Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 2160044, 19 pages
https://doi.org/10.1155/2022/2160044

*Research Article*

# Contrastive Learning-Based Haze Visibility Enhancement in Intelligent Maritime Transportation System

**Xianjun Hu**,[1] **Jing Wang**,[2] **and Guilian Li**[3]

[1]*College of Electronic Engineering, Naval University of Engineering, Wuhan 430033, China*
[2]*China Aerospace Science and Industry Corporation, Beijing 100048, China*
[3]*Shell Finding Housing (Beijing) Technology Co., Ltd., Beijing 100085, China*

Correspondence should be addressed to Xianjun Hu; xianjunhoo@163.com

Received 14 June 2022; Revised 17 July 2022; Accepted 22 July 2022; Published 30 September 2022

Academic Editor: Wen Liu



With the rapid development of artificial intelligence and big traffic data, the data-driven intelligent maritime transportation has received significant attention in both industry and academia. It is capable of improving traffic efficiency and reducing traffic accidents in maritime applications. However, video cameras often suffer from severe haze weather, leading to degraded visual data and ineffective maritime surveillance. It is thus necessary to restore the visually degraded images and to guarantee maritime transportation efficiency and safety under hazy imaging conditions. In this work, a contrastive learning framework is proposed for haze visibility enhancement in intelligent maritime transportation systems. In particular, the proposed learning method could fully learn both local and global image features, which are beneficial for visual quality improvement. A total of 100 clean images containing water traffic scenes were selected as the synthetic test dataset, and good dehazing results were achieved on both visual and indexing results (e.g., peak signal to noise ratio (PSNR): $23.95 \pm 3.48$ and structural similarity index (SSIM): $0.924 \pm 0.065$ for different transmittance and atmospheric light values). In addition, extensive experiments on real-world 100 water hazy images demonstrate the effectiveness of the proposed method (e.g., natural image quality evaluator (NIQE): $4.800 \pm 0.634$ and perception-based image quality evaluator (PIQE): $46.320 \pm 10.253$). The enhanced images could be effectively exploited for promoting the accuracy and robustness of ship detection. The maritime traffic supervision and management could be accordingly improved in the intelligent transportation system.

## 1. Introduction

The visual perception system is a fundamental means of environmental perception for intelligent ships and intelligent surveillance systems [1]. In good weather conditions, the visual perception system can obtain high-quality images of the surface traffic scene, which can assist intelligent ships and intelligent supervision systems in completing their missions. In hazy conditions, however, the images collected by the visual perception system are typically blurry, making it difficult to distinguish targets such as ships and bridges on the water, which has a severe negative impact on the development and application of intelligent shipping technologies. In recent years, several image dehazing algorithms have been proposed in response to the problem of image quality degradation in hazy scenes. Image enhancement-based, physical model-based, and deep learning-based methods make up most of the currently available algorithms for image dehazing tasks both domestically and internationally.

To more accurately and systematically describe the imaging principle in hazy scenes, Narasimhan et al. [2] proposed a well-known atmospheric scattering model. The proposed model considers that the scene information collected by visible light imaging equipment in a hazy scene consists of two parts, i.e., the attenuated part of the incident light obtained by the attenuated light reflected from the object and the atmospheric light imaging part caused by the scattering of other light. The mathematical expression of the hazy imaging model is as follows:

$$I(x) = J(x)t(x) + A(1 - t(x)), \tag{1}$$

where $I$ represents the hazy image collected by the imaging device, $J$ denotes the potential haze-free image, $x$ is the image pixel, $t$ represents the transmission, and $A$ denotes the atmospheric light value.

When collecting images in hazy water surface traffic scenes, the abundant water vapor and particulate impurities on the water surface will cause serious scattering and refraction effects on light. Therefore, the reflected light on the water surface target will be severely attenuated during the propagation of the light path to the image acquisition device. As the scene distance increases and the hazy concentration increases, the attenuation phenomenon becomes more and more obvious, which is named the incident light attenuation phenomenon. In addition, in a hazy environment, some reflected light (e.g., light sources and atmospheric light) also enter the imaging device due to the scattering of water vapor and particles, making the scene blurred and the details of distant objects unclear. This part of the imaging content is atmospheric light imaging. Under the dual influence of incident light attenuation and atmospheric light imaging, the images collected by smart ship vision sensors in hazy scenes often have problems of low contrast in distinguishing surface targets.

Deep learning has yielded positive results in the field of dehazing and has significant potential for future research. As a result of the influence of abundant water vapor on the water surface, the image with haze on the water surface is more likely to lose image structure and detail information, and the existing land dehazing algorithms are applied directly. It is often difficult to obtain the desired effect on images with haze on the water surface. Therefore, the dehazing network for water surface hazy images must be redesigned in accordance with the characteristics of water surface traffic scenes. This paper focuses on developing an image dehazing model for water surface traffic scenes using contrastive learning. We take into account the characteristics of a single water scene and a small number of targets, and the dehazing model uses both hazy and clear images during training to extract image features from a comparative learning network framework. Then, the clear image corresponding to the extracted image features and the hazy image is restored. The main contribution of our method differs from others in the following aspects:

(i) We propose a contrastive learning-based framework to improve the visibility of intelligent maritime transportation systems in hazy environments

(ii) We propose a global feature comparison module, a local feature comparison module, a feature fusion module, and a hybrid loss function to improve the dehazing performance based on contrastive learning models

(iii) Extensive experimental results show that our method has advanced performance compared to state-of-the-art methods in image dehazing performance and ship detection tests

The remaining portions of this article are organized as follows. Section 2 introduces related work on image dehazing. The majority of Section 3 introduces the framework for enhancing the visibility of hazy images that utilize CNN. Experiments on both synthetic and real-world hazy images are implemented in Section 4. In Section 5, we conclude our primary contributions.

## 2. Related Work

In this section, we briefly review the research on dehazing methods, including image enhancement-based, physical model-based, and deep learning-based methods.

*2.1. Image Enhancement-Based Methods.* The dehazing algorithm based on image enhancement employs the conventional digital image enhancement technology to improve image quality by increasing the contrast of the haze image, thereby achieving the purpose of image dehazing. Meanwhile, these methods ignore the physical imaging model of image quality degradation in the haze environment. In particular, histogram equalization (HE) [3] and Retinex [4] are the most widely used image dehazing algorithms, which are both computationally efficient and effective.

Histogram equalization [3] refers to an image enhancement technique that adjusts image pixels based on image histogram information. After histogram equalization, the image's pixel gray value range is expanded, and the gray value distribution is uniform, enhancing the contrast of the hazy image, enhancing the image's visual quality, and achieving the image dehazing. According to the different scopes of histogram equalization, the corresponding dehazing algorithms can be divided into global and local histogram equalization dehazing algorithms. The dehazing algorithms based on global histogram equalization [4–7] use the entire hazy image as a pixel adjustment unit, convert its pixel grayscale to a state of uniform distribution, enhance the global contrast of the hazy image, and thus achieve the dehazing effect. However, it is easy for the global equalization process to ignore the local details of the image, and it is easy to lose the local details of an image scene with uneven haze density. Local histogram equalization is an adaptive method for HE. Perform histogram equalization processing, and then superimpose the equalization results of each local area, to improve the visual quality of hazy images. The local histogram equalization methods [8–11] can pay greater attention to the image's finer details, but it will also increase the amount of calculation required. Kim et al. [8] proposed a subblock stacking algorithm to improve local histogram equalization while avoiding the problem of information loss caused by local division to reduce the amount of computation required. However, HE-based methods may increase the contrast of the background and decrease the contrast of the useful signal; the gray level of the transformed image is decreased, some details are lost, and the contrast is unnaturally heightened. Retinex [3] is a color theory proposed by Land et al., which has been validated by a large number of experiments, simulates the retinal cortex of the human brain, and has attracted a large number of researchers conducting in-depth studies. According to the Retinex theory, Jobson et al. [12] proposed a single-scale Retinex algorithm

