methods, where the first to fourth columns are the results from the UIEB dataset and the fifth to eighth columns are the



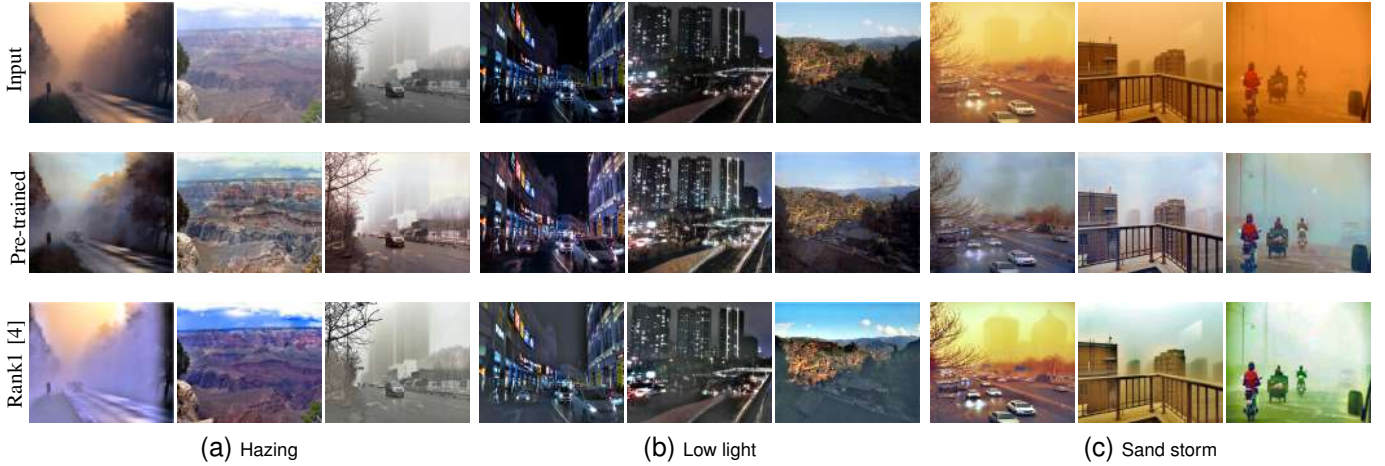


Fig. 11. There are the application results of the pre-trained model for low quality images, including haze, low light, and sand storm, where Rank1 represent the results of [4].

TABLE VI  
THE ABLATION STUDIES ON THE UIEB AND EUVP DATASETS, WHERE THREE MODELS ARE TRAINED, THE PLAIN UNET, META UNET AND MODEL-BASED META UNET.

Models	UIEB		EUVF	
	PSNR	SSIM	PSNR	SSIM
plain UNet	17.54	0.8125	18.76	0.8239
meta UNet	21.12	0.8054	21.24	0.8211
MetaUE	<b>22.01</b>	<b>0.8609</b>	<b>22.12</b>	<b>0.8322</b>

results from the EUVP. We can observe that the fine-tuned models based on the physical model give the best results of the three models, confirming the effectiveness of our model.

## VI. CONCLUSION AND DISCUSSION

In this paper, we presented a model-based deep-learning model to improve the visual qualities of underwater images, which was demonstrated with better interpretability and generalization ability. To be specific, we built up a multi-variable convolution model to learn the composition of the underwater image, including the clean image, transmission map, and background light. The meta-learning was used to first capture the shared prior knowledge of the degraded underwater images and then learn to adapt to a specific underwater environment by fine-tuning on the real underwater image dataset. Numerical experiments showed that our MetaUE outperformed several SOTA underwater image enhancement methods in both quality and robustness.

