

Research Article

A Method of Image Restoration for Distortion of Object in Water-Air Cross-Media

Yuhe Gao,¹ Jishen Jia,^{1,2} Lei Cai ,³ Meng Zhou,¹ Haojie Chai,³ and Jinze Jia⁴

¹School of Mathematical Sciences, Henan Institute of Science and Technology, Xinxiang 453003, China

²Henan Digital Agriculture Engineering Technology Research Center, Xinxiang 453003, China

³School of Artificial Intelligence, Henan Institute of Science and Technology, Xinxiang 453003, China

⁴School of Management, Henan Institute of Technology, Xinxiang 453003, China

Correspondence should be addressed to Lei Cai; cailei2014@126.com

Received 26 October 2023; Revised 2 April 2024; Accepted 5 April 2024; Published 22 April 2024

Academic Editor: Janez Perš

Copyright © 2024 Yuhe Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Uneven water-air media distribution or irregular liquid flow can cause changes in light propagation, leading to blurring and distortion of the extracted image, which presents a challenge for object recognition accuracy. To address these issues, this paper proposes a repair network to correct object image distortion in water-air cross-media. Firstly, convolutional combination performs feature extraction on water-air cross-media images, which retains the same features at the same scale and marks feature points with large differences. Then, an attention correction module for geometric lines is proposed to correct geometric lines in water-air cross-media images by comparing and sensing the marked feature points with large differences and utilizing the line similarity in positive and negative samples. Finally, the blurring artifact elimination module eliminates artifacts caused by image blurring and geometric line correction by using multiscale fusion of individual U-Net information streams. This completes the image restoration of object distortion under water-air cross-media. The proposed method is feasible and effective for restoring aberrated objects in water-air cross-media environments, with numerous experiments conducted on water-air cross-media image datasets.

1. Introduction

When the amphibious robot detects from the sea to the land, the image under the water-air cross-media is greatly affected by the external environment. On the one hand, the normal viewing angle is air-to-air, and the line-of-sight medium is air without change. Water-to-air is to detect things in the air through seawater in the water. The medium is uneven, and the light will undergo complex phenomena such as refraction, reflection, and diffraction, which will change the propagation path of the light, and the speed of the light in the medium will also change, which will lead to the geometric distortion of the image under the water-air cross-media, and then produce image distortion. On the other hand, the irregular motion of fluid causes image blur and artifact phenomenon, and bad areas such as weak texture scene and confusion of ringing artifacts appear. As a result, the feature discrimination of the image under the water-air cross-media

is low and the feature distortion is generated, which seriously affects the effective identification of the investigation object.

It is difficult to repair object distortion image in water-air cross-media. Firstly, various types of distortion [1, 2] existing in image data (such as radial distortion and tangential distortion) need to be accurately identified and modelled in order to achieve effective repair. Traditional image correction methods use camera calibration technology to obtain internal parameters and distortion parameters of the camera [3, 4] and then use these parameters to correct the image. While these methods work well under sufficient prior conditions, they are not expected to work well under limited conditions. Secondly, due to the difference between the images obtained under different water quality, the detail texture in the image data is not obvious. Traditional image restoration uses various filters [5–7] to eliminate image blur artifacts and restores the clarity of the image by filtering the image in the frequency domain or airspace [8]. However,

a filter-based approach may result in loss of detail or the introduction of other artifacts.

In order to solve the above problems, this paper proposed a method of object distorted image inpainting under water-air cross-media. Firstly, the features of positive and negative samples of the image were extracted and fused. The convolution combination is used to extract the features of positive and negative samples, and the extracted features of the same scale are fused. Then, a water-air cross-media image geometric line attention correction module was designed. Self-attention perception of geometric lines in the image will produce distorted line correction. Finally, a water-air cross-media image blurring artifact elimination module was designed. For the blur of the image and the blur caused by geometric line correction, the multiscale feature fusion module is used to partially eliminate the blur artifacts in the image.

To address these issues, this paper proposes a restoration network on water-air cross-media images of objects. However, the existing methods do not adequately consider the irregular distortion and blur feature extraction of images. The following contributions are proposed:

- (1) Aiming at the problem of feature aberrations of the objects existing in water-air cross-media images, this paper proposes a positive and negative sample feature extraction attention module. By comparing the spatial location features of clear image objects and water-air cross-media image objects, the feature points with smaller spatial location differences are retained, and the feature points with larger spatial location differences are marked for attention. Thus, the aberration features can be better focused
- (2) Due to the uneven distribution of water quality, the propagation path of light is altered, causing uncertain aberrations at the water-air cross-media. In this paper, we propose a geometric line attention correction module to address this issue. By reducing the distance between feature points with significant differences in positive and negative samples, this module achieves aberration correction. It enables the flexible and efficient capturing of global-local line features, effectively dealing with distorted lines in various scenes
- (3) To address the issue of irregular fluid flow, which leads to artifacts or blurring in water-air cross-media images, as well as the presence of feature residuals at the boundary of the corrected region after geometric line attention correction, this paper proposes a blur artifact elimination module. This module utilizes the multiscale fusion of individual U-Net information streams to efficiently remove the blurred artifact portion of the image, enabling effective multiscale deblurring

