Unpaired Overwater Image Defogging Using Inverted Dark Channel Prior-Guided Cycle-Consistent Generative Adversarial Network

Yaozong Mo[a], Chaofeng Li[a,*], Tuxin Guan[b], Qiuping Jiang[b], Wenqi Ren[c], Wenwu Wang[d], Xiao-jun Wu[e]

[a]*Institute of Logistics Science and Engineering, Shanghai Maritime University, 201306, Shanghai, China*
[b]*Faculty of Information Science and Engineering, Ningbo University, 315211, Ningbo, China*
[c]*School of Cyber Science and Technology, Sun Yat-sen University, 518000, Shenzhen, China*
[d]*Center for Vision Speech and Signal Processing, Department of Electrical and Electronic Engineering, University of Surrey, Surrey GU2 7XH, Surrey, U.K.*
[e]*School of Artificial Intelligence and Computer Science, Jiangnan University, 214122, Wuxi, China*

Abstract

Existing image defogging methods have made significant advancements. However, these approaches are primarily optimized for land scenes, resulting in suboptimal performance when applied to overwater images due to the distinct characteristics of overwater scenes. In this paper, we propose an inverted dark channel Prior-Guided Cycle-consistent Generative Adversarial Network (PG-CycleGAN) for overwater image defogging. Specifically, an inverted dark channel prior map is designed to suppress the sky and highlight objects over the water. Building on this prior map, we develop a prior encoder to extract object-related features. Additionally, we propose a Prior-Guided Residual Block (PGRB) and a Prior-Guided TriUpsample (PGTU) module, which effectively integrate the extracted prior features for both feature encoding and upsampling. This integrated approach enhances the network's ability to accurately restore overwater objects, leading to improved defogging performance. Furthermore, we develop a prior map-guided GAN loss and a prior map-guided cycle-consistency loss, which guide the network to recover objects with greater fidelity while minimizing unnecessary restoration of the sky region. Through extensive experimental comparisons, our method demonstrates superior performance over existing state-of-the-art

★This work was supported by the National Natural Science Foundation of China under Grant 62176150.
*Corresponding author: wxlichaofeng@126.com (Chaofeng Li)

approaches in terms of qualitative analysis, quantitative metrics, and improvements in object detection.

Keywords: Overwater image defogging, generative adversarial networks, unpaired defogging dataset.

---

1. Introduction

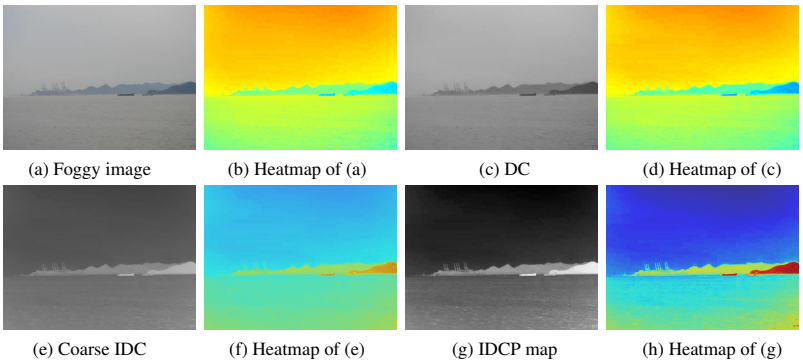
Fog can considerably degrade images taken on lakes, rivers, and seas, leading to adverse effects on downstream computer vision tasks like object detection and tracking. Although many promising image defogging algorithms have been developed, they rarely focus on the scene of overwater and are limited in defogging overwater image. Therefore, it is necessary and valuable to conduct research on the defogging of overwater images.

Based on the Atmospheric Scattering Model (ASM) [1], a foggy image can be formulated as :

$$I(x_i) = J(x_i)t(x_i) + A(1 - t(x_i)). \tag{1}$$

Here, x_i denotes the pixel position, $I(x_i)$ is degraded foggy image, and $J(x_i)$ is fog-free image. A and $t(x_i)$ are the air light and the transmission. There are two main types of defogging approaches: prior-based and learning-based. The prior-based algorithms, such as Non-Local Dehazing (NLD) [2] and Dark Channel Prior (DCP) [3], recover clean images by using physical priors through the model (1). However, these priors are not always reliable. For instance, DCP [3] suggest that a minimum of one channel in the RGB space tends to approach zero in a fog-free image. As a result, DCP may have limitations in addressing scene objects that closely resemble the atmospheric light.

Unlike prior-based methods which perform well only for certain scenarios, the learning-based approaches such as All-in-One Dehazing Network (AODNet) [4] and Detail-Enhanced Attention Network (DEANet) [5] exhibit better generalization ability. However, most of them rely on supervised learning and frequently encounter challenges when handling real-world foggy images due to the synthetic training samples. To address this issue, researchers have focused on developing semi-supervised and unsupervised image defogging algorithms, such as Principled Synthetic-to-real Dehazing



(a) Foggy image (b) Heatmap of (a) (c) DC (d) Heatmap of (c)

(e) Coarse IDC (f) Heatmap of (e) (g) IDCP map (h) Heatmap of (g)

Figure 1: Comparison between the Dark Channel (DC), coarse Inverted Dark Channel (IDC), and the Inverted Dark Channel Prior (IDCP) map.

(PSD) framework [6], to mitigate the domain shift problem. While these methods have shown promising performance, little attention has been given to the challenge of overwater image defogging.

Different from the images captured on land, images captured on water usually contain large sky and water areas. Although objects such as boats on the water often occupy only a small area, their recovery is more important than the recovery of sky, as it directly affects subsequent downstream tasks such as object detection and tracking. Considering this feature, we propose an Inverted Dark Channel Prior (IDCP) map, which facilitates the network to focus on the recovery of objects. Fig. 1 compares the DC and the prior map. In dark channels, the sky and water regions tend to be brighter than the objects on the water because their pixel values are generally close to the atmospheric light. Based on this feature, the objects will be highlighted when the dark channel is inverted, as shown in Fig. 1 (e). To emphasize this feature, we utilize the min-max normalization to enhance the contrast, as illustrated in Fig. 1 (g), and we call it the inverted dark channel prior map.

To integrate the prior knowledge into the defogging network, we propose a prior encoder designed to extract critical features from overwater images, such as ships. Subsequently, we design a Prior-Guided Residual Block (PGRB) and a Prior-Guided TriUpsample Module (PGTU). Specifically, the PGRB compresses prior information into a 1-D vector, which is used to refine the residual block's intermediate feature maps

through channel attention, enabling the block to focus on key features. Meanwhile, the PGTU integrates prior information with three upsampling methods, motivated by the following observations: transposed convolution tends to generate checkerboard artifacts, while bilinear interpolation produces smooth results but limits the network's representational power due to its fixed upscaling procedure. PixelShuffle [7] addresses the overlap issue with sub-pixel convolution, but it is prone to introducing artifacts, as observed in our experiments. To overcome these limitations, the PGTU module employs a channel-split strategy to combine three distinct upsampling branches, followed by the use of channel shuffle and group convolution to effectively merge the upsampled features. The prior information is then leveraged to guide the upsampling process, enhancing ship-related features and ensuring that these features are better preserved and more clearly represented in the final output.