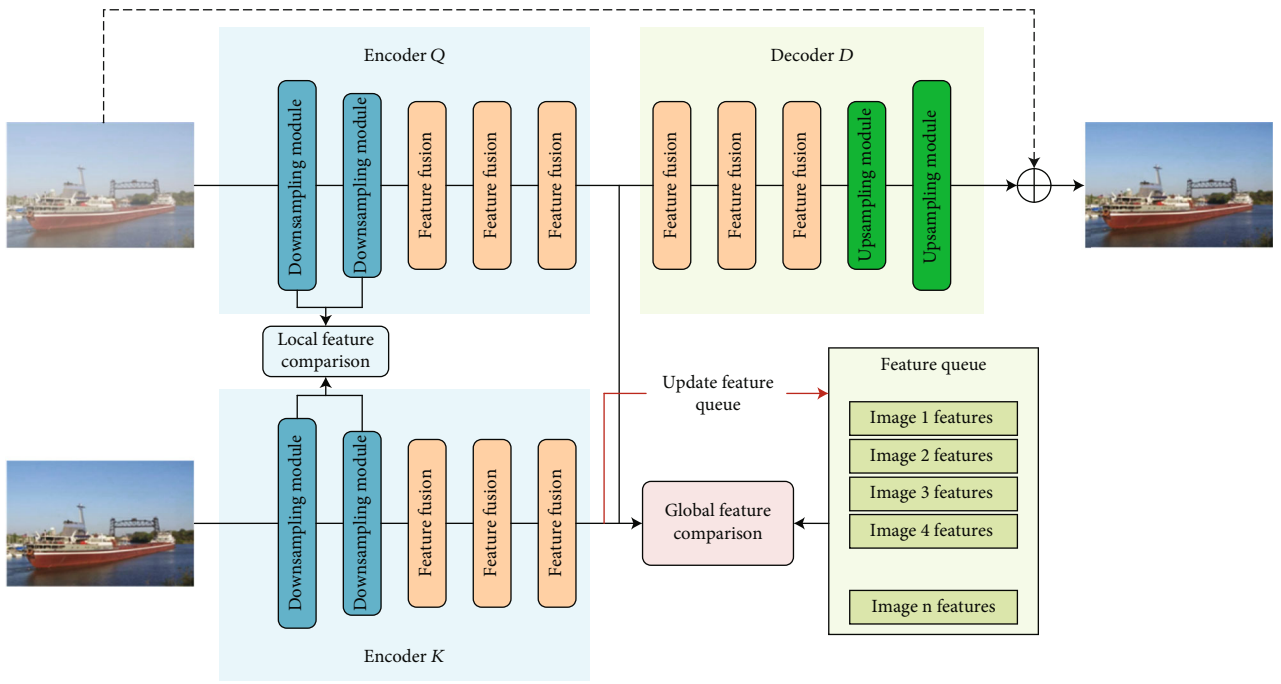



Figure 1: The flowchart of our contrastive learning-based dehazing model, which consists of two encoders and one decoder, named $Q$, $K$, and $D$, respectively.

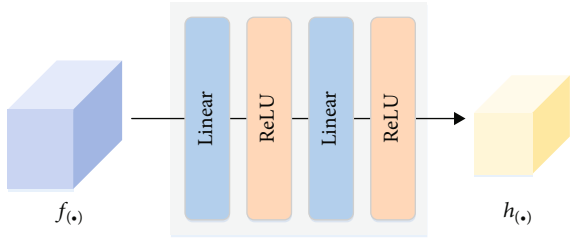



Figure 2: The pipeline of multilayer perceptron (MLP).

that used a Gaussian function and a low-pass filter to extract the illumination component of the image and had a good effect on image restoration. The multiscale Retinex algorithm [13] and its extension [14] employ RGB channel color information to prevent color loss. The enhancement-based method optimizes the image's contrast and color using the image enhancement technique. The final enhanced image has a wide dynamic range and distinct details, which can enhance the visual quality to some degree. However, these methods do not consider the physical process of image degradation in hazy weather, resulting in the loss of image background information. They cannot effectively remove the influence of haze, typically leading to issues such as detail distortion and insufficient dehazing.

*2.2. Physical Model-Based Methods.* Aiming to address the issue of insufficient hazy removal based on the enhancement approach, researchers have constructed mathematical models imaging according to the causes of hazy weather. Then, the mathematical model is determined in reverse, and the clear image's solution formula is obtained. These algorithms are known as physical model-based dehazing approaches, and they consist mostly of picture dehazing techniques based on image depth of field and prior knowledge. Dehazing based on image depth of field utilizes mostly professional equipment or specialized algorithms to extract the depth information of the image, calculates the parameters in the hazy imaging model, and then derives the formula to restore the hazy image. Oakley et al. [15] used the data collected by the radar sensor on the aircraft to determine the image's depth of field information and then estimated the transmittance to restore the image's clarity. By analyzing the link between light wavelength and picture contrast loss, Tan et al. [16] optimized the technique described by Oakley et al. and effectively used it to color haze image restoration. Narasimhan et al. [17] investigated multiangle photographs of the same scene in different weather conditions, calculated the depth information of the image using the information difference between the images, and then restored the haze-free image. Kopf et al. [18] directly employed 3D technology to determine the image's depth of field to estimate the image's transmittance and restore its clarity. The method based on prior information involves doing statistical analysis on the image data, combining certain physical models to propose some prior knowledge or assumptions, and then using these priors and assumptions to dehaze the image. He et al. [19] conducted statistical research on a large number of natural photos and discovered that in a limited local region of RGB color photographs, there must be a specific pixel in a specific channel whose pixel gray value is exceedingly low, even approaching zero. This concept is known as dark channel prior (DCP). Combining DCP and an atmospheric scattering model, He et al. introduced a picture dehazing technique based on DCP, which has a pronounced dehazing impact and is recommended by a large number of

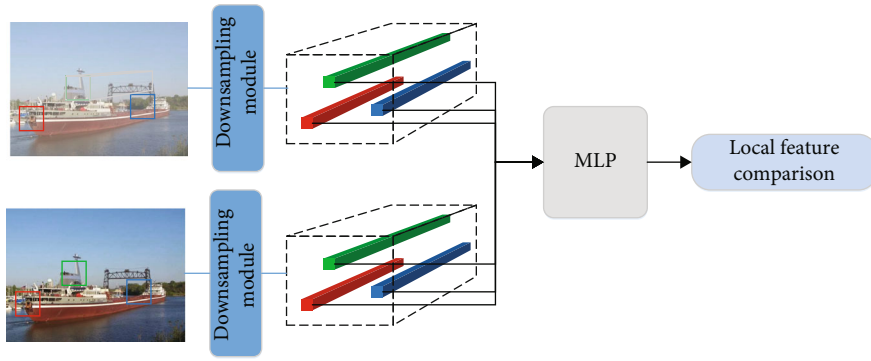



Figure 3: The pipeline of local feature comparison module.

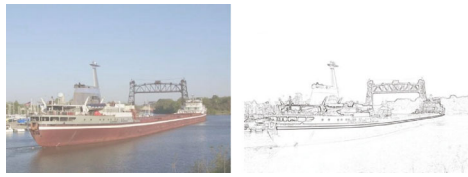


Figure 4: Comparison of original image and image gradient.

researchers. Additionally, several related better algorithms are regularly proposed. However, the color of the image dehazed by the DCP algorithm is typically dark. Due to the flaws of the previous theory, the DCP method frequently experiences distortion while dealing with clear sky and sea surface regions. These methods begin with the principle of dehazing, build a physical model to replicate the image degradation process, and then use the model to restore a clear image. The restored image created by adopting this method for image dehazing appears more natural and retains more of the original scene information. Due to the complexity of the image dehazing problem and the imaging variations between land and water scenes, the image dehazing algorithm based on a physical model [15–22] has difficulties for hazy water traffic scenes.