2. Related Work

Feature extraction and feature fusion are commonly used techniques in machine learning and pattern recognition. It

is used to extract valuable information and features from raw data and fuse these features into higher-level representations. A more comprehensive and accurate representation of images can be achieved by combining various feature extraction methods or features at different scales. Feature fusion enhances the performance and reliability of image processing and computer vision tasks by leveraging the strengths of different features. It serves as an effective approach to enhance image analysis and recognition, enabling the system to be more robust and accurate in complex and diverse scenarios. In [9], ResNet-50 is used for feature extraction in the encoding stage, and cascaded upsampling is used to recover the feature map resolution in the decoding stage, fusing multiscale image features and spectral feature pyramid structures layer by layer. Qiu et al. [10] proposed a parallel network consisting of a two-stream feature extraction and fusion module and a context extraction and transcription module that fuses content and location features extracted by two feature extraction networks. In [11], attention-based multistage multitask fusion was used for feature extraction, having low- and high-level fusion with matching attention, and efficient FPN were used for point-level fusion of cross-sensors and single sensors, respectively. Qi et al. [12] used ResNet50 as the feature extractor for the network, introducing channel attention and spatial attention to the extracted high-level underwater estimation features. This study [13] proposed a double-pyramid repair framework. The pyramid attention mechanism in the decoder, which acquires finer patches directly from the learning layer, complements the layer-by-layer pyramid convolutional feature extraction in the encoder, thus facilitating feature representation. Chandrashekar et al. [14] proposed improved deep learning architecture consisting of U-Net and attention gating. This article [15] proposed the introduction of multiscale cyclic residual convolution in generative networks and the use of attentional skip connections to enhance the information interaction between features of different scales. Li et al. [16] used optimized ResNet-34 for feature extraction. Ref. [17] proposed a new feature selection method for class label-specific mutual information. Each class tag selects a set of specific features, maximizes the information shared between the selected features and the object class tags, and narrates the shared information among all classes. In general, feature extraction and fusion techniques have the potential to enhance the performance of computer vision tasks. However, it is important to acknowledge that there exist challenges and limitations. For instance, combining different features can introduce redundancy, thereby degrading performance. Therefore, this paper focuses on utilizing feature fusion to address the issue of size variability between features and mitigate any negative impact on performance.

Distortion correction is used to remove the distortion effect in the image and make the image more consistent with the real-world geometry. In addition to traditional camera calibration methods, deep learning-based distortion correction methods have also emerged in recent years. These methods employ deep convolutional neural networks to learn distortion patterns within an image. This is achieved by constructing a convolutional neural network that maps

the input distorted image to a geometrically correct image. Through this approach, the network gains the ability to understand and rectify the distortions present in the input image. These methods are trained by a large amount of calibration data and have high correction accuracy. In recent research [18], the quantification and correction of aberrations are discussed. It is mentioned that adaptive focusing can compensate for these aberrations but is only effective within a limited area of isoplanar patches. Ref. [19] proposed a new underwater image saliency detection framework that computes and estimates reliable underwater image saliency maps using Weberian descriptors based on quaternionic digit distance, pattern discriminability, and local contrast. In [20], Gao et al. used a single image to calibrate the camera through a specially designed checkerboard. Cai et al. [21] proposed a dynamic multiscale feature fusion method for underwater object recognition, which learned the spatial semantic features of the object through dynamic conditional probability matrix to improve the accuracy of underwater distorted objects. Mozaffari et al. [22] proposed a high-quality eye-tracking reference frame to improve real-time active eye movement correction system for revisit accuracy between consecutive imaging sessions. Some recent work [23] proposed a feature-level correction scheme with a correction layer embedded in the jump connection. Image features are precorrected by separating two parallel and complementary structures, content reconstruction and structure correction. Xu et al. [24] established a distortion model for linear processing of the radial distortion generated by images, iteratively estimated the unknown image mapping model online based on the classical Slotine-Li adaptive algorithm, and estimated the internal and external parameters in real time. This article [25] fitted the resonance scanner orientation data to a cosine model for correcting image distortion and sampling jitter, as well as accurately interleaving the image lines collected in the clockwise and anticlockwise resonance scanner portions of the rotation cycle for accurate correction. This study [26] introduced graph-based registration and mixing procedures to formulate and solve feature matching in each seafloor image pair through graph matching, so as to incorporate structural information between features. In their work, Decker and Zhang [27] introduced a novel application of dynamic time warping. This application allows for the estimation of the direction and relative strength of elliptical horizontal transverse isotropy anisotropy. Tian and Srinivasa [28] used fluctuation equations to develop a water surface spatial aberration model and algorithms to track water aberrations for image restoration of underwater images. To sum up, both traditional camera calibration methods and deep learning-based calibration methods have their own advantages and scope of application. However, for the uncertainty of aberrations in water-air cross-media images, an accurate camera model cannot be relied upon. Moreover, traditional methods may result in images that still contain distortions or have insignificant correction effects after restoration. Therefore, this paper proposes a geometric line attention correction module. This module utilizes attention perception of geometric lines to reduce the gap between positive and negative sample line features, thereby achieving improved geometric line restoration.