Furthermore, we propose a prior map-guided cycle-consistency loss and a prior map-guided GAN loss, both of which are weighted by the inverted dark channel prior map during the training phase. This weighting mechanism imposes a higher penalty on overwater objects, such as ships, compared to the water and sky regions. By emphasizing the importance of overwater objects in the loss functions, we ensure that the model prioritizes their restoration, while mitigating the influence of less relevant background regions, such as water and sky, thereby enhancing overall defogging performance.

In brief, this work presents the following contributions:

- An overwater scene-oriented unsupervised image defogging framework with an inverted dark channel prior map is proposed that emphasizes the objects on the water, based on the characteristic of the overwater images which often contain large sky and water areas.
- A prior-guided residual block and a prior-guided TriUpsample module are designed to effectively integrate key object features from the prior map, thereby improving the network's representation of overwater objects.
- We further propose a prior map-guided GAN loss and a prior map-guided cycle-consistency loss to facilitate the network in recovering the objects on the water.

These two loss functions successfully address and preserve object-related information in defogged images.

- Through extensive experimental comparisons, our method demonstrates superior performance over existing state-of-the-art approaches in terms of qualitative analysis, quantitative metrics, and object detection improvements.

2. Related Work

In this section, we succinctly overview both prior-based and learning-based methods relevant to single-image defogging.

2.1. Prior-Based Methods

Prior-based methods leveraging natural image statistics for single-image defogging. Fattal *et al.* [8] formulated an improved image degradation model that incorporates a shading component and surface reflectance coefficients. He *et al.* [3] suggest that a minimum of one channel in the RGB space tends to approach zero in a fog-free image. Zhu *et al.* [9] developed a linear model to obtain depth map through supervised learning. Fattal *et al.* [10] explored color-lines which assert that pixels within local image patches often demonstrate a one-dimensional distribution. Berman *et al.* [2] assume that a limited number of distinct colors can approximate those present in a clean image. Liu *et al.* [11] developed a new approach that reformulates the image defogging problem to increase local visibility and global contrast. Ling *et al.* [12] proposed a defogging method based on the saturation line prior, which exploits the linear relationship between saturation and brightness components of pixels with the same surface reflectance in fog-free images. Although these methods show effectiveness in image defogging, they may fail when dealing with certain scenarios. For example, the DCP [3] often encounters challenges with the sky or white buildings. Additionally, using the haze-line prior [2] for dense fog can cause color distortion. In comparison, we optimize the network by building the prior into the loss functions, and our data-driven model offers good generalization ability.

2.2. Learning-Based Methods

Data-driven algorithms based on convolutional neural networks (CNNs) for image defogging have become popular in recent years. For instance, Ren *et al.* [13] comprising both a coarse and a fine network component dedicated to predicting the transmission map. Li *et al.* [4] reconstruct the ASM by combining the transmission map and air light into a single variable. Then, they developed a CNN to estimate this variable. Instead of using a single input, the multi-source input fusion strategy [14] is considered an effective approach. For example, Ren *et al.* [15] developed a fusion-based network to recover the clean image, they predict a confidence map by integrating three extracted channels from the input image. Qin *et al.* [16] also adopted a fusion strategy, where they applied channel attention and pixel attention after the high-level feature for giving more weight to important features. Nevertheless, the utilization of a multi-branch network architecture escalates network complexity and elongates the training process. Chen *et al.* [5] proposed a detail-enhanced attention network that improves dehazing by enhancing feature representation. It integrates detail-enhanced convolution for efficient feature extraction and content-guided attention to emphasize critical information. Cui *et al.* [17] integrate spatial-domain processing, spectral-domain analysis, and a dual-domain interaction mechanism to effectively capture both local details and global context. Lyu *et al.* [18] tackle challenging non-homogeneous scenes by leveraging image priors from multiple color spaces. They design parallel sub-networks and a feature fusion module to effectively restore color and details in foggy images.

Despite the favorable results of supervised learning-based approaches, their application to real-world scenes is limited because of the inherent domain shift caused by synthetic training sets [19]. Therefore, recent efforts have focused on integrating semi-supervised and unsupervised approaches to address real-world challenges [20]. These methods improve model robustness and enhance downstream tasks, such as object detection [21], by leveraging unlabeled data and adapting to diverse scene conditions. For example, Engin *et al.* [22] introduced an enhanced CycleGAN [23] which adopt unpaired real-world images as training set. Li *et al.* [24] integrated a supervised subnetwork with an unsupervised subnetwork. Shao *et al.* [25] incorporate defogging blocks alongside image translation blocks for handling real-world scenes. Zhao *et al.*

[26] utilized an unsupervised CNN to refine the defogged image of DCP [3]. Sun *et al.* [27] explored an unsupervised multi-branch network with high-frequency enhancement for image dehazing. It improves feature consistency, stabilizes learning, and enhances edge and texture restoration in dense fog. Recently, Liang *et al.* [28] developed a self-supervised image dehazing framework guided by depth estimation to enhance fog removal.

The above mentioned approaches offer desirable defogging results on real-world images, but little literature focuses on the image defogging of overwater scenes. Zheng *et al.* [29] proposed an enhanced CycleGAN to tackle the overwater image defogging problem. Nevertheless, this method simply replaces the training set of land foggy images as used in other methods with overwater images. Its defogging performance is limited as this method does not leverage the properties of the overwater scenes. To this end, our method focuses on the characteristics of overwater foggy images, and improves the defogging performance and visual quality of the defogged images.

3. Method

In this section, the proposed PG-CycleGAN will be presented in detail. Fig. 2 illustrates the architecture of the PG-CycleGAN, it consists of two cycles: foggy-clean-foggy and clean-foggy-clean. In the foggy-clean-foggy cycle, the defogging network G receives a foggy image x and corresponding prior map I_P^{foggy} as inputs, learning the defogged and prior information separately. These features are then integrated through the PGTU module and PGRB, and subsequently decoded to produce a clean image $G(x)$. This clean image is then passed through the fogging network F to generate a reconstructed foggy image $F(G(x))$. Finally, the prior map I_P^{foggy} is employed to compute the prior map-guided cycle-consistency loss L_{pgcyc}, ensuring the accurate restoration of overwater objects, as well as the prior map-guided GAN loss L_{pgG}, which guarantees the authenticity of these objects. The clean-foggy-clean cycle is similar to the foggy-clean-foggy cycle.

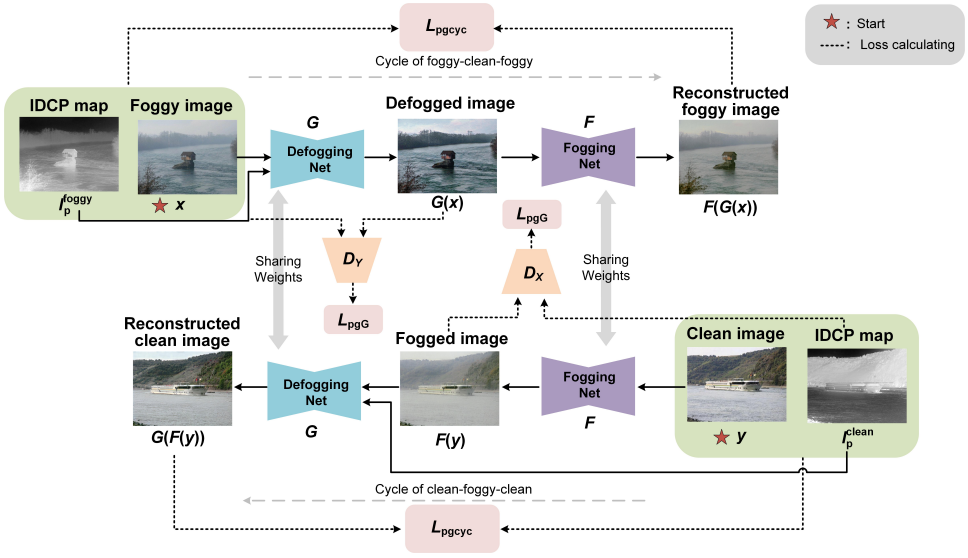


Figure 2: The details of the PG-CycleGAN. It has two generators, a defogging network G and a fogging network F, as well as two discriminators, D_X and D_Y.