In addition to underwater images, there exist other degraded images that could be regarded as a superimposition of a clean image with environmental conditions such as sand dust, haze, and low light, etc. Various methods have been investigated to deal with such image restoration problems, e.g., low light enhancement [61], [62] and image dehazing [63], [64]. Indeed, these problems can be also modeled by the physical model (1) and recovered by our model-based deep learning model. As illustrated in Fig. 11, our pre-trained model works well on these challenging applications, which outperforms the recent

Rank1 prior method for scene recovery in [4]. Thus, it is beneficial to construct the network model based on a solid physical model, which can help to guide the network towards achieving desired outcomes. Although our MetaUE has demonstrated the improved generalization capabilities, there is still an obvious disparity in the performance of the fine-tuned model across different datasets; see Table IV and Table V. Our future work includes developing more efficient underwater image enhancement models in dealing with various real underwater images.

## REFERENCES

- [1] M. Jian, X. Liu, H. Luo, X. Lu, H. Yu, and J. Dong, "Underwater image processing and analysis: A review," *Signal Processing: Image Communication*, vol. 91, pp. 116–128, 2021.
- [2] D. Akkaynak, T. Treibitz, T. Shlesinger, Y. Loya, R. Tamir, and D. Iluz, "What is the space of attenuation coefficients in underwater computer vision?" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4931–4940.
- [3] D. Akkaynak and T. Treibitz, "Sea-Thru: A method for removing water from underwater images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1682–1691.
- [4] J. Liu, W. Liu, J. Sun, and T. Zeng, "Rank-one prior: Toward real-time scene recovery," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 14 802–14 810.
- [5] J. Liu, R. W. Liu, J. Sun, and T. Zeng, "Rank-one prior: Real-time scene recovery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [6] X. Ding, Z. Liang, Y. Wang, and X. Fu, "Depth-aware total variation regularization for underwater image dehazing," *Signal Processing: Image Communication*, vol. 98, pp. 116–128, 2021.
- [7] G. Hou, J. Li, G. Wang, Z. Pan, and X. Zhao, "Underwater image dehazing and denoising via curvature variation regularization," *Multimedia Tools and Applications*, vol. 79, no. 27, pp. 20 199–20 219, 2020.
- [8] Q. Jiao, M. Liu, P. Li, L. Dong, M. Hui, L. Kong, and Y. Zhao, "Underwater image restoration via non-convex non-smooth variation and thermal exchange optimization," *Journal of Marine Science and Engineering*, vol. 9, no. 6, pp. 570–581, 2021.
- [9] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "Water Gan: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 387–394, 2018.
- [10] C. Li, J. Guo, and C. Guo, "Emerging from water: Underwater image color correction based on weakly supervised color transfer," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 323–327, 2018.