The primary advantage of deep learning methods in image restoration lies in their ability to automatically discern patterns and principles of image restoration from extensive training data. Through exposure to images within the training set, the deep learning model is capable of assimilating structural information and applying it to the restoration process. This data-driven approach is more adaptable and versatile compared to traditional, manually crafted image restoration algorithms. Furthermore, the deep learning approach can address various types of image restoration tasks, including denoising, deblurring, and missing data recovery. By thoughtfully designing the network structure and loss function, it can be tailored to different restoration tasks, yielding superior restoration results. Ljubenovic et al. [29] focus on a procedure in reducing degradation effects, frequency-dependent blurring, and noise in terahertz time-domain spectroscopy images in reflective geometry. Ref. [30] surveyed and generalized the existing relatively mature and representative underwater image processing models. The current status and future trends of underwater image processing are objectively assessed, and some insights into underwater vision and research directions for future development are provided. In recent research [31], an iterative filtering adaptive network was introduced into end-to-end learning to address the challenges posed by spatial variations and significant out-of-focus blurring. Liu et al. [32] proposed an attention-guided global-local adversarial learning network, which generates coarse fusion results under the attention weight graph and obtains regions of interest, edge loss functions, and spatial feature transformation layers to refine the fusion process. Their study [33] proposed an underwater image enhancement method based on potential low-order decomposition and image fusion and implemented an improved Laplace sharpening method and a gamma correction technique to adaptively compensate for color for removing color distortion. Jiang et al. [34] integrate visual and temporal knowledge at both global and local scales using convolutional recurrent neural networks. The proposed microable directional event filtering module enables the extraction of rich a priori boundaries from event streams. This method proves to be effective in handling real-world motion blur. This article [35] proposed to extract small object features by hybrid expansion convolutional network. Spatial semantic features are learned by adaptive correlation matrix and fused with spatial semantic features and visual features for underwater fuzzy object recognition. Zhang et al. [36] proposed a dual-path joint correction network. The method utilizes the multiscale U-Net to adaptively fusion features from different paths to generate enhanced images. In [37], Liu et al. introduced the attention mechanism of the model in terms of four dimensions: multiscale attention, channel attention, structural attention, and region of interest attention using dense blocks as a framework. The data is trained with the help of weakly supervised model. It is worth noting that the diversity of the datasets traditionally used for removing blur artifacts is insufficient to cover all possible types of blur artifacts. Furthermore, there are still some limitations in restoration, particularly for detailed textures. Therefore, this paper proposes a multiscale fusion of

the U-Net network information streams to remove the blur artifacts. The different scales of the U-Net network can learn distinct features and patterns, and their integration can provide a more comprehensive and accurate image restoration.

3. Proposed Method

The proposed method in this chapter consists of three parts, as illustrated in Figure 1. The overall architecture is based on U-Net. The first part is a feature extraction attention module for positive and negative samples of water-air cross-media images. It involves extracting features from the samples using convolutional combinations, where features with small differences are retained and features with large differences are labeled. The second part is the attention correction module for geometric line of water-air cross-media images. It focuses on correcting features that cause distortion by reducing the distance between labeled features with significant differences. Lastly, the third part is the blur artifact elimination module for water-air cross-media images. It addresses blurring artifacts generated by both self-contained image blurring and geometric line correction. The module partially eliminates these artifacts by utilizing multiscale fusion of individual U-Net information streams.

To enhance the capacity and effectiveness of the proposed method, we introduce additional skip connections at the input image locations, as represented by the dotted lines in Figure 1. The distorted image inpainting method under the water-air cross-media thus provides more paths for image feature information to be transmitted more quickly to subsequent levels. This improvement helps to enhance feature transmission and learning, enabling the network to better capture the structure and characteristics of the input data, as well as the nonlinear mapping relationship between the input and output, leading to better performance of the model. Overall, this approach results in a more expressive and efficient model.