*2.3. Deep Learning-Based Methods.* Deep learning methods have been successfully used in different fields based on computer vision [23–36]. To distinguish the main core of the dehazing method, we still list the methods involving learning as a separate category, but the inclusion of enhancement- and physics-based modules in the deep model will be detailed. Cai et al. [23] employed convolutional neural networks for the first time in the task of image dehazing and proposed the DehazeNet network model. This network model estimates image transmission. The network model outputs the transmittance corresponding to the hazy image, then uses prior knowledge to estimate the atmospheric light value, and finally uses the atmospheric scattering model to restore the haze-free image. Ren et al. [24] developed a multiscale convolutional network model MSCNN for estimating the transmittance of hazy image data. The MSCNN model consists of two parts. The first part employs a larger convolution kernel to extract image features to generate a coarse transmittance map, while the second part employs a smaller convolution kernel to optimize the coarse transmittance map generated in the first part. Using a model based on atmospheric scattering, a haze-free image is reconstructed. The DehazeNet and MSCNN network models share a characteristic. To train the neural network, it is necessary to preestimate the transmittance of the training set images and then use the estimated transmittance map to restore clear images; however, multiple processing is susceptible to error superposition. Li et al. [25] disregarded the global atmospheric light and transmittance parameters in the atmospheric scattering model, directly used convolutional neural networks to learn the mapping from hazy images to clear images, and trained an end-to-end image dehazing network AODNet, as well as a network for image dehazing in the cloud. It has been demonstrated that dehazing images with the AODNet model can enhance the performance of advanced vision tasks, such as object detection algorithms. However, AODNet's single network structure makes it challenging to deal with complex water hazy scenes, resulting in subpar image performance after haze removal. Chen et al. [26] proposed a model for dehazing the GCANet network from end to end. In GCANet, smooth dilation is used to improve the atrous convolution, which eliminates the image grid artifacts that are easily generated during atrous convolution, and a gated fusion subnet is employed to weight and fuse the hazy image features extracted from different layers of the network model. Although atrous convolution improves the receptive field of the network, it also brings additional computational effort. Zhang et al. [27] proposed an end-to-end densely connected pyramid dehazing network DCPDN by fusing a pyramid pooling module with a densely connected module. Using dense connection blocks and a U-Net structure, the network learns the transmittance and atmospheric light value of hazy images and then outputs the corresponding clear images. However, water hazy imaging is complex and lacks reference depth information compared to land. It makes it difficult to deploy DCPDN in maritime intelligent supervision. The joint learning process can ensure that the dehazing results learned by the network model are more congruent with the physical model in use. Li et al. [28] designed a generative adversarial network model for image dehazing based on the concept of the conditional generative adversarial network [29] and introduced VGG image feature perception loss commonly used in image

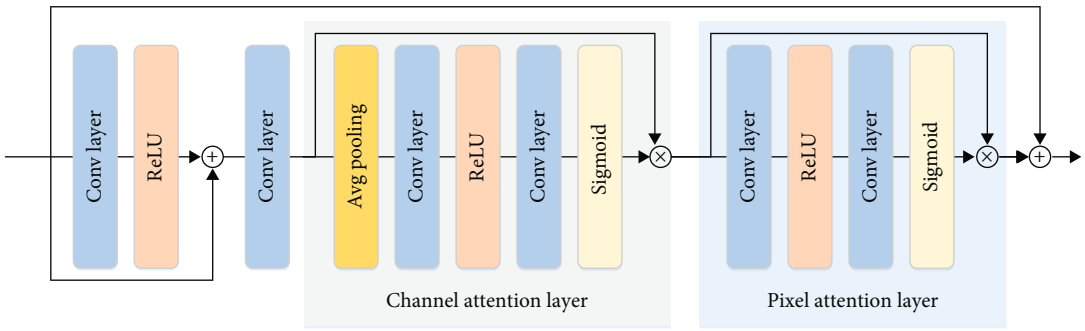
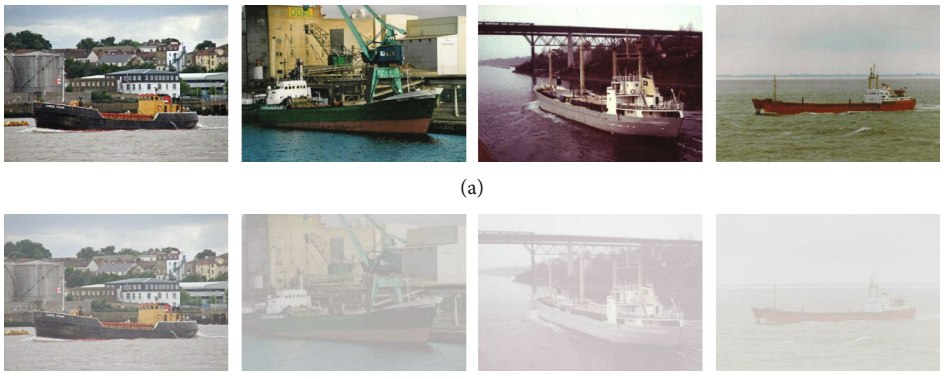



Figure 5: The pipeline of feature fusion module network module.



(a)

(b)

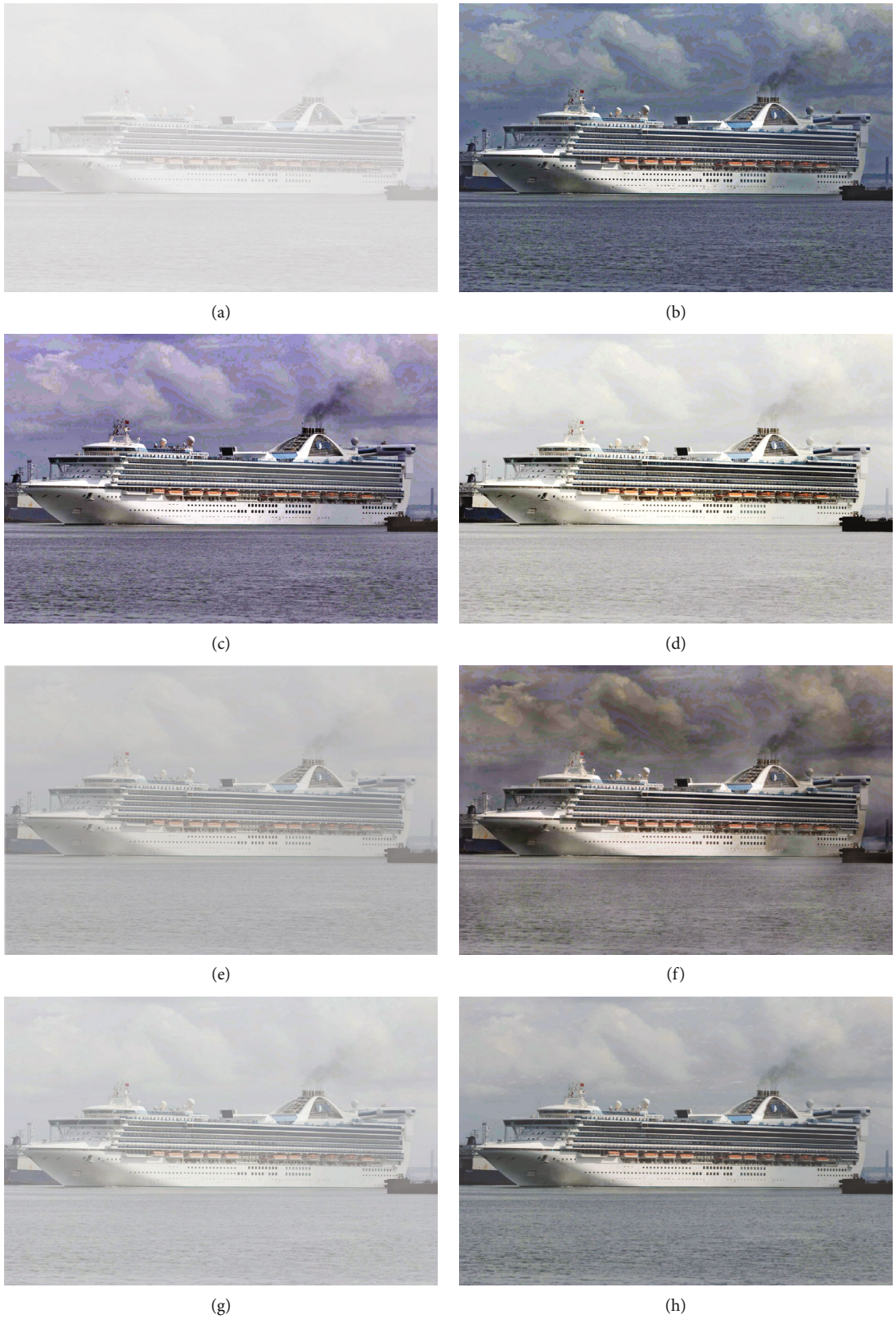
Figure 6: Examples of synthetic hazy images. (a) Original. (b) Synthetic hazy.