3.1. Inverted Dark Channel Prior

He *et al.* [3] argued that the DCP loses its validity in scenarios where the objects possess characteristics that are similar to the air light. This limitation is evident in Fig. 3 (b), where the pixel values of water and sky are not close to 0.

A characteristic of the overwater scenes is that the water and sky usually occupy the majority of the area. In dark channels, they tend to be brighter than objects on the water because their pixel values are similar to the air light. Based on this feature, the objects will be highlighted when the DC of the overwater image is inverted, as depicted in Fig. 3 (c). To emphasize this feature, we try to enhance the contrast of the inverted dark channel. Fig. 3 shows the difference between the DC, coarse IDC, coarse IDC after gamma correction, coarse IDC after histogram equalization (HE), and coarse IDC after min-max normalization (i.e., the inverted dark channel prior map). It can be seen that though the contrast is enhanced, the gamma correction may lower the pixel value of the objects, and the histogram equalization may amplify background noise. In contrast, the min-max normalization maintains the highlight of the object while reducing noise. Fig. 4 is the intensity histogram over 4,531 inverted dark channels of an overwater image

(a) Input (b) DC (c) coarse IDC (d) Gamma (e) HE (f) IDCP map

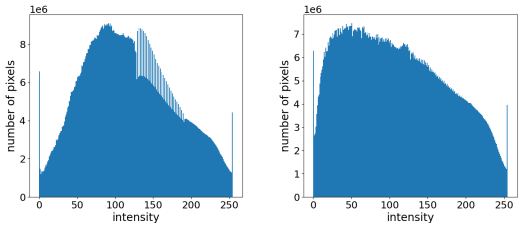
Figure 3: Difference on overwater images between the DC, coarse IDC, coarse IDC after gamma correction (gamma=3), coarse IDC after histogram equalization, and the IDCP map. Here the dark channels are without performing minimum filtering.

dataset [29] (details described in Sec. 4.1; the images are rescaled to 256×256 for computational efficiency). It can be seen that the min-max normalization stretches the pixel values in the middle region while keeping the sky regions darker and the objects brighter, this statistic provides strong support to our proposed inverted dark channel prior.

Specifically, given a foggy image I whose pixel value is in the range of [0, 1], its corresponding dark channel is:

$$I_{dc} = \min_{y_i \in N(x_i)} \left(\min_{c \in \{r,g,b\}} I^c(y_i) \right). \tag{2}$$

Here, x_i and y_i denote the pixel coordinates, and I^c is c-th color channel. $N(x_i)$ refers to the local neighborhood centered around x_i. To obtain the inverted dark channel



(a) unnormalized (b) min-max normalized

Figure 4: Intensity histogram of 4,531 inverted dark channels. (a) histogram of unnormalized inverted dark channels. (b) histogram of min-max normalized inverted dark channels.

prior map, we first remove the minimum filtering $\min_{y_i \in N(x_i)}(\cdot)$ from the procedure for dark channel calculation, and invert the pixel value as follows:

$$I_{idc} = 1 - \min_{c \in \{r,g,b\}} I^c. \tag{3}$$

Here, I_{idc} represents the coarse inverted dark channel. The reasons for removing the minimum filter are two. One is that it reduces the required computational resources, as performing the sliding window algorithm is time-consuming. The other is that the minimum filter may lead to unclear boundaries of the objects in the defogging results because it exploits all the surrounding pixels.

Then, the min-max normalization operation is applied to I_{idc} to obtain the inverted dark channel prior map.

$$I_p = \frac{I_{idc} - min(I_{idc})}{max(I_{idc}) - min(I_{idc})}. \tag{4}$$

After getting the inverted dark channel prior map, we further define the prior map-guided cycle-consistency loss:

$$L_{pgcyc} = \|(F(G(x)) - x)I_p^{foggy}\|_1 + \|(G(F(y)) - y)I_p^{clean}\|_1, \tag{5}$$

and the prior map-guided GAN loss:

$$L_{pgG} = \mathbb{E}_{y \sim p_{data}(y)}[(D_Y(y) - 1)^2] + \mathbb{E}_{x \sim p_{data}(x)}[D_Y(G(x)I_p^{foggy})^2], \tag{6}$$

where x and y are foggy and clean image, I_p^{foggy} and I_p^{clean} are the corresponding prior maps, and $\|\cdot\|_1$ is an L_1 norm. In L_{pgcyc} and L_{pgG}, the prior map increases the loss

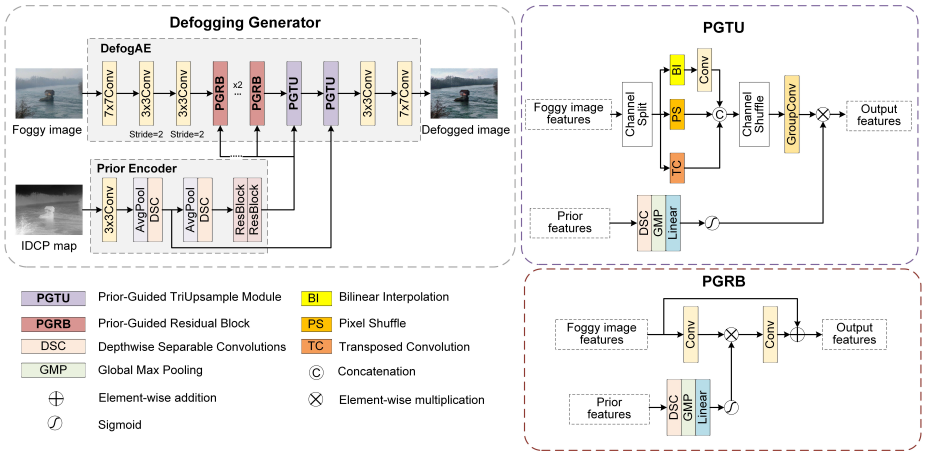


Figure 5: The architecture of the defogging network.

associated with objects, ensuring that the network prioritizes the recovery of object features. Meanwhile, it reduces the loss related to the sky, guiding the network to focus less on relevant background areas. By modulating these losses, the prior map effectively guides the network to recover objects with greater fidelity while minimizing unnecessary restoration of the sky region. This targeted guidance enhances the overall performance of the network, particularly in preserving key features in complex scenes.

3.2. PG-CycleGAN

1) Overall Structure: The details of the defogging generator G are depicted in Fig. 5. It includes a Defogging Auto-Encoder (DefogAE) and a prior encoder. Specifically, the prior encoder first employs a 3×3 convolution for initial feature extraction from the prior map, followed by two downsampling modules, each containing an average pooling layer and a depthwise separable convolution [30]. Two residual blocks [31] are then cascaded to enhance the learning capacity of the prior encoder and generate the prior features.