3.1. Feature Extraction and Attention Module of Positive and Negative Samples of Water-Air Cross-Media Images. Due to the influence of water surface fluctuation in the water-air environment, the feature discrimination in the water-air cross-media images is low, and the image distortion part is easy to be ignored by the algorithm. In this chapter, the positive-negative sample feature extract attention module (PSFEAM) of the water-air cross-media images is constructed. For the positive and negative samples of the water-air cross-media images, the spatial location features of the object are extracted, respectively, and the feature points with large differences in spatial location are marked with attention. The overall process of feature extract and attention module for positive and negative samples is shown in Figure 2.

In this chapter, the water-air cross-media image and its different scale views are used as positive samples x_i ($i = 1, 2, 3$), and the real ground image and its different scale views are used as negative samples x_i^- . The feature extraction module (FEM) is trained from multiscale positive sample images $X = \{x_1, x_2, x_3\}$ and multiscale negative sample images $X^- = \{x_1^-, x_2^-, x_3^-\}$.

The whole network framework contains six feature extraction modules, two of which are a group for feature extraction of positive and negative samples of the same scale. The input side inputs $256 \times 256 \times 32$, $128 \times 128 \times 64$, and $64 \times 64 \times 128$ pixel size positive samples x_i and negative samples x_i^- . Each scale image contains two convolution branches FEM_i and FEM_i^- , applying the same convolution structure. The convolutional structure uses a stack of two 3×3 and 1×1 convolutional layers. An additional 1×1 convolutional layer is used to further refine the connected features.

The principle of extracting feature information from positive and negative samples is as follows:

$$\begin{aligned} T_{\text{feature}} &= f_{\text{extract}}(X; H, W), \\ T_{\text{feature}}^- &= f_{\text{extract}}^-(X^-; H, W), \end{aligned} \quad (1)$$

where f_{extract} represents the extraction function of positive sample features and f_{extract}^- represents the extraction function of negative sample features. The H and W denote the width and height of the input image. The feature extraction module processes each scaled image to extract relevant features, denoted as T . Through the convolutional layer, small feature differences between positive and negative samples are retained, while large feature differences are emphasized in order to delineate the spatial location of water-air cross-media objects in the image.

This subsection is aimed at retaining the feature points extracted from the positive and negative samples with small differences, while paying attention to the feature points with significant differences. The feature points extracted from the negative samples are standard feature points since the negative samples are ground-truth images. However, in the case of positive samples, which are distorted due to the water-air cross-media environment, some feature points may have altered spatial locations. In this section, the feature points with significant differences are marked for attention in preparation for the correction module in the next section.

3.2. Attention Correction Module for Geometric Lines in Water-Air Cross-Media Images. After extracting the image features and performing feature fusion, it is necessary to correct the distorted parts in the water-air cross-media image to minimise the influence of the image feature information by the interference of image distortion. This chapter proposes a geometric line attention correction module (GACM). The method corrects lines that cause image distortion by using self-attentive perception of feature points. It focuses on feature points with large differences between positive and negative samples and reduces their feature distances. In this way, water-air cross-media image repair can be handled effectively. The overall process of the geometric line attention correction module is shown in Figure 3.

Geometric line attention correction is performed using an attention mechanism for various distorted geometric lines in the image. The geometric line attention correction module consists of three channel attention modules (CAM) and two spatial attention modules (SAM) to extract the local-global geometric line features of the image. It contains one global