dehazing to the original GAN loss function and L1 gradient loss to improve the effect of image dehazing. But it still highly relies on paired data, resulting in poor robustness and generalization ability. Ren et al. [30] first utilized image white balance, contrast enhancement, and gamma correction digital image processing techniques to preprocess the original hazy image, which can extract various color or contrast features within the image, and then input it as input. In a network model with an encoding-decoding structure, the network model estimates the weights of various transformations of the hazy image before outputting the corresponding haze-free image via weighted fusion. However, when the hazy imaging environment is complex, the weight distribution is difficult to balance the advantages of different image processing techniques, and the generated dehazing images are visually unnatural. Yang et al. [31] argue that different statistically based priors are not applicable in all circumstances, but physical models can accurately describe hazy images. Therefore, Yang et al. proposed the disentangled dehazing network model, which directly designs different modules in the network model to simultaneously estimate the transmittance, atmospheric light value, and haze-free images in the atmospheric scattering model without requiring training with matching datasets. One can achieve a satisfactory image restoration effect. There is the same limitation as with DCPDN; the difficult-to-estimate depth information on the water will cause the image after dehazing to display excessive or insufficient dehazing. Engin et al. [32] proposed an image dehazing model cycle-dehaze that is based on cycle-consistent generative adversarial networks and does not rely on estimates that are independent of atmospheric light values and throw rates during training. When dehazing high-resolution images, the image fusion technology based on the Laplacian pyramid is utilized, which yields more effective clear and haze-free images. However, the image after cycle-dehaze dehazing will exhibit local color distortion and abnormal contrast changes.

## 3. Contrastive Learning-Based Image Dehazing Model

The abnormally dense water vapor in the water surface traffic scene causes the water surface hazy image to be more distorted than the land scene. The proposed contrastive learning-based dehazing network framework can use both hazy and clear images to train the network model, as both positive and negative samples have been thoughtfully designed. By comparing positive and negative samples, our model can improve its ability to extract image features and restore high-quality images of the water's surface.

Contrastive learning first learns a general representation of images from an unlabeled dataset, which is then fine-tuned with a small number of labeled images to enhance dehazing performance. Simply put, contrastive representation learning can be viewed as comparison-based learning. Contrastive learning is acquired by simultaneously



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 7: Continued.

(i)

Figure 7: Comparison experiment of different methods on synthetic hazy images in surface ship images. From top-left to bottom-right: (a) synthetic hazy image and restored images, generated by (b) DCP [19], (c) nonlocal [37], (d) F-LDCP [20], (e) AODNet [25], (f) GCANet [26], (g) MSCNN [24], (h) ours, and (i) ground truth, respectively.

maximizing the consistency between transformed views of the same image (e.g., cropping, flipping, and color transformations) and minimizing the consistency between transformed views of different images. In contrast, generative learning is a discriminative model that learns a mapping of certain (false) labels and then reconstructs the input samples. In contrastive learning, representations are acquired through the comparison of input samples. Contrastive learning does not learn the signal from a single data sample, but rather by comparing multiple samples. Through such comparative training, the encoder can learn image-level representations as opposed to image-level generative models.

As shown in Figure 1, the contrastive learning-based dehazing model proposed in this paper consists of two encoders and one decoder (i.e., $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{D}$). In the training phase, the encoders $\boldsymbol{Q}$ and $\boldsymbol{K}$, respectively, receive the hazy image and the clear image as input, and then compare the local and global features of the image features extracted by the encoder based on the idea of contrastive learning. Feature maps are added to the feature queue, loss function computation for contrastive learning, and gradient updates for the encoder. Finally, the extracted features are converted back to RGB images by the encoder $\boldsymbol{D}$. In the inference test phase, encoder $\boldsymbol{K}$ is no longer used, and only the combination of encoder $\boldsymbol{Q}$ and decoder $\boldsymbol{D}$ is used as the image dehazing model.

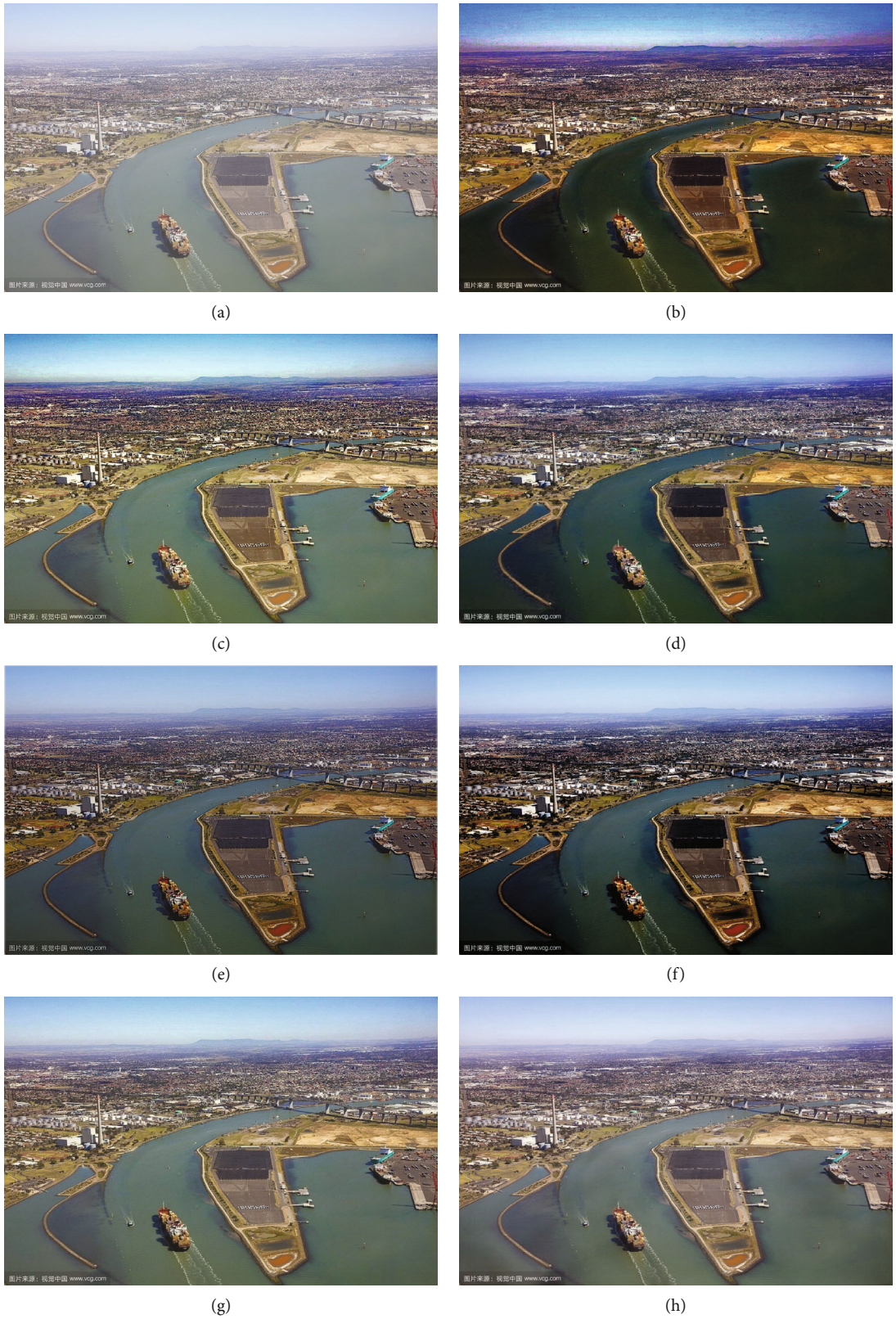
The contrastive learning-based image dehazing model proposed in this paper is a typical encoder-decoder model. The encoder accepts RGB images as input and extracts features from the input images via a series of downsampling and feature fusion modules. Typically, the image features extracted by the encoder are more complex and abstract expressions of the image that can represent the image's most essential characteristics. The model proposed in this chapter, which introduces the concept of contrastive learning, utilizes two encoders with the same structure to extract the features of different samples. Encoders are used to extract the features of the water surface hazy image and the water surface clear image, respectively. If the hazy image and the clear image on the water surface are from the same scene, then the image features extracted by the encoder should be similarly based on the concept of contrastive learning, and vice versa. The encoder's extracted features will be added to the feature queue, which is used to compare the loss function calculation and gradient update of the encoder. We use two encoders to extract the features of clear and hazy images that can assist the encoder in learning and expressing the true scene information of the images. The decoder's structure is the inverse of the encoder's, and its purpose is to restore the image features extracted by the encoder to a clear image. Due to the multilayer feature fusion module, the image features extracted by the encoder are more complex and abstract. Residual connections are used to connect the shallow output of the model to the decoder to better restore low-level image characteristics such as texture and shape.