In contrast to the lightweight prior encoder, the DefogAE demonstrates enhanced learning capabilities. It consists of a 7×7 convolution followed by two 3×3 convolutions with a stride of 2, which are used for the initial feature extraction and downsampling of the foggy input image. Four PGRB are then employed to integrate prior

information from the prior encoder, enhancing feature representation and ensuring that the network focuses more on relevant regions of the image. The PGTU module is subsequently used for upsampling during feature decoding, incorporating channel attention from the prior encoder to selectively amplify features related to the objects of interest, while suppressing irrelevant background regions. The integration of prior information in both feature encoding and upsampling effectively refines the network's defogging results, particularly in challenging overwater scenes. Finally, the defogged image is generated through a 3×3 convolution, followed by a 7×7 convolution.

2) Prior-Guided TriUpsample Module: The PGTU module is a combination of three parallel upscaling branches: transposed convolution, bilinear interpolation, and PixelShuffle [7], with their output filter banks concatenated into a single output. As shown in Fig. 5, for the foggy image features, we first apply channel split operation to obtain three feature map groups. Subsequently, three different upsampling methods are employed to upsample these feature maps:

$$X^c_{TPB} = [TC(F_0^{c/3}), PS(F_1^{c/3}), C_{1\times1}(BI(F_2^{c/3}))]. \tag{7}$$

Here, TC, PS, and BI represents transposed convolution, PixelShuffle [7], and bilinear interpolation, respectively. $C_{1\times1}(\cdot)$ is convolution layer with 1×1 kernel size, and $[\cdot]$ denotes feature concatenation operation. $F_i^{c/3}(i = 0, 1, 2)$ denotes each channel group of foggy image features. Additionally, X^c_{TPB} denotes the combination of the three upsampled feature groups.

The channel split operation is conducive to reducing computational cost. Nevertheless, such an approach restricts the information exchange among channel groups. To overcome this constraint, a combination of channel shuffle and grouped convolution is adopted, thereby augmenting the interaction of diverse upsampled feature groups:

$$X^c = GC_{3\times3}(CS(X^c_{TPB})), \tag{8}$$

where $CS(\cdot)$ denotes a channel shuffle operation, $GC_{3\times3}(\cdot)$ denotes a group convolution layer with the kernel size of 3×3, and X^c denotes the obtained features.

Finally, the upsampled features from the foggy image X^c integrate the channel at-

tention derived from the prior encoder:

$$X_{refined} = \sigma(L(GMP(DSC_{3\times3}(X^c_{prior})))) \otimes X^c, \tag{9}$$

where $L(\cdot)$ denotes linear layer, $GMP(\cdot)$ is global max pooling layer, $DSC_{3\times3}(\cdot)$ represents depthwise separable convolution [30] with a 3×3 kernel size in the depthwise convolution, σ denotes the sigmoid function, $\otimes$ denotes elemtent-wise multiplication, and $X_{refined}$ is the final refined features. Depthwise separable convolution optimizes feature extraction by decomposing standard convolution into lightweight depthwise and pointwise operations. Subsequently, the global max pooling layer extracts the most significant features and compresses them into a 1-dimensional vector. Finally, the sigmoid function generates attention weights, directing the model to focus on the most informative regions within each channel.

3) Prior-Guided Residual Block: To facilitate the flow of information across DefogAE and prior encoder, we design a prior-guided residual block to assist DefogAE in capturing object-related information, as illustrated in Fig. 5. Let X_{foggy} and X_{prior} denote the features from the foggy image and the inverted dark channel prior map respectively, we first calculating the channel attention using X_{prior}:

$$X_{CA} = \sigma(L(GMP(DSC_{3\times3}(X_{prior})))), \tag{10}$$

where X_{CA} denotes a set of learnable 1-D weights that represent the relative importance of each channel, allowing the model to focus more on informative channels and suppress less relevant ones.

Then, we apply the computed channel-level weights to the feature maps at the middle of the residual block:

$$X_{res} = C_{3\times3}(C_{3\times3}(X_{foggy}) \otimes X_{CA}) \oplus X_{foggy}, \tag{11}$$

where $C_{3\times3}(\cdot)$ represents the convolution operation with a 3×3 kernel size, $\oplus$ denotes element-wise addition and X_{res} denotes the output features of the PGRB. By incorporating prior knowledge, the residual block focuses on key aspects of the image, such as ships or floating objects on the water surface, which are crucial for the overwater image defogging task. Additionally, integrating prior knowledge enhances the model's ability to leverage domain-specific information, improving its robustness and generalization.

4) Discriminator: Unlike the patch-based discriminator used in [23], we employ a pixel-level discriminator in our PG-CycleGAN, as illustrated in Fig. 6. This choice is motivated by the fact that both the prior map and the pixel-level discriminator generate output arrays that match the size of the input image, allowing for finer-grained feature matching. In contrast, a patch-based discriminator produces smaller output arrays, which may not align as effectively with the input image.

Specifically, our discriminator consists of three convolutional layers, each with a kernel size of 1×1 and a stride of 1. The number of filters increases from 64 in the first layer to 128 in the second layer, and finally reduces to 1 in the last layer. The prior map is applied to the discriminator's output scores as a spatial weighting mask, which emphasizes important regions during loss computation. As a result, misclassifications in less relevant areas such as the sky incur relatively small penalties, while misclassifications in overwater object regions lead to higher loss values. This encourages the network to focus on refining object-related details.

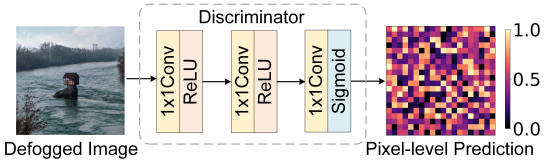


Figure 6: The architecture of the discriminator.

3.3. Loss Function

We employ three loss functions for optimizing the PG-CycleGAN, i.e., the prior map-guided cycle-consistency loss L_{pgcyc}, the prior map-guided GAN loss L_{pgG}, and the perceptual loss L_{VGG}.

Prior map-guided cycle-consistency loss and **prior map-guided GAN loss** are defined in Equations (5) and (6), they give higher penalties to objects on the water and lower penalties to the sky.

Perceptual loss is also used to enhance the textural information of the recovery image, which is based on VGG16 [32] and defined as:

$$L_{VGG} = \|\phi(x) - \phi(F(G(x)))\|_2^2 + \|\phi(y) - \phi(G(F(y)))\|_2^2. \tag{12}$$

Here, ϕ denotes the features extracted from the second and fifth pooling layers of the VGG16 network.

Total loss function. Integrating all the aforementioned losses, the comprehensive loss function is formulated as:

$$L = \lambda_1 L_{pgcyc} + \lambda_2 L_{VGG} + \lambda_3 L_{pgG}, \tag{13}$$

where λ_i ($i = 1, 2, 3$) represents the weight. In our empirical experiments, it is found that the network is sensitive to L_{pgG} which can significantly improve the defogging performance after fine-tuning. Thus, we set λ_1 and λ_2 to 1 which can stabilize the training, and set λ_3 to 2 which facilitates more thorough defogging by the network.