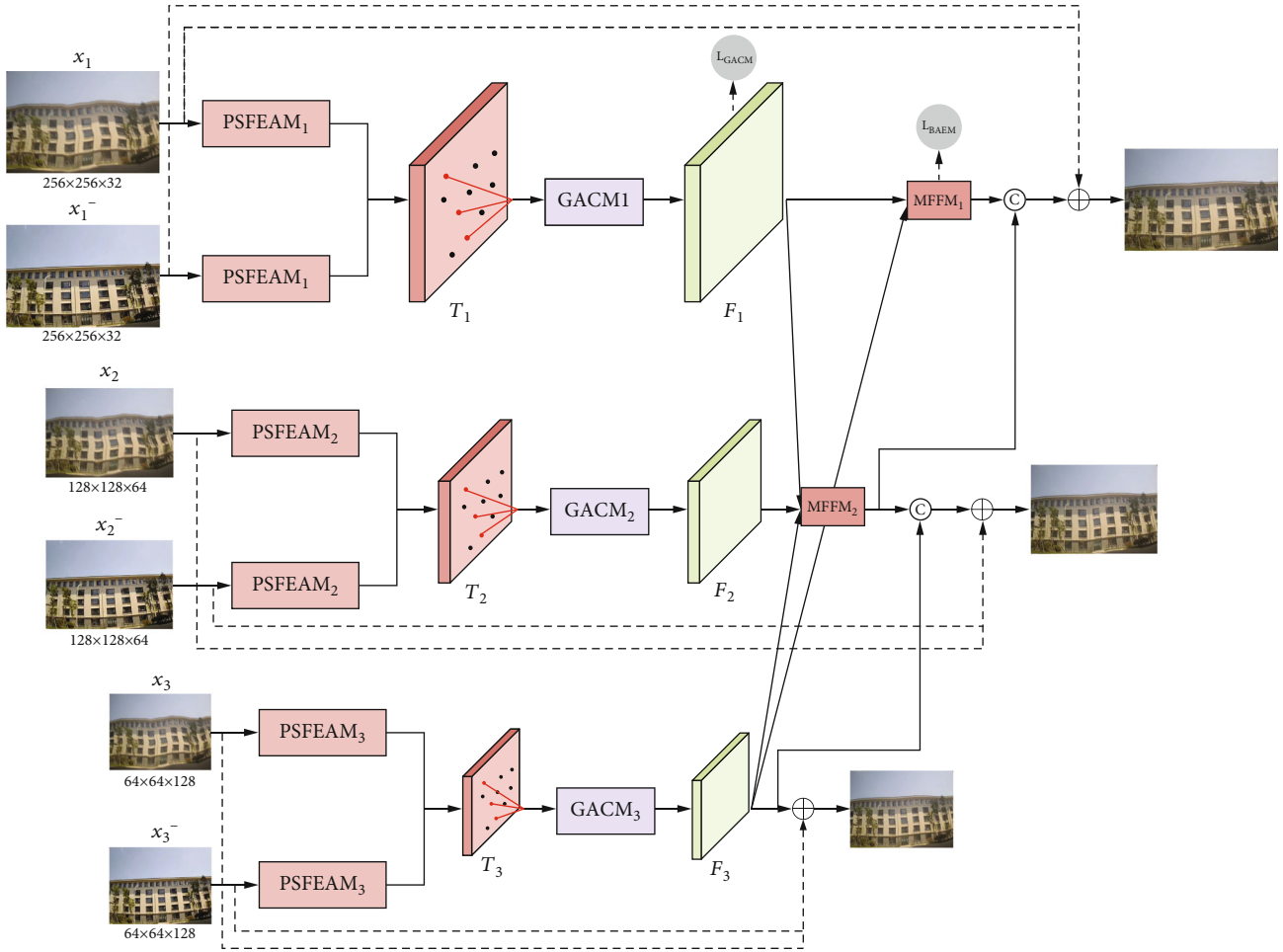


FIGURE 1: General flow chart of object distortion image repair method in water-air cross-media.

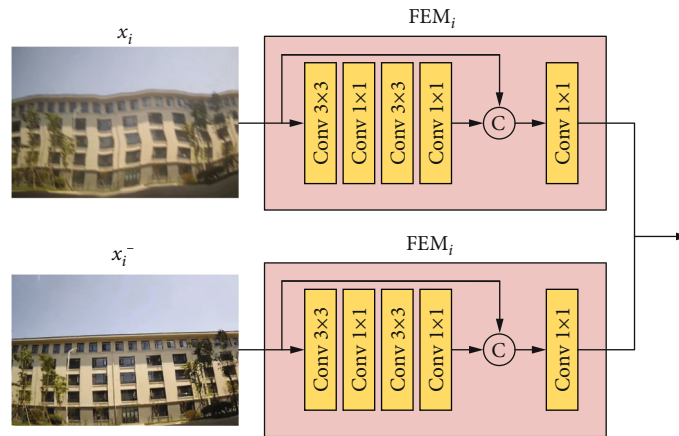


FIGURE 2: PSFEAM overall architecture.

pooling operation and two 1×1 convolutional layers for generating $C \times 1$ attention maps given different scales of feature maps $C \times H \times W$. One CAM architecture is shown in Figure 4.

The outputs of all CAMs are connected to SAM. The SAM integrates the raw encoder features and CAM features

through a two-branch convolution. In one branch, the global average pooling is performed and the convolutional layers are combined to obtain a global information feature map. The other branch is global maximum pooling, which performs convolutional layer combination to obtain significant geometric line information feature maps. In the spatial

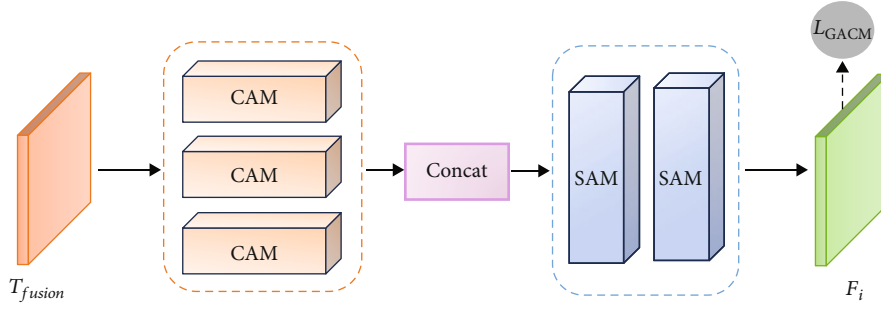


FIGURE 3: GACM overall architecture.