*3.1. Global Feature Comparison.* To maintain the consistency of image feature distribution in the queue, it is necessary to extract image features using a stable encoder. Encoder $\boldsymbol{Q}$ is the model's primary encoder and is influenced by the gradient update. The change of encoder $\boldsymbol{Q}$ during each iteration is uncontrollable, and encoder parameters with significant changes will be incapable of producing image features with consistent distribution. Consequently, this model utilizes the momentum update method to update the encoder $\boldsymbol{K}$ based on the encoder $\boldsymbol{Q}$ model's parameters. Compared to the encoder $\boldsymbol{Q}$, the change in the encoder $\boldsymbol{K}$ is more subtle and stable, and it can provide image features with a consistent distribution when the number of iterations is relatively similar. The mathematical representation of the momentum update is as follows:

$$\theta_K = m \cdot \theta_K + (1 - m) \times \theta_Q, \tag{2}$$

where $\theta_K$ is the model parameter of the encoder $\boldsymbol{K}$, $\boldsymbol{Q}$ is the model parameter of the encoder $\boldsymbol{Q}$, and $m \in [0, 1]$ is the parameter of controlling the update speed, which is usually set to a number close to 1 to control the slow change of the encoder $\boldsymbol{K}$.

The global feature comparison of the model in this chapter is to use all the features of the extracted images for comparison. Suppose $f_q$ and $f_k$ are the encoder $\boldsymbol{Q}$ and the sample



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 8: Continued.

(i)

Figure 8: Comparison experiment of different methods on synthetic hazy images in Inland waterway. From top-left to bottom-right: (a) synthetic hazy image and restored images, generated by (b) DCP [19], (c) nonlocal [37], (d) F-LDCP [20], (e) AODNet [25], (f) GCANet [26], (g) MSCNN [24], (h) ours, and (i) ground truth, respectively.

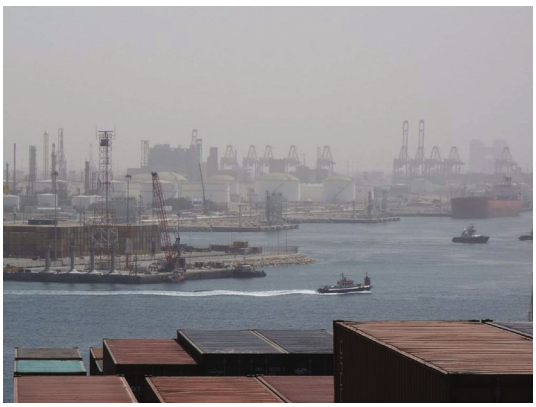
features extracted by the encoder $\boldsymbol{K}$, respectively. Global feature comparison does not directly use $f_q$ and $f_k$ for feature comparison, but first uses MLP (multilayer perceptron) to remap the feature $f_q$ and $f_k$ to obtain new feature $h_q$ and $h_k$ and then use $h_q$ and $h_k$ as a new feature for comparison loss calculation and gradient return. Experiments show that by nonlinearly transforming the features extracted by the encoder through MLP, more advanced and abstract image features can be extracted, and the performance of contrastive learning can be improved. The image features extracted by the original encoder are more suitable for image restoration tasks in addition to retaining the characteristics of contrastive learning. The MLP network structure used in this chapter is shown in Figure 2. We employ two "linear" and two "ReLU" operations in the MLP to increase its ability to map the learning data. The experimental results show that the nonlinear transformation of the features extracted by the encoder through MLP can extract more advanced and abstract image features and improve the performance of contrastive learning.

*3.2. Local Feature Comparison.* Most existing dehazing techniques only uniformly process the image as a whole, even though land scenes typically contain a wealth of scene information. Typically, images of water surfaces include expansive views of the sky and the water's surface. Scenes depicting water surface traffic emphasize waterborne objects, such as ships and shore structures. Different regions of the image contain distinct content and exhibit distinct properties. Therefore, using only global features as the target of contrastive learning is likely to result in insufficient dehazing of ships and other image targets, resulting in the loss of local detail information. The dehazing model for water surface images of water surface traffic scenes proposed in this paper includes a local feature comparison module to address the aforementioned problems.

This chapter proposes a local feature comparison module that operates on the encoder's initial two downsampling modules. Since the pixel blocks in the same position of the clear image and the hazy image of the same scene should have the same texture features, the local feature comparison module uses the pixel blocks in the same position of the clear image and the hazy image of the same water scene as the positive sample pair. As negative samples, pairs of pixel blocks at distinct positions are used. According to the properties of the convolutional neural network, the features extracted by the deeper network are more abstract, whereas the shallower network can readily extract shallow features such as texture, shape, and color. Therefore, the local feature comparison module proposed in this chapter will operate on the first two downsampling modules of the encoder. Figure 3 depicts the local feature comparison process within the output features of the first downsampling module. Similar to the global feature comparison module, MLP is used to remap the output features before the local feature comparison, with the remapped features serving as the comparison features for calculating the loss function.

We use a random method to select pixel blocks guarantees that each location within an image has an equal opportunity to participate in local feature comparison. Due to the large proportion of sky and water surface areas in the water surface traffic scene, it is simple to randomly extract multiple sky area or water surface area pixel blocks at once. According to the sample pair design of the local feature comparison module, pixel blocks that belong to the sky area or the water surface area (the characteristics can be considered consistent) will become negative sample pairs, which does not meet the definition of positive and negative samples and will impact the contrastive learning model. To address this issue, this chapter uses the gradients of the original image input as selection weights for pixel blocks. As depicted in Figure 4, the sky area and water surface area of the water surface image are typically flat, the change in the picture is relatively gradual, and the gradient is small, whereas the ships and other targets relevant to the water surface traffic task exhibit a significant change in the image and contain more images. The gradient is comparatively steep. Assuming that the image contains a total of pixel blocks, the gradient sum of the pixel block and its probability of being selected for feature comparison are

$$p(i) = \frac{\Omega(i)}{\sum_{l=1}^{N} \Omega(l)}. \tag{3}$$



(a)

(b)

(c)

(d)

Figure 9: Continued.

(e)

(f)

Figure 9: Continued.

(g) (h)

Figure 9: Comparison experiment of different methods on real hazy images in marine ports. From top to bottom: (a) real hazy image and restored images, generated by (b) DCP [19], (c) nonlocal [37], (d) F-LDCP [20], (e) AODNet [25], (f) GCANet [26], (g) MSCNN [24], and (h) ours, respectively.

*3.3. Feature Fusion Module.* The encoder-decoder of the image dehazing model proposed in this chapter contains three feature fusion modules, respectively. As shown in Figure 5, the feature fusion module includes a residual connection and a feature attention module, in which the feature attention includes a channel attention layer and a pixel attention layer, which can more flexibly deal with haze in images on the water surface areas with different concentrations. Residual learning bypasses less important information such as mist or low-frequency regions of the image through multiple residual connections, allowing the model to focus on more effective information, further improving network performance and training stability. Due to the complex water surface meteorology and the inconsistent effects of haze on different channels and different pixel positions, the use of channel attention layer and pixel attention layer can guide the network to pay attention to the information features of denser haze and high-frequency areas of the image in the water surface image.

The flowchart of our contrastive learning-based dehazing model, which consists of two encoders and one decoder, is named $\boldsymbol{Q}$, $\boldsymbol{K}$, and $\boldsymbol{D}$, respectively.