4. Experiments and Discussions

To evaluate the performance and generalization ability of the PG-CycleGAN, we conducted qualitative comparisons, non-reference image quality assessment (NR-IQA), and comparisons of improvements in object detection performance. The evaluations were conducted against several state-of-the-art methods, including the prior-based method DCP [3], supervised methods MSBDN [33], 4KDehazing [34], C2PNet [35], and DEANet [5], as well as semi-supervised and unsupervised methods such as PSD [6], SLAD [36], UMENet [27], and DGD [28].

4.1. Datasets

In this work, we use real-world foggy images for training instead of the widely used synthetic images. The dataset, OverwaterHaze, was proposed by Zheng *et al.* [29]. The training set consists of 2,090 unpaired foggy images and 2,441 clean images, while the test set contains 188 foggy images. All images are captured in real-world overwater scenes with a resolution of 640×480 pixels.

Additionally, the Singapore Maritime Dataset (SMD) [37] is used to evaluate the maritime object detection performance before and after defogging. The dataset comprises 692 foggy images and includes a single category, “boat".

4.2. Implemented Details

All training samples are rescaled to 512×512 and subjected to random cropping to 256×256 for data augmentation. We employ the ADAM optimizer [38] with a learning rate of 0.0001. Training follows a cosine annealing schedule, spanning a total of 150 epochs. The process is accelerated using an Nvidia GeForce RTX 3090 GPU, taking approximately 30 hours.

4.3. Performance Evaluation

A set of experiments is conducted to evaluate the performance of the proposed PG-CycleGAN, including visual quality comparison, non-reference image quality assessment, and object detection performance evaluation.

1) Qualitative comparison: Fig. 7 presents the defogging results obtained using different approaches on overwater images. It can be observed that the defogging results of DCP [3] are generally darker, and due to the inaccurate estimation of the transmission map, artifacts are presented in the sky region. Supervised learning-based methods, such as C2PNet [35] and DEANet [5], suffer from poor generalization performance due to the domain shift between synthetic training data and real foggy images, which limits their effectiveness in removing fog from real-world images. Although unsupervised and self-supervised defogging methods generally yield better results, the defogged images produced by PSD [6] exhibit noticeable color shifts, while UMENet [27] also introduces artifacts in the sky region.

In contrast, the proposed PG-CycleGAN achieves superior defogging results. By incorporating prior knowledge, the network is better able to focus on restoring overwater objects, as highlighted in the zoomed-in region in Fig. 7. Compared to other methods, our approach leaves less residual fog, allowing for clearer restoration of ship details on the water. Moreover, the defogged images appear more realistic and align better with human visual perception.

2) Non-reference image quality assessment: Non-reference image quality assessment metrics [39][40] have been effectively utilized to evaluate the quality of real-world images. In this paper, we adopt FADE [41], HazDesNet [42], and CV-CNN [43] to quantitatively assess the defogging performance. These metrics have been developed

(a) Foggy (b) DCP (c) MSBDN (d) 4KDehazing (e) C2PNet (f) DEANet (g) PSD (h) SLAD (i) UMENet (j) DGD (k) Ours

Figure 7: Qualitative evaluation of defogging performance on real-world overwater images

for non-reference evaluation and defogging quality assessment. FADE and HazDesNet, in particular, are capable of assessing defogging performance by predicting the density of the fog. CV-CNN assesses the perceptual quality of defogged images through the complex-valued responses. Lower FADE and HazDesNet values signify less fog in the defogged images, while higher CV-CNN values indicate better defogging performance.

Table 1 presents a quantitative comparison on the OverwaterHaze dataset, where bold numbers with superscripts denote the rankings. It can be observed that our method achieves the best results across all three metrics, with a 0.24 lead over the second-best

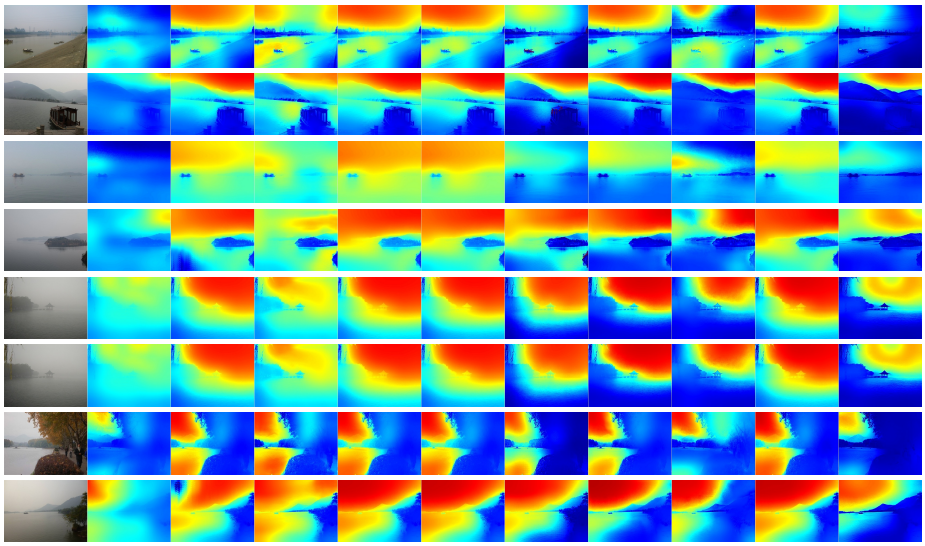
Table 1: Quantitative evaluation of defogging performance on the OverwaterHaze dataset.

Method	**Publication**	**Type**	**FADE↓**	**HazDesNet↓**	**CV-CNN↑**
DCP	TPAMI'10	Prior	1.08	0.34	0.59
MSBDN	CVPR'20	Supervised	2.28	0.53	0.61
4KDehazing	CVPR'21		1.54	0.44	0.65
C2PNet	CVPR'23		3.01	0.58	0.63
DEANet	TIP'24		2.90	0.57	0.70
PSD	CVPR'21	Semi/Unsupervised	1.24	0.43	0.65
SLAD	IJCAI'22		1.07	0.40	0.70
UMENet	PR'24		**0.97**[2]	**0.33**[2]	**0.73**[2]
DGD	PR'25		1.92	0.51	0.72
Ours		Unsupervised	**0.73**[1]	**0.30**[1]	**0.77**[1]

method on FADE, indicating that our method leaves the least residual fog in the defogged images. Additionally, the results on CV-CNN show that our defogging method generates images with higher quality, aligning better with human visual perception.

To provide a more intuitive comparison of fog residue in the defogged images, we present the fog density map comparison (generated by HazDesNet) in Fig. 8. In this map, colors approaching deep red indicate higher levels of residual fog, while colors approaching deep blue represent lower levels of fog residue. As observed, the defogged images produced by our method predominantly exhibit deep blue, signifying minimal residual fog. In contrast to other methods, PG-CycleGAN places greater emphasis on defogging overwater objects. Considering all the performance evaluations, it can be concluded that our method leaves less fog in the images and produces better visual quality.

3) Evaluation of object detection improvement: Object detection performance can serve as an effective indicator of image enhancement quality [44]. Therefore, we employ YOLOv10b [45] to validate the influence of defogging process on object detection performance, where the confidence and Non-Maximum suppression Intersection over Union (NMS IoU) threshold are set to 0.25 and 0.45, which are the same as those in the original code (available on the project website: https://github.com/THU-MIG/yolov10). In this comparison, a defogging network was employed as a preprocessing step before applying the YOLOv10 detection network. It is worth noting that



(a) Foggy (b) DCP (c) MSBDN (d) 4KDehazing (e) C2PNet (f) DEANet (g) PSD (h) SLAD (i) UMENet (j) DGD (k) Ours

Figure 8: Comparison of the HazDesNet predicted fog density maps on the defogging results. Deep blue indicates less fog and deep red indicates more fog residues.

due to the domain gap between synthetic training data and real-world data, supervised image defogging methods provide limited improvement in the performance of object detection algorithms. In some cases, they may even degrade detection performance by introducing noise during the defogging process [46]. Therefore, we focus solely on comparing prior-based methods and semi-supervised and unsupervised methods.