attention module, MaxPool is able to highlight important features in water-air cross-media images, while AvgPool is able to extract overall information. The parallel combination of the two can balance the needs of highlighting local details and overall scene features. At the same time, the use of parallelism can obtain more feature representation without adding too much extra computational burden. In the total pooling branch, the weight map is generated by sigmoid to enhance the geometric line regions. The final geometric line attention feature map is output. To localise the distorted geometric lines in the water-air cross-media image, two SAM outputs are applied as additional line supervision. One SAM architecture is shown in Figure 5.

The feature map $F_i (i = 1, 2, 3)$ is output after geometric line attention correction module. For the significantly different feature points labelled in the positive-negative sample feature extract attention module, the

$$s_{\text{sim}}(l', l) = \left\langle \frac{l'}{\|l'\|}, \frac{l}{\|l\|} \right\rangle, \quad (2)$$

where $s_{\text{sim}}(l', l)$ is the similarity between the distortion geometric line l' in the positive sample and the object restoration line l in the negative sample. Geometric lines in water-air cross-media images are corrected by comparing lines in positive and negative samples to adjust the correction parameters. By weighing the attentional weights of the distorted geometric lines in softmax based on $s_{\text{sim}}(l', l)$, the attention level of each distorted geometric line can be expressed as

$$s_{\text{sim}}^l(l', l) = \kappa \text{softmax}_{\text{sim}}(l', l), \quad (3)$$

where κ is a constant.

The geometric line attention correction module loss L_{GACM} is computed using the Sobel operator from which the x_i^- corresponding edges are extracted as ground-truth X_{edge}^- .

$$L_{\text{GACM}} = \left\| F_i - X_{\text{edge}}^- \right\|_2. \quad (4)$$

L_{GACM} is the difference between the GACM output F_i and the ground-truth X_{edge}^- .

3.3. Blurring Artifact Elimination Module for Water-Air Cross-Media Images. There were fluid irregularities in the water-air cross-media environment that caused artifacts or blurring in the image. The geometric line attention correction module also leads to feature residuals at the boundary of the correction site. In this chapter, a blur artifact elimination module (BAEM) is proposed for water-air cross-media images. The method uses feature maps of different scales to eliminate the blur artifacts in the image through the multiscale feature fusion module (MFFM), so as to achieve efficient multiscale deblurring. The overall process of the blur artifact elimination module is shown in Figure 6.

The high-level network has a relatively large receptive field and strong ability for characterization of semantic information, but the feature map has low resolution and weak characterization of geometric information (lack of spatial geometric feature details). The low-level network has a relatively small receptive field and strong geometric detail information characterization ability, and although the resolution is high, the semantic information characterization ability is weak. At this point, it is necessary to integrate the network characteristics and perform multiscale feature fusion.

The multiscale feature fusion module performs feature fusion across different scales. This enables the flow of information from various scales within a single U-Net model. The whole network contains two MFFM; MFFM₁ is the first level of multiscale feature fusion, and MFFM₂ is the second level of multiscale feature fusion. Specifically, the MFFM formula is as follows:

$$\begin{aligned} \text{MFFM}_1^{\text{out}} &= \text{MFFM}_1 \left(F_1^{\text{out}}, (F_2^{\text{out}})^{\uparrow}, (F_3^{\text{out}})^{\uparrow} \right), \\ \text{MFFM}_2^{\text{out}} &= \text{MFFM}_2 \left((F_1^{\text{out}})^{\downarrow}, F_2^{\text{out}}, (F_3^{\text{out}})^{\uparrow} \right), \end{aligned} \quad (5)$$

where $\text{MFFM}_n^{\text{out}}$ denotes the output of the n th MFFM. Applying upsampling (\uparrow) and downsampling (\downarrow) allows features from different scales to be connected.