*3.4. Loss Function.* The loss function $\mathscr{L}_{total}$ of the image dehazing model proposed in this chapter consists of four parts, which can be given by

$$\mathscr{L}_{total} = \omega_1 \mathscr{L}_G + \omega_2 \mathscr{L}_L + \omega_3 \mathscr{L}_M + \omega_4 \mathscr{L}_P, \tag{4}$$

where $\mathscr{L}_G$ is global feature contrast loss, $\mathscr{L}_L$ is local feature contrast loss, $\mathscr{L}_M$ is mean square error loss, $\mathscr{L}_P$ is perceptual loss, and $\omega_{1-4}$, respectively, represent the weight of each loss function. We conduct multiple experiments to determine the weights of each loss function. The weights of $\omega_1$, $\omega_2$, $\omega_3$, and $\omega_4$ are 0.5, 0.5, 0.05, and 0.01, respectively.

*3.4.1. Global Feature Comparison Loss.* To constrain the model to learn image features better, this chapter uses the InfoNCE loss function to calculate the global contrast loss. According to the design of the positive and negative sample pairs in this paper, the expression of the global feature contrast loss function is

$$\mathscr{L}_G = -\log \frac{\exp\left(h(x) \times h(y)/\tau\right)}{\exp\left(h(x) \times h(y)/\tau\right) + \sum_{i=0}^{s} \exp\left(h(x) \times h(y_i)/\tau\right)}, \tag{5}$$

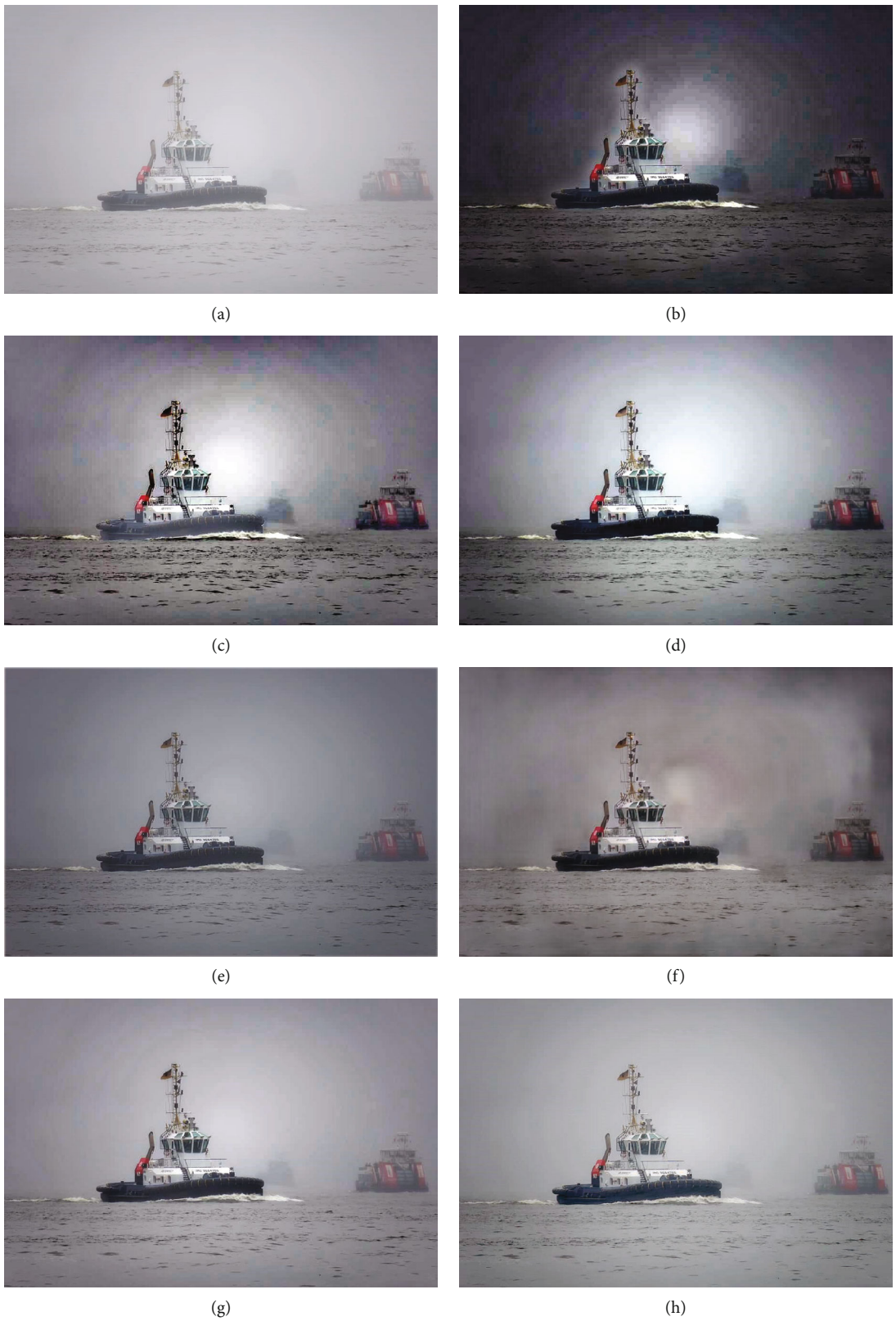
where $x$ and $y$ represent the hazy image and the corresponding clear image, respectively, $(\cdot)$ present the global image features extracted by the encoder and MLP, $\tau$ is a constant and is used to adjust the distribution of the loss, and $s$ is the number of image features contained in the feature queue.

*3.4.2. Local Feature Contrast Loss.* Since the water surface traffic scene has few targets and a single scene, to make the model better learn the information of the salient target area in the image, this chapter proposes a local feature comparison module. Unlike the global feature comparison, which uses the entire image as a comparison feature, the local feature comparison module focuses on the features of local areas of the image and extracts features from the shallow network of the encoder for comparative learning, which can better learn the details and textures of the image structural information. The expression of the local feature contrast loss function proposed in this chapter is

$$\mathscr{L}_L = \sum_{l=1}^{L} \sum_{i=0}^{N} -\log \frac{\exp\left(h^l(x^i) \times h^l(y^i)/\tau\right)}{\sum_{j=0}^{N} \exp\left(h^l(x^i) \times h^l(y^j)/\tau\right)}, \tag{6}$$

where $L$ represents the number of layers of the local feature comparison module, and in this chapter $L = 2$. $N$ is the number of pixel blocks of the local feature comparison module. $h^l(a^i)$ represents the local feature of the image extracted by the $i$th pixel block of the image $a$ in the local feature comparison module through the encoder $l$ layer and MLP.

*3.4.3. Mean Square Error Loss.* Mean squared error loss, also known as L2 loss, is a commonly used loss function in image restoration tasks. In the image dehazing task, the square sum



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 10: Comparison experiment of different methods on real hazy images in maritime surveillance. From top to bottom: (a) real hazy image and restored images, generated by (b) DCP [19], (c) nonlocal [37], (d) F-LDCP [20], (e) AODNet [25], (f) GCANet [26], (g) MSCNN [24], and (h) ours, respectively.

Table 1: Quantitate comparison on synthetic hazy images.

| Metrics | PSNR | SSIM | FSIM | $FSIM_C$ | VSI |
|---|---|---|---|---|---|
| DCP [19] | 17.26 ± 3.70 | 0.818 ± 0.080 | 0.926 ± 0.040 | 0.922 ± 0.041 | 0.973 ± 0.015 |
| Nonlocal [37] | 17.62 ± 3.15 | 0.793 ± 0.111 | 0.905 ± 0.076 | 0.898 ± 0.077 | 0.964 ± 0.023 |
| F-LDCP [20] | 18.96 ± 4.79 | 0.860 ± 0.123 | 0.939 ± 0.078 | 0.933 ± 0.077 | 0.977 ± 0.031 |
| AODNet [25] | 18.09 ± 3.62 | 0.800 ± 0.133 | 0.883 ± 0.080 | 0.881 ± 0.083 | 0.968 ± 0.027 |
| GCANet [26] | 18.35 ± 4.60 | 0.823 ± 0.096 | 0.930 ± 0.037 | 0.920 ± 0.043 | 0.973 ± 0.020 |
| MSCNN [24] | 16.59 ± 4.36 | 0.809 ± 0.114 | 0.923 ± 0.059 | 0.920 ± 0.061 | 0.975 ± 0.019 |
| Ours | **23.95 ± 3.48** | **0.924 ± 0.065** | **0.980 ± 0.018** | **0.977 ± 0.020** | **0.992 ± 0.007** |