- [11] X. Ye, Z. Li, B. Sun, Z. Wang, R. Xu, H. Li, and X. Fan, "Deep joint depth estimation and color correction from monocular underwater images based on unsupervised adaptation networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 11, pp. 3995–4008, 2019.
- [12] M. Yang, K. Hu, Y. Du, Z. Wei, Z. Sheng, and J. Hu, "Underwater image enhancement based on conditional generative adversarial network," *Signal Processing: Image Communication*, vol. 81, pp. 115–123, 2020.
- [13] R. Liu, Z. Jiang, S. Yang, and X. Fan, "Twin adversarial contrastive learning for underwater image enhancement and beyond," *IEEE Transactions on Image Processing*, vol. 31, pp. 4922–4936, 2022.
- [14] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, vol. 5, no. 3, 2018, pp. 7159–7165.
- [15] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.
- [16] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern Recognition*, vol. 98, pp. 107–118, 2020.
- [17] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," *IEEE Transactions on Image Processing*, vol. 29, no. 2, pp. 4376–4389, 2020.
- [18] X. Xue, Z. Hao, L. Ma, Y. Wang, and R. Liu, "Joint luminance and chrominance learning for underwater image enhancement," *IEEE Signal Processing Letters*, vol. 28, no. 1, pp. 818–822, 2021.
- [19] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color space embedding," *IEEE Transactions on Image Processing*, vol. 30, pp. 4985–5000, 2021.
- [20] K. Zhang, W. Zuo, and L. Zhang, "Deep plug-and-play super-resolution for arbitrary blur kernels," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1671–1681.
- [21] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6360–6376, 2021.
- [22] Y. Fang, H. Zhang, H. S. Wong, and T. Zeng, "A robust non-blind deblurring method using deep denoiser prior," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 735–744.
- [23] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.
- [24] P. Dhariwal and A. Nichol, "Diffusion models beat GANs on image synthesis," *Advances in Neural Information Processing Systems*, vol. 34, pp. 8780–8794, 2021.
- [25] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 10 684–10 695.
- [26] J. Wyatt, A. Leach, S. M. Schmon, and C. G. Willcocks, "Anoddpm: Anomaly detection with denoising diffusion probabilistic models using simplex noise," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 650–656.
- [27] H. Chung, E. S. Lee, and J. C. Ye, "Mr image denoising and super-resolution using regularized reverse diffusion," *IEEE Transactions on Medical Imaging*, 2022.
- [28] E. Hoogeboom, D. Nielsen, P. Jaini, P. Forré, and M. Welling, "Argmax flows and multinomial diffusion: Learning categorical distributions," *Advances in Neural Information Processing Systems*, vol. 34, pp. 12 454–12 465, 2021.
- [29] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap, "Meta-learning with memory-augmented neural networks," in *International Conference on Machine Learning*. PMLR, 2016, pp. 1842–1850.
- [30] Y.-X. Wang, D. Ramanan, and M. Hebert, "Meta-learning to detect rare objects," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9925–9934.
- [31] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5149–5169, 2021.
- [32] H. Zhu, L. Li, J. Wu, W. Dong, and G. Shi, "MetalQA: Deep meta-learning for no-reference image quality assessment," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 14 143–14 152.
- [33] H. Zhao, F. Ji, Q. Li, Q. Guan, S. Wang, and M. Wen, "Federated meta-learning enhanced acoustic radio cooperative framework for ocean of things," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 3, pp. 474–486, 2022.
- [34] W. Wang, L. Duan, Y. Wang, Q. En, J. Fan, and Z. Zhang, "Remember the difference: Cross-domain few-shot semantic segmentation via meta-memory transfer," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 7065–7074.
- [35] L. Hong, X. Wang, Z. Xiao, G. Zhang, and J. Liu, "Wsuie: Weakly supervised underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8237–8244, 2021.
- [36] I. H. Laradji, A. Saleh, P. Rodriguez, D. Nowrouzezahrai, M. R. Azghadi, and D. Vazquez, "Weakly supervised underwater fish segmentation using affinity lcfn," *Scientific reports*, vol. 11, no. 1, pp. 173–179, 2021.
- [37] A. Saleh, M. Sheaves, D. Jerry, and M. R. Azghadi, "Transformer-based self-supervised fish segmentation in underwater videos," *arXiv preprint arXiv:2206.05390*, 2022.
- [38] S. Cai, G. Li, and Y. Shan, "Underwater object detection using collaborative weakly supervision," *Computers and Electrical Engineering*, vol. 102, p. 108159, 2022.
- [39] P. M. Uplavikar, Z. Wu, and Z. Wang, "All-in-One underwater image enhancement using domain-adversarial learning," in *CVPR Workshops*, 2019, pp. 1–8.
- [40] Y. Liu, H. Xu, B. Zhang, K. Sun, J. Yang, B. Li, C. Li, and X. Quan, "Model-based underwater image simulation and learning-based underwater image enhancement method," *Information*, vol. 13, no. 4, pp. 2176–2289, 2022. [Online]. Available: <https://www.mdpi.com/2078-2489/13/4/187>
- [41] Z. Wang, L. Shen, M. Xu, M. Yu, K. Wang, and Y. Lin, "Domain adaptation for underwater image enhancement," *IEEE Transactions on Image Processing*, 2023.
- [42] Y. Wang, J. Guo, H. Gao, and H. Yue, "UIEC<sup>2</sup>-Net: CNN-based underwater image enhancement using two color space," *Signal Processing: Image Communication*, vol. 96, pp. 116–130, 2021.
- [43] S. Liu, H. Fan, S. Lin, Q. Wang, N. Ding, and Y. Tang, "Adaptive learning attention network for underwater image enhancement," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5326–5333, 2022.
- [44] W.-H. Zhou, D.-M. Zhu, M. Shi, Z.-X. Li, M. Duan, Z.-Q. Wang, G.-L. Zhao, and C.-D. Zheng, "Deep images enhancement for turbid underwater images based on unsupervised learning," *Computers and Electronics in Agriculture*, vol. 202, pp. 107–122, 2022.
- [45] S. Chai, Z. Fu, Y. Huang, X. Tu, and X. Ding, "Unsupervised and untrained underwater image restoration based on physical image formation model," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 2774–2778.
- [46] Z. Guo, D. Guo, Z. Gu, H. Zheng, B. Zheng, and G. Wang, "Unsupervised underwater image clearness via transformer," in *OCEANS 2022-Chennai*. IEEE, 2022, pp. 1–4.
- [47] Q. Jiang, Y. Zhang, F. Bao, X. Zhao, C. Zhang, and P. Liu, "Two-step domain adaptation for underwater image enhancement," *Pattern Recognition*, vol. 122, pp. 108–124, 2022.
- [48] A. R. Espinosa, D. McIntosh, and A. B. Albu, "An efficient approach for underwater image improvement: Deblurring, dehazing, and color correction," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 206–215.
- [49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [50] N. Silberman and R. Fergus, "Indoor scene segmentation using a structured light sensor," in *2011 IEEE International Conference on Computer Vision Workshops (ICCV workshops)*. IEEE, 2011, pp. 601–608.
- [51] X. Zhao, T. Jin, and S. Qu, "Deriving inherent optical properties from background color and underwater image enhancement," *Ocean Engineering*, vol. 94, pp. 163–172, 2015.
- [52] C. D. Mobley and C. D. Mobley, *Light and water: radiative transfer in natural waters*. Academic Press, 1994.
- [53] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.
- [54] W. Zhang, Y. Wang, and C. Li, "Underwater image enhancement by attenuated color channel correction and detail preserved contrast enhancement," *IEEE Journal of Oceanic Engineering*, 2022.