Precision and recall are fundamental metrics in object detection, used to assess the accuracy and completeness of a model's predictions. Precision measures the proportion of true positive detections among all predicted positives, while recall evaluates the proportion of true positives among all actual objects. Mean Average Precision (mAP) is a commonly used metric in object detection that provides a comprehensive evaluation of model performance by incorporating both precision and recall across various confidence thresholds. In our evaluation, we calculate mAP for each object category, with AP_{50} and AP_{75} representing the average precision at Intersection over Union (IoU) thresholds of 50% and 75%, respectively. Additionally, $AP_{50:95}$ is computed over a range of IoU thresholds from 5% to 95%, with increments of 5%.

Table 2 presents the quantitative comparison of object detection performance before

Table 2: Quantitative comparison of object detection performance improvement.

	Precision↑	**Recall↑**	**AP_{50}↑**	**AP_{75}↑**	**$AP_{50:95}$↑**
Foggy	0.57	0.46	53.6	37.4	34.1
DCP	**0.72**[2]	0.43	56.0	**38.4**[2]	35.4
PSD	0.62	**0.50**[2]	57.4	**38.4**[2]	**36.1**[2]
SLAD	0.60	0.46	55.6	37.4	34.8
UMENet	0.68	0.48	**58.7**[2]	37.3	35.6
DGD	0.67	0.45	54.5	35.7	34.2
Ours	**0.75**[1]	**0.52**[1]	**66.8**[1]	**43.5**[1]	**40.4**[1]

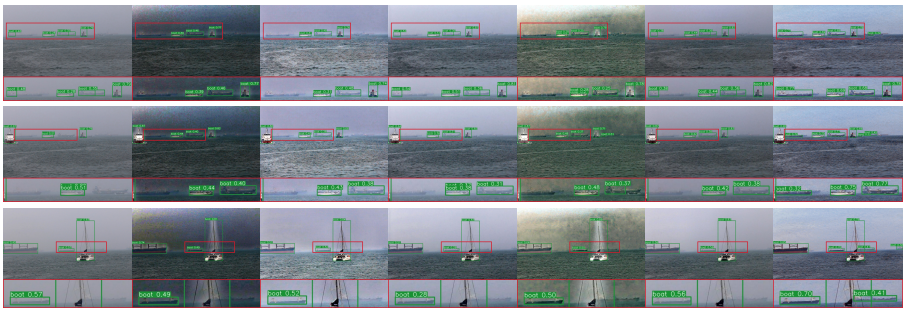
and after defogging. The results indicate that the proposed algorithm outperforms all other defogging methods across multiple evaluation metrics. Specifically, it achieves a 0.03 improvement in Precision over DCP, a 0.02 increase in Recall compared to UMENet, and a 4.3% improvement in $AP_{50:95}$ over PSD. These results highlight the effectiveness of the proposed approach in enhancing detection performance in foggy conditions.

Moreover, the qualitative comparisons presented in Fig. 9 further substantiate the effectiveness of our approach. It can be observed that although the image clarity improves after dehazing, artifacts appear around ships in the results generated by DCP and UMENet. Furthermore, the compared dehazing methods do not adequately account for the unique characteristics of overwater objects, leading to insufficient recovery of texture details and contours in critical regions such as ships, consequently resulting in limited improvements in object detection accuracy. In contrast, our proposed method significantly enhances the performance of the detection network, as evidenced by the accurate detection of previously missed objects and higher confidence scores.

5. Analysis and Discussion

5.1. Efficacy of the PG-CycleGAN

To validate the impact of the proposed PGTU module, PGRB, and prior map-guided losses, we conduct a comparative analysis.



(a) Foggy (b) DCP (c) PSD (d) SLAD (e) UMENet (f) DGD (g) Ours

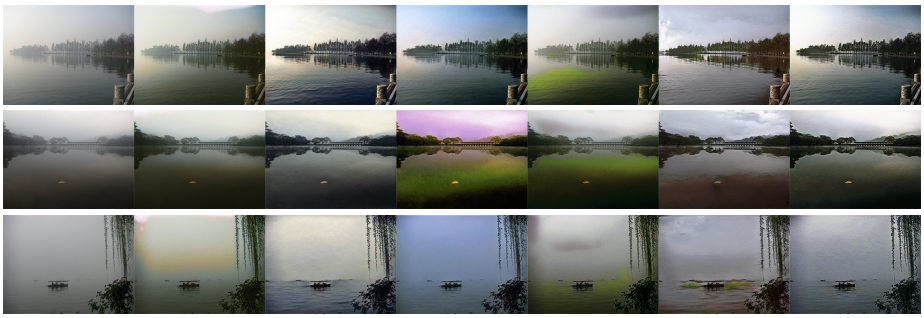
Figure 9: Quanlitative comparison of object detection performance improvement.

1) Evaluation of PG-CycleGAN with different settings: To evaluate the contribution of each component within PG-CycleGAN, we conducted a series of ablation experiments by individually removing or modifying four key modules: the PGTU, PGRB, prior map-guided cycle-consistent loss, and prior map-guided GAN loss. As summarized in Table 3, five ablation settings are evaluated: Setting i replaces the PGTU with a transposed convolution module for upsampling; Setting ii substitutes the PGRB with standard residual blocks; Setting iii replaces the prior map-guided cycle-consistent loss with the original cycle-consistent loss [47]; Setting iv substitutes the prior map-guided GAN loss with a conventional PatchGAN loss [23]; and Setting v replaces the prior map-guided GAN loss with a prior map-guided PatchGAN loss, where the prior map is downsampled to match the spatial resolution of the discriminator output.

Table 3: Quantative comparison of PG-CycleGAN with different settings.

Setting	**i**	**ii**	**iii**	**iv**	**v**	**vi**
FADE↓	1.67	**0.81**[2]	0.87	1.59	1.43	**0.73**[1]
HazDesNet↓	0.55	**0.33**[2]	0.35	0.51	0.45	**0.30**[1]
CV-CNN↑	0.69	**0.74**[2]	0.73	0.65	0.70	**0.77**[1]

Table 3 shows that removing PGTU causes the largest drop in FADE and HazDesNet metrics, indicating transposed convolution alone is insufficient for effective fog removal. When using the conventional PatchGAN loss, the CV-CNN metric experiences



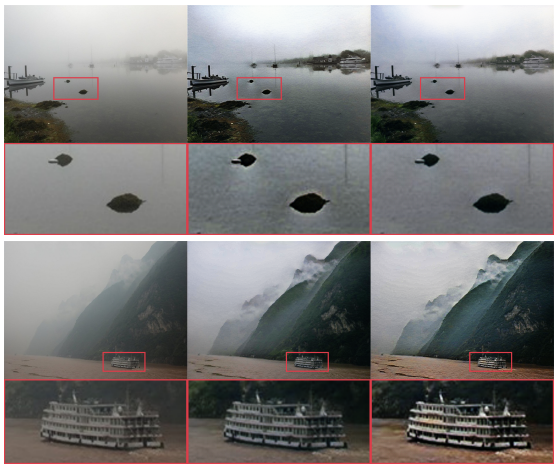
(a) Foggy (b) Setting i (c) Setting ii (d) Setting iii (e) Setting iv (f) Setting v (g) Setting vi

Figure 10: Qualitative comparison of PG-CycleGAN with different settings.

the largest decline. This is attributed to the lack of prior map guidance, which causes the GAN loss to rely on global features, thereby neglecting important local details and degrading image quality. Although incorporating the inverted dark channel prior map into the PatchGAN loss leads to some improvement, the performance still falls short of that achieved with our proposed prior map-guided GAN loss.