The loss function of the blur artifact elimination module is computed using a multiscale content loss function that employs the Euclidean distance as a distance metric between feature vectors. The loss function L_{BAEM} is defined as follows:

$$L_{\text{BAEM}} = \sum_{k=1}^K \frac{1}{t_k} \|X - X^-\|^2, \quad (6)$$

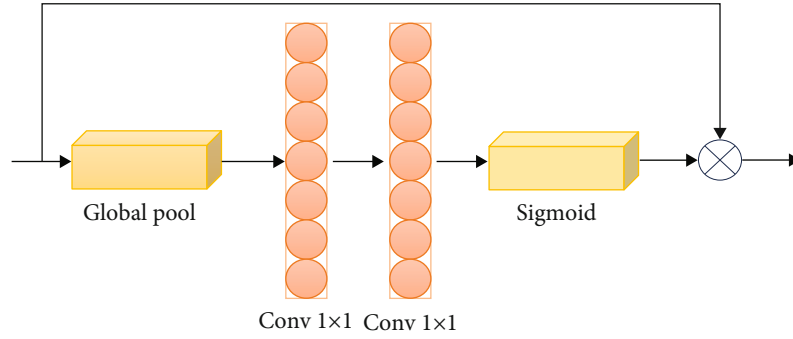


FIGURE 4: One CAM architecture.

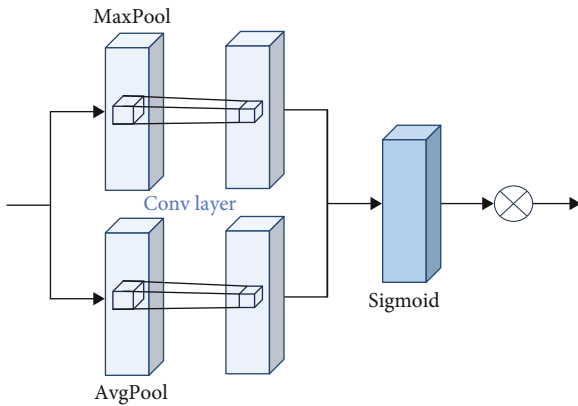


FIGURE 5: One SAM architecture.

where K is the number of levels. Divide the loss by the number of total elements t_k for normalisation. Additionally, the use of the squared error as the distance measure in the loss function aims to enhance the dissimilarities between different images in the feature space, thereby promoting a more stable training process. The squared error helps mitigate the impact of noise, image distortion, and other factors that may interfere with the Euclidean distance measurement, thus improving the model's robustness. Moreover, employing squared error simplifies mathematical computations, rendering it more convenient for calculations.

The final loss function for training the network is determined as follows:

$$L_{\text{total}} = L_{\text{GACM}} + \lambda L_{\text{BAEM}}, \quad (7)$$

where λ denotes a scaling factor for controlling the loss ratio of the geometric line attention correction module and the blur artifact removal module.

4. Experiment

4.1. Experimental Setup

4.1.1. Experimental Environment. In this experiment, training, validation, and testing were performed on a small server with Intel® Core™ i7-1165G7 CPU, RTX 3090 GPU, and 64G RAM. In order to reflect the objectivity of the proposed method in the comparison experiments, a U-Net

convolutional neural network was used for deep network construction using Visual Studio Code deep learning tool implementation.

4.1.2. Dataset. The main research problem of this paper is to address the challenges posed by object distorted images in water-air cross-media using a multiscale feature attention approach. Since there is no publicly available dataset of water-air cross-media scenes, the data used in this paper is a self-made dataset generated by taking water-air cross-media scene images from an underwater camera outward in a transparent pool (considering the problem of media variation in transparent pools, the aberrated images from this shot did not pass through the media of the transparent pool). The camera shoots in air to obtain a ground-truth image, which is the negative sample. The camera shoots in water to get a water-air cross-media figure, which is the positive sample. The turbidity is due to the fact that when the camera is moving, the camera is near the bottom of the water where the sediment is, and the camera is near the surface of the water where the turbidity is lower.

In order to evaluate the image restoration performance of the proposed methods for water-air cross-media scenes, all the methods are tested on the water-air cross-media distorted image dataset (WCDID). Since the shooting environment is daytime with fine weather, the applicability of the dataset and the range of applications are considered. In order to make the proposed object aberration image restoration network under water-air transmedia better trained, the data is augmented by data augmentation (simulating the shooting time by augmenting the brightness and darkness and augmenting the rotated images to increase the dataset of the object under water-air environment). The data is split into a training set, test set, and validation set with a ratio of 7:2:1, based on a training set of 1407 images. After applying data augmentation, the image data is expanded fivefold, resulting in a total of 10050 images, including 7035 training images, 2010 test images, and 1005 validation images. Within the self-made dataset, there are 3765 water-air cross-media images in conventional environments, 3290 in low turbidity environments, and 2995 in high turbidity environments. In the test datasets, there are 895 water-air cross-media pictures in conventional environments, 575 in low turbidity environments, and 540 in high turbidity environments.