- [55] W. Zhang, P. Zhuang, H.-H. Sun, G. Li, S. Kwong, and C. Li, "Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement," *IEEE Transactions on Image Processing*, vol. 31, pp. 3997–4010, 2022.
- [56] H. Li, J. Li, and W. Wang, "A fusion adversarial underwater image enhancement network with a public test dataset," *arXiv preprint arXiv:1906.06819*, 2019.
- [57] A. Duarte, F. Codevilla, J. D. O. Gaya, and S. S. Botelho, "A dataset to evaluate underwater image restoration methods," in *OCEANS 2016-Shanghai*. IEEE, 2016, pp. 1–6.
- [58] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 6062–6071, 2015.
- [59] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 2015.
- [60] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *Journal of Visual Communication and Image Representation*, vol. 26, pp. 132–145, 2015.
- [61] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5637–5646.
- [62] M. Wang, Y. Huang, J. Xiong, and W. Xie, "Low-light images in-the-wild: A novel visibility perception-guided blind quality indicator," *IEEE Transactions on Industrial Informatics*, 2022.
- [63] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, and D. Tao, "Self-augmented unpaired image dehazing via density and depth decomposition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 2037–2046.
- [64] R. W. Liu, Y. Guo, Y. Lu, K. T. Chui, and B. B. Gupta, "Deep network-enabled haze visibility enhancement for visual iot-driven intelligent transportation systems," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1581–1591, 2022.