Fig. 10 provides a qualitative comparison of PG-CycleGAN under different configurations. It can be observed that when PGTU is removed, the residual fog in the images increases significantly. In Fig. 10 (d) and (e), it can be seen that the removal of the prior map-guided loss leads to noticeable color distortions in the defogged images. Even when employing a prior map-guided PatchGAN loss, noticeable artifacts persist in the generated images. This is because the patch-based discriminator, even with prior guidance, operates at a coarser spatial scale, which limits its ability to leverage the fine-grained information provided by the prior map. In contrast, the pixel-level discriminator offers more precise spatial alignment, resulting in better detail preservation and overall performance.

To further evaluate the impact of the minimum filter on network performance during prior map computation, we conducted an ablation study by alternately incorporating and removing the minimum filter in the calculation process. As illustrated in Fig. 11, it can be observed that although the overall defogging effects appear comparable, a detailed examination of the magnified regions reveals distinct differences. Specifically, when the minimum filter is applied, residual fog persists around the edges of overwater



(a) Foggy (b) w/ minimum filter (c) w/o minimum filter

Figure 11: Qualitative comparison of PG-CycleGAN with and without applying the minimum filter to the inverted dark channel prior map.

objects, and the texture clarity of object regions is slightly reduced compared to the scenario where the minimum filter is removed. This phenomenon occurs because the minimum filter aggregates the smallest pixel values within local regions, preventing the network from accurately focusing on overwater objects and causing attention to shift toward approximate surrounding areas.

2) Evaluation of PG-CycleGAN with different upsampling methods: To evaluate the impact of different upsampling strategies within the PGTU module, we conducted additional ablation experiments, as presented in Table 4. The tested methods include Transposed Convolution (TC), Bilinear Interpolation (BI), and PixelShuffle (PS), each applied individually and in pairwise combinations. The results show that due to the lack of learning capability, the BI method tends to leave more fog in the defogged images, resulting in inferior performance across all evaluation metrics. In contrast, the combination of two learning-based methods, TC and PS, yields improved results and achieves the second-best performance. Overall, the proposed PGTU module delivers the most effective and visually realistic dehazing results, attaining the highest scores on FADE, HazDesNet, and CV-CNN metrics.

3) Study on the influence of loss weights: To evaluate the effect of loss weight con-

Table 4: Quantative comparison of PG-CycleGAN with different upsampling methods.

Setting	TC	BI	PS	PGTU	FADE↓	HazDesNet↓	CV-CNN↑
i	✓				1.67	0.55	0.69
ii		✓			1.74	0.60	0.72
iii			✓		1.32	0.48	0.68
iv	✓	✓			1.13	0.39	0.72
v		✓	✓		0.97	0.42	0.73
vi	✓		✓		**0.83**[2]	**0.35**[2]	**0.75**[2]
vii				✓	**0.73**[1]	**0.30**[1]	**0.77**[1]

Table 5: Quantitative performance of FADE on the OverwaterHaze dataset using PG-CycleGAN trained with different loss weights.

λ_i	Weight					
	0.01	0.1	1	2	3	5
λ_1	0.90	0.82	**0.73**[1]	**0.78**[2]	0.85	0.88
λ_2	0.93	**0.77**[2]	**0.73**[1]	0.79	0.87	0.92
λ_3	0.97	0.85	**0.79**[2]	**0.73**[1]	0.81	1.08

figurations on network performance, we conducted experiments on the OverwaterHaze dataset with varying weight settings, as shown in Table 5 and Fig. 12. It is observed that setting λ_1 and λ_2 to 1 yields the lowest FADE score, indicating optimal dehazing performance, while the best result for λ_3 is achieved at a value of 2. Notably, variations in λ_3 lead to the most substantial changes in performance, which can be attributed to the intrinsic instability of the GAN loss during training. This sensitivity suggests that the prior map-guided GAN loss plays a particularly critical role in guiding the generator towards producing more perceptually realistic defogged images.

5.2. Limitations

As shown in Fig. 13, when the local area in the original input image is near to white, which usually happens on white hulls, the pixel in the corresponding inverted dark channel prior map will tend to 0. The reason is that if all the pixels of the three channels in the input possess a large value, the local region will show brightly in the dark channel, and then appear dark after the dark channel is inverted. This may lead to suboptimal solutions in network optimization, as it cannot give the objects a high

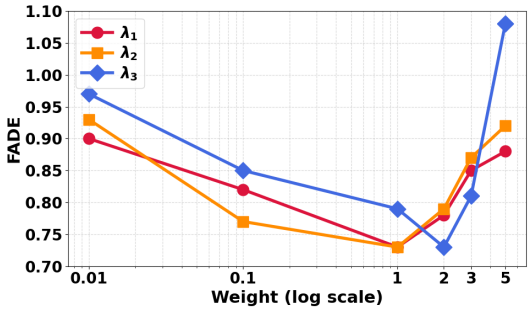


Figure 12: Performance plot for PG-CycleGAN trained with different λ values.

penalty. The following solutions could be used to overcome this problem, i.e., refining the weights of the prior map-guided losses, or adding an extra PatchGAN [23] discriminator to optimize the model.

(a) Input image (b) Dark channel (c) Prior map

Figure 13: The inverted dark channel prior map does not work well in white areas.

6. Conclusion

In this article, we present a PG-CycleGAN for overwater image defogging. We design an inverted dark channel prior map to highlight objects over the water and suppress the sky region, which is further incorporated into two loss functions that impose a greater penalty on the restoration of overwater objects. Additionally, a prior-guided residual block and a prior-guided TriUpsample module are developed, which effectively integrate the extracted prior features for both feature encoding and upsampling.

These components work together to enhance the network's ability to accurately restore overwater objects, improving overall defogging performance, particularly in challenging overwater environments. Through visual quality comparison, non-reference image quality assessment, and object detection performance evaluation, we demonstrate that our method outperforms existing state-of-the-art image defogging approaches.

By effectively highlighting overwater objects and distinguishing them from backgrounds, the proposed inverted dark channel prior map shows strong potential for high-level visual tasks in overwater scenes. In the future, we aim to integrate the prior map into detection frameworks to enhance overwater object contour extraction and improve detection performance. Additionally, the prior map may also be extended to support related tasks in overwater scenes such as instance segmentation and object tracking, broadening its applicability in real-world overwater surveillance and navigation systems.

References

[1] S. K. Nayar, S. G. Narasimhan, Vision in bad weather, in: Proceedings of the seventh IEEE international conference on computer vision, Vol. 2, IEEE, 1999, pp. 820–827.

[2] D. Berman, S. Avidan, et al., Non-local image dehazing, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1674–1682.