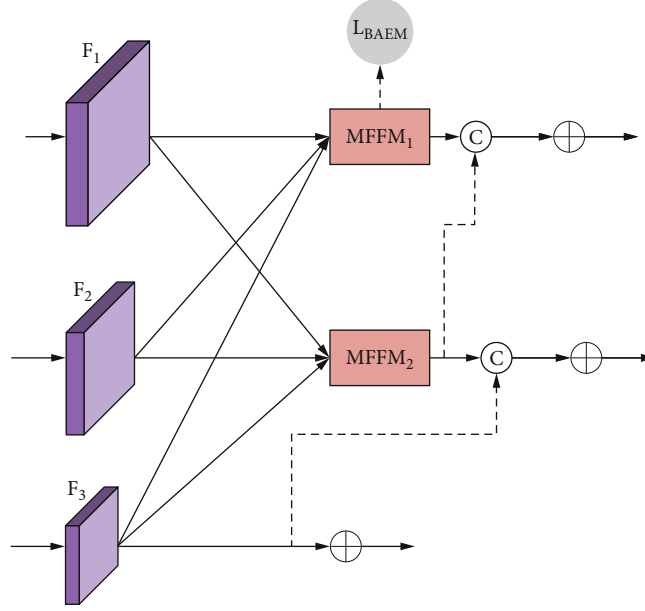


FIGURE 6: BAEM architecture.

The dataset in this paper has been filtered appropriately according to the water-air cross-media scenario, and the data quality is relatively high. At the same time, the data are effectively augmented to reach a certain quantity in the relatively high data, so the WCDID dataset meets the requirements of the water-air cross-media environment background and has a good representation in both “quality” and “quantity.”

4.1.3. Learning Rates and Training Settings. Usually, the initial learning rate is set to a smaller value, and then, the learning rate is dynamically adjusted during the training process. We train the network model using ground-truth images and water-air cross-media images as inputs to the network model. The algorithm proposed in this paper sets the initial learning rate to 0.0001 during the training process and uses poly to adjust the dynamic learning rate. In order to prevent the model from oscillating or overfitting during the training process, the decay is set to 0.0005 in the paper, and the stochastic gradient descent method is used to gradually decrease the learning rate with the number of training steps for training. Each iteration of the training input sample batch size is 8, and the number of epochs is 300. In the initialisation training stage of the model, the training process will appear local optimum without reaching the global optimum. In this paper, setting momentum to 0.9 can solve the problem to some extent.

4.1.4. Evaluation Indicators. There is no single evaluation metric that is perfect for assessing this scenario when evaluating water-air cross-media image restoration. After reviewing a large amount of relevant literature, we chose to use evaluation metrics that are commonly used in existing research in this field. In order to evaluate the performance of the proposed algorithm, the evaluation metrics of the proposed algorithm are chosen as mean square error (MSE),

peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and line straightness metric (LineAcc) [38].

This MSE is an intuitive metric that evaluates the differences between images by calculating the squared difference between their pixel values. The MSE is a common measure of how close the restored image is to the ground-truth image and is given by the following formula:

$$\text{MSE} = \frac{\sum_{(x,y)} (G(x,y) - D(x,y))^2}{t \times d}, \quad (8)$$

where G is the ground-truth image, D is the restored image, and $t \times d$ is the size of each image.

The PSNR is an expression of the ratio between the maximum possible value (power) of the signal and the power of the distorted noise that affects the quality of its representation and is given by the following equation:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{(\text{MAX}(G))^2}{(1/(t \times d)) \sum_{x=1}^t \sum_{y=1}^d (G(x,y) - D(x,y))^2} \right). \quad (9)$$

SSIM is an image quality metric for estimating the visual impact of variations in image brightness, contrast, and structure and is formulated as follows:

$$\text{SSIM} = \frac{(2\mu_R\mu_G + C_1)(2\sigma_{RG} + C_2)}{(\mu_R^2 + \mu_G^2 + C_1)(\sigma_R^2 + \sigma_G^2 + C_2)}, \quad (10)$$

where μ_R , μ_G , σ_R , σ_G , and σ_{RG} are the local mean, standard deviation, and intercovariance of the ground-truth image G

and the restored image D . C_1 and C_2 are constants that can be defined as

$$\begin{aligned} C_1 &= (0.01 \times L)^2, \\ C_2 &= (0.03 \times L)^2, \end{aligned} \quad (11)$$

where L is the specified dynamic range value.

LineAcc evaluates the change in curvature of the marker line (the line that marks the significant line in the test dataset). The specific LineAcc algorithm is shown in Figure 7.

The formula is as follows:

$$\text{LineAcc} = 1 - \frac{1}{n} \sum_{i=0}^{n-1} \left(\frac{y_{d_i} - y_{d_{i-1}}}{x_{d_i} - x_{d_{i-1}}} - \frac{y_{g_o} - y_{g_n}}{x_{g_o} - x_{g_n}} \right), \quad (12