Table 2: Quantitate comparison on real hazy images.

| Metrics | BRISQUE | BTMQI | NIQE | PIQE |
|---|---|---|---|---|
| DCP [19] | 0.517 ± 0.153 | 4.216 ± 1.187 | 5.178 ± 1.363 | 53.251 ± 16.278 |
| Nonlocal [37] | 0.498 ± 0.090 | 2.940 ± 0.808 | 4.936 ± 1.680 | 53.986 ± 15.663 |
| F-LDCP [20] | 0.517 ± 0.120 | 3.451 ± 1.174 | 5.111 ± 1.686 | 54.715 ± 15.805 |
| AODNet [25] | 0.519 ± 0.153 | 4.945 ± 1.289 | 5.324 ± 1.505 | 48.814 ± 12.024 |
| GCANet [26] | 0.472 ± 0.119 | 3.547 ± 1.118 | 5.240 ± 0.949 | 49.630 ± 12.061 |
| MSCNN [24] | 0.492 ± 0.065 | 3.659 ± 1.490 | 5.565 ± 1.492 | 53.663 ± 12.747 |
| Ours | 0.536 ± 0.121 | 2.942 ± 1.097 | 4.800 ± 0.634 | 46.320 ± 10.253 |

of the difference between the restored image and the clear image is used as the error value, which can better constrain the clear image restored by the model to be close to the real image at the pixel level. The expression for the mean squared error loss is

$$\mathscr{L}_M = \|y - \hat{x}\|_2^2, \tag{7}$$

where $\hat{x}$ represents the clear water surface image finally output by the model in this chapter.

*3.4.4. Content-Aware Loss.* Unlike the mean squared error loss, which calculates the gap between the restored image and the clear image, the perceptual loss calculates the high-level perceptual and semantic difference between the restored image and the clear image. We suggest perceptual loss to constrain the high-level features of the image can better preserve the image content and overall spatial structure and generate images with good visual perception effects. The content-aware loss proposed in this chapter uses a VGG-19 network pretrained on the ImageNet dataset to extract image features. The expression for the content-aware loss is

$$\mathscr{L}_P = \|\phi(y) - \phi(\hat{x})\|_2^2, \tag{8}$$

where $\phi(\cdot)$ represents the image features extracted by the 11th convolutional layer of the VGG-19 network.

## 4. Experiments

In this section, we conduct comparison experiments with six classical dehazing algorithms to verify the effectiveness of the proposed image dehazing model, including the physical model-based (i.e., DCP [19], nonlocal [37], and LDCP [20]) and the deep learning-based methods (i.e., AODNet [25], GCANet [26], and MSCNN [24]). The results are then evaluated in terms of both visual and quantitative metrics.

*4.1. Dataset.* The proposed contrast learning-based image dehazing network enhances the hazy images on the water surface by learning the transformation between hazy and clear images, which requires a large number of paired water surface images for network training. However, it is extremely difficult to capture the paired hazy and corresponding clear images of the same scene, which leads to a lack of real water surface image dehazing datasets. To tackle this problem, we employ synthetic methods to form a paired dehazing image dataset artificially. The basic clear images are mainly obtained from the OverwaterHaze dataset [38], the SeaShips dataset [39], Singapore Maritime dataset [40], and newly collected water surface traffic scene images.

To increase the diversity of data samples and improve the robustness of the network, we employ deep learning models and traditional methods to synthesize the hazy images simultaneously. Some of the synthesized images are shown in Figure 6.

*4.2. Dehazing Performance.* In this section, we select some competitive and classic dehazing methods and test them

Figure 11: Examples of detection results on the dehazed synthetic hazy images. From top to bottom: (a) clear image, (b) synthetic hazy images, and restored images, generated by (c) DCP [19], (d) nonlocal [37], (e) F-LDCP [20], (f) AODNet [25], (g) GCANet [26], and (h) ours, respectively.

FIGURE 12: Examples of detection results on the dehazed real-world hazy images. From top to bottom: (a) real hazy images and restored images, generated by (b) DCP [19], (c) nonlocal [37], (d) F-LDCP [20], (e) AODNet [25], (f) GCANet [26], and (g) ours, respectively.

on both synthetic and real images for comparison and then analyze the experimental result from visual and quantitate perspectives separately.

*4.2.1. Visual Comparison.* The comparison of synthetic hazy images between the proposed and other methods is shown in Figures 7 and 8. The output images of DCP, nonlocal, and GCANet look a little dark and inaccurate, the sky areas suffer from color deviation, and details are loss when the haze is thick. AODNet and MSCNN keep the most details, but when the haze becomes relatively thick, they cannot enhance the images with excellent restoration. F-LDCP successfully restores the hazy images with significant detail preservation, but the image becomes too bright. Compared with these methods, our proposed method dehazes the synthetic hazy images naturally with closer results to ground truth.

In addition to synthetic images, the proposed network is also tested on real hazy images, the results of which are shown in Figures 9 and 10. Due to the failure of the dark channel theory in the sky and water surface regions, the dehazing results of DCP in the water surface scenes look dark. Besides, nonlocal and F-LDCP suffer from color distortion, and the recovered images of AODNet perform badly in illumination, which leads to the vagueness of the ships. GCANet suffers from severe noise interference, while MSCNN is not able to reduce the noise sufficiently. Compared with previous methods, our network has better visual effects and makes the ships more prominent, which gives the convenience of following visual tasks.

*4.2.2. Quantitate Comparison.* To verify the performance of the proposed method, multiple quality evaluation metrics are employed in the experiment, including PSNR, SSIM [41], FSIM [42], FSIMC [42], and VSI [43]. FSIM evaluates the image quality by calculating the similarity of the image structure, which is the element that people pay more attention to in visual observation. FSIMC is the version of FSIM that takes the color factors into consideration. VSI mainly evaluates the image quality by comparing the difference between the saliency feature maps of two images, which pays more attention to prominent objects such as ships and buoys in waterborne transportation scenes. The results of the quantitate comparison are shown in Table 1. It is shown that the proposed method achieves the best performance on synthetic hazy images.

For testing on real hazy images, we introduced the NIQE [44], BRISQUE [45], BTMQI [46], and PIQE [47] as quality evaluation metrics, and the results are shown in Table 2. It can be seen that our proposed method archived competitive performance most of the time. The massive experiments indicate that the dehazed image of our method contains less noise with better image quality. Moreover, the structures of the dehazed images are close to ground truth.

*4.3. Results on Ship Detection with Synthetic Low-Lightness.* As an important means of waterborne transportation, ships are the main targets of intelligent ships sensing and intelligent maritime surveillance [47]. The current object detection algorithms can detect ships accurately based on clear visual data. However, the detection performance under hazy conditions will be greatly reduced. In this section, we will investigate the effect on ship detection after dehazing processing. To ensure the accuracy of the experiment, the experiments will be conducted using synthetic hazy images and real hazy images, respectively, and the detection algorithm we employ is pretrained YOLOv5.

The experimental results are shown in Figures 11 and 12. The bounding boxes indicate the location and the visual length of the ship, and the text and numbers on the boxes are the predicted categories and probabilities, respectively. In both synthetic and realistic hazy images, the accuracy of ship detection is greatly reduced, especially when the ships are far from the camera or too dense to capture the ships completely. The accuracy is significantly improved after the dehazing process. Besides, compared with other methods, the output dehazed images of the proposed method archived better performance on detection tasks due to the outperforming visual effect.

## 5. Conclusion

In this paper, we first analyze the characteristics of hazy maritime images and the reasons why existing dehazing algorithms fail in such scenes. To solve the issue of image dehazing in water surface traffic scenes, we propose a contrastive learning-based image dehazing model. To be special, the dehazing model consists of a global feature comparison module, a local feature comparison module, and a feature fusion module, which can fully utilize the information of hazy images and clear images to enhance the model's dehazing performance. In addition, a hybrid loss function is introduced to direct model training. Then, the water surface image dehazing experiments using synthetic and real images demonstrated that our dehazing model is capable of generating restored images with satisfying visual effects, as well as having significant advantages in terms of objective evaluation indicators. Experimental results demonstrate the superior dehazing ability of the proposed method, which archived competitive performance on both visual and quantitative evaluation. Besides, the ship detection experiment indicates that the dehazing effect of the proposed method is beneficial for the subsequent visual task in the maritime transportation system. However, there are still issues with the visual perception system of water surface traffic scenes that merit further investigation.