[3] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE transactions on pattern analysis and machine intelligence 33 (12) (2010) 2341–2353.

[4] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, Aod-net: All-in-one dehazing network, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4780–4788.

[5] Z. Chen, Z. He, Z.-M. Lu, Dea-net: Single image dehazing based on detail-

enhanced convolution and content-guided attention, IEEE Transactions on Image Processing (2024).

[6] Z. Chen, Y. Wang, Y. Yang, D. Liu, Psd: Principled synthetic-to-real dehazing guided by physical priors, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 7180–7189.

[7] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, Z. Wang, Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 1874–1883.

[8] R. Fattal, Single image dehazing, ACM transactions on graphics (TOG) 27 (3) (2008) 1–9.

[9] Q. Zhu, J. Mai, L. Shao, A fast single image haze removal algorithm using color attenuation prior, IEEE transactions on image processing 24 (11) (2015) 3522–3533.

[10] R. Fattal, Dehazing using color-lines, ACM transactions on graphics (TOG) 34 (1) (2014) 1–14.

[11] X. Liu, H. Li, C. Zhu, Joint contrast enhancement and exposure fusion for real-world image dehazing, IEEE Transactions on Multimedia (2021) 1–1.

[12] P. Ling, H. Chen, X. Tan, Y. Jin, E. Chen, Single image dehazing using saturation line prior, IEEE Transactions on Image Processing 32 (2023) 3238–3253.

[13] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.-H. Yang, Single image dehazing via multi-scale convolutional neural networks, in: European conference on computer vision, Springer, 2016, pp. 154–169.

[14] Q. Jiang, X. Yi, L. Ouyang, J. Zhou, Z. Wang, Towards dimension-enriched underwater image quality assessment, IEEE Transactions on Circuits and Systems for Video Technology (2024) 1–1.

[15] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, M.-H. Yang, Gated fusion network for single image dehazing, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3253–3261.

[16] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia, Ffa-net: Feature fusion attention network for single image dehazing, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 11908–11915.

[17] Y. Cui, Q. Wang, C. Li, W. Ren, A. Knoll, Eenet: An effective and efficient network for single image dehazing, Pattern Recognition 158 (2025) 111074.

[18] Z. Lyu, Y. Chen, Y. Hou, Mcpnet: Multi-space color correction and features prior fusion for single-image dehazing in non-homogeneous haze scenarios, Pattern Recognition 150 (2024) 110290.

[19] Y. Kang, Q. Jiang, C. Li, W. Ren, H. Liu, P. Wang, A perception-aware decomposition and fusion framework for underwater image enhancement, IEEE Transactions on Circuits and Systems for Video Technology 33 (3) (2022) 988–1002.

[20] Q. Jiang, Y. Kang, Z. Wang, W. Ren, C. Li, Perception-driven deep underwater image enhancement without paired supervision, IEEE Transactions on Multimedia (2023).

[21] J. Jin, Q. Jiang, Q. Wu, B. Xu, R. Cong, Underwater salient object detection via dual-stage self-paced learning and depth emphasis, IEEE Transactions on Circuits and Systems for Video Technology (2024).

[22] D. Engin, A. Genç, H. Kemal Ekenel, Cycle-dehaze: Enhanced cyclegan for single image dehazing, in: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2018, pp. 825–833.

[23] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.

[24] L. Li, Y. Dong, W. Ren, J. Pan, C. Gao, N. Sang, M.-H. Yang, Semi-supervised image dehazing, IEEE Transactions on Image Processing 29 (2019) 2766–2779.

[25] Y. Shao, L. Li, W. Ren, C. Gao, N. Sang, Domain adaptation for image dehazing, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 2808–2817.

[26] S. Zhao, L. Zhang, Y. Shen, Y. Zhou, Refinednet: a weakly supervised refinement framework for single image dehazing, IEEE Transactions on Image Processing 30 (2021) 3391–3404.

[27] H. Sun, Z. Luo, D. Ren, B. Du, L. Chang, J. Wan, Unsupervised multi-branch network with high-frequency enhancement for image dehazing, Pattern Recognition 156 (2024) 110763.

[28] Y. Liang, S. Li, D. Cheng, W. Wang, D. Li, J. Liang, Image dehazing via self-supervised depth guidance, Pattern Recognition 158 (2025) 111051.

[29] S. Zheng, J. Sun, Q. Liu, Y. Qi, S. Zhang, Overwater image dehazing via cycle-consistent generative adversarial network, in: Proceedings of the Asian Conference on Computer Vision, 2020.

[30] F. Chollet, Xception: Deep learning with depthwise separable convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1251–1258.

[31] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[32] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).

[33] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, M.-H. Yang, Multi-scale boosted dehazing network with dense feature fusion, in: Proceedings of

the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 2157–2167.

[34] Z. Zheng, W. Ren, X. Cao, X. Hu, T. Wang, F. Song, X. Jia, Ultra-high-definition image dehazing via multi-guided bilateral learning, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 16180–16189.

[35] Y. Zheng, J. Zhan, S. He, J. Dong, Y. Du, Curricular contrastive regularization for physics-aware single image dehazing, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023, pp. 5785–5794.

[36] Y. Liang, B. Wang, W. Zuo, J. Liu, W. Ren, Self-supervised learning and adaptation for single image dehazing, in: Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI-22), 2022, pp. 1137–1143.

[37] S. Moosbauer, D. König, J. Jäkel, M. Teutsch, A benchmark for deep learning based object detection in maritime environments, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 916–925.

[38] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[39] T. Guan, C. Li, K. Gu, H. Liu, Y. Zheng, X.-j. Wu, Visibility and distortion measurement for no-reference dehazed image quality assessment via complex contourlet transform, IEEE Transactions on Multimedia 25 (2023) 3934–3949.

[40] Q. Jiang, Y. Gu, C. Li, R. Cong, F. Shao, Underwater image enhancement quality evaluation: Benchmark dataset and objective metric, IEEE Transactions on Circuits and Systems for Video Technology 32 (9) (2022) 5959–5974.

[41] L. K. Choi, J. You, A. C. Bovik, Referenceless prediction of perceptual fog density and perceptual image defogging, IEEE Transactions on Image Processing 24 (11) (2015) 3888–3901.

[42] J. Zhang, X. Min, Y. Zhu, G. Zhai, J. Zhou, X. Yang, W. Zhang, Hazdesnet: An end-to-end network for haze density prediction, IEEE Transactions on Intelligent Transportation Systems 23 (4) (2022) 3087–3102.

[43] T. Guan, C. Li, Y. Zheng, X. Wu, A. C. Bovik, Dual-stream complex-valued convolutional network for authentic dehazed image quality assessment, IEEE Transactions on Image Processing 33 (2024) 466–478. doi:10.1109/TIP.2023.3343029.

[44] J. Cheng, Z. Wu, S. Wang, C. Demonceaux, Q. Jiang, Bidirectional collaborative mentoring network for marine organism detection and beyond, IEEE Transactions on Circuits and Systems for Video Technology 33 (11) (2023) 6595–6608.

[45] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, G. Ding, Yolov10: Real-time end-to-end object detection, arXiv preprint arXiv:2405.14458 (2024).

[46] S. Sun, W. Ren, T. Wang, X. Cao, Rethinking image restoration for object detection, Advances in Neural Information Processing Systems 35 (2022) 4461–4474.

[47] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 2223–2232.