(i) The water surface traffic scenes obtained by intelligent ships and intelligent supervision systems via visual perception systems are typically in the form of video media as opposed to still images. When processing video dehazing, existing image dehazing algorithms typically loop a single frame image. When dealing with high-resolution video, it is difficult to meet real-time requirements using this technique. Consequently, research into video dehazing techniques and the enhancement of video dehazing performance is of great practical importance for enhancing the usability of intelligent systems

(ii) Due to the complexity of the water traffic environment, in addition to haze, the visual perception system is susceptible to adverse weather conditions, such as rain and snow, and low illumination. The method described in this paper can enhance the performance of the visual perception system to a limited degree, but it cannot eliminate the impact of harsh environments. Therefore, studying more visual perception enhancement technologies under the mixed influence of severe weather conditions can effectively enhance the visual perception capabilities of intelligent ships and intelligent supervision systems, as well as advance the development of intelligent shipping technology

## Data Availability

The image data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China under grant no. 41774021.

## References

[1] Y. Guo, Y. Lu, and R. W. Liu, "Lightweight deep network-enabled real-time low-visibility enhancement for promoting vessel detection in maritime video surveillance," *The Journal of Navigation*, vol. 75, no. 1, pp. 230–250, 2022.

[2] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for-visionm in bad weather," in *Proceedings of the IEEE Conferenceon Computer Vision and Pattern Recognition*, pp. 598–605, Hilton Head, SC, USA, 2000.

[3] S. M. Pizer, E. P. Amburn, J. D. Austin et al., "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.

[4] E. H. Land, "The retinex theory of color vision," *Scientific American*, vol. 83, no. 10, pp. 1986–3078, 1977.

[5] Y.-T. Kim, "Contrast enhancement using brightness preserving bihistogram equalization," *IEEE Transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[6] S.-D. Chen and A. R. Ramli, "Minimum mean brightness error bihistogram equalization in contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 4, pp. 1310–1319, 2003.

[7] C. Wang and Z. Ye, "Brightness preserving histogram equalization with maximum entropy: a variational perspective," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 4, pp. 1326–1334, 2005.

[8] J.-Y. Kim, L.-S. Kim, and S.-H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 475–484, 2001.

[9] Y. Wang and Z. Pan, "Image contrast enhancement using adjacent-blocks-based modification for local histogram equalization," *Infrared Physics and Technology*, vol. 86, pp. 59–65, 2017.

[10] F. Lamberti, B. Montrucchio, and A. Sanna, "CMBFHE: a novel contrast enhancement technique based on cascaded multistep binomial filtering histogram equalization," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 3, pp. 966–974, 2006.

[11] B. Liu, W. Jin, Y. Chen, C. Liu, and L. Li, "Contrast enhancement using non-overlapped sub-blocks and local histogram projection," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 583–588, 2011.

[12] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Transactions on Image Processing*, vol. 6, no. 3, pp. 451–462, 1997.

[13] Z. Rahman, D. J. Jobson, and G. Woodell, "Multiscale Retinex for color image enhancement," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 1003–1006, Lausanne, Switzerland, 1996.

[14] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965–976, 1997.

[15] J. P. Oakley and B. L. Satherley, "Improving image quality in poor visibility conditions using a physical model for contrast degradation," *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 167–179, 1998.

[16] K. K. Tan and J. P. Oakley, "Physics-based approach to color image enhancement in poor visibility conditions," *Journal of the Optical Society of America an Optics Image Science and Vision*, vol. 18, no. 10, pp. 2460–2467, 2001.

[17] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *International Journal of Computer Vision*, vol. 48, no. 3, pp. 233–254, 2002.

[18] J. Kopf, B. Neubert, B. Chen et al., "Deep photo," *ACM Transactions on Graphics*, vol. 27, no. 5, pp. 1–10, 2008.

[19] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

[20] Y. Zhu, G. Tang, X. Zhang, J. Jiang, and Q. Tian, "Haze removal method for natural restoration of images with sky," *Neurocomputing*, vol. 275, pp. 499–510, 2018.

[21] Q. Shu, C. Wu, Z. Xiao, and R. W. Liu, "Variational regularized transmission refinement for image dehazing," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 2781–2785, Taipei, Taiwan, 2019.

[22] S.-C. Huang, B.-H. Chen, and Y.-J. Cheng, "An efficient visibility enhancement algorithm for road scenes captured by intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2321–2332, 2014.

[23] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: an end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.

[24] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M. H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *Proceedings of the European Conference on Computer Vision*, pp. 154–169, Amsterdam, The Netherlands, 2016.

[25] B. Li, X. Peng, Z. Wang, Z. Wang, J. Xu, and D. Feng, "AOD-Net: all-in-one dehazing network," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4770–4778, Venice, Italy, 2017.

[26] D. Chen, M. He, Q. Fan et al., "Gated context aggregation network for image dehazing and deraining," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pp. 1375–1383, Waikoloa, HI, USA, 2019.

[27] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3194–3203, Salt Lake City, UT, USA, 2018.

[28] R. Li, J. Pan, Z. Li, and J. Tang, "Single image dehazing via conditional generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8202–8211, Salt Lake City, UT, USA, 2018.

[29] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, https://arxiv.org/abs/1411.1784.

[30] W. Ren, L. Ma, J. Zhang et al., "Gated fusion network for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3253–3261, Salt Lake City, UT, USA, 2018.

[31] X. Yang, Z. Xu, and J. Luo, "Towards perceptual image dehazing by physics-based disentanglement and adversarial training," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32no. 1, pp. 702–717, Salt Lake City, UT, USA, 2018.

[32] D. Engin, A. Genc, and H. K. Ekenel, "Cycle-dehaze: enhanced CycleGAN for single image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 938–946, Salt Lake City, UT, USA, 2018.

[33] Y. Lu, Y. Guo, and M. Liang, "CNN-enabled visibility enhancement framework for vessel detection under haze environment," *Journal of advanced transportation*, vol. 2021, Article ID 5598390, 14 pages, 2021.

[34] 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*, 2022.

[35] X. Hu, J. Wang, C. Zhang, and Y. Tong, "Deep learning-enabled variational optimization method for image dehazing in maritime intelligent transportation systems," *Journal of Advanced Transportation*, vol. 2021, Article ID 6658763, 18 pages, 2021.

[36] R. W. Liu, Y. Guo, J. Nie et al., "Intelligent Edge-Enabled Efficient Multi-Source Data Fusion for Autonomous Surface Vehicles in Maritime Internet of Things," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 3, pp. 1574–1587, 2020.

[37] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1674–1682, Las Vegas, NV, USA, 2016.

[38] S. Zheng, J. Sun, Q. Liu, Y. Qi, and S. Zhang, "Overwater image dehazing via cycle-consistent generative adversarial network," in *Proceedings of the Asian Conference on Computer Vision*, vol. 9no. 11, pp. 1877–1916, Kyoto, Japan, 2020.

[39] Z. Shao, W. Wu, Z. Wang, W. Du, and C. Li, "Seaships: a large-scale precisely annotated dataset for ship detection," *IEEE Transactions on Multimedia*, vol. 20, no. 10, pp. 2593–2604, 2018.

[40] D. K. Prasad, D. Rajan, L. Rachmawati, E. Rajabaly, and C. Quek, "Video processing from electro-optical sensors for object detection and tracking in a maritime environment: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 8, pp. 1993–2016, 2017.

[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[42] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[43] L. Zhang, Y. Shen, and H. Li, "VSI: a visual saliency-induced index for perceptual image quality assessment," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4270–4281, 2014.

[44] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2013.

[45] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[46] K. Gu, S. Wang, G. Zhai et al., "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure," *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 432–443, 2016.

[47] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *Proceedings of the 2015 Twenty First National Conference on Communications*, pp. 1–6, Mumbai, India, 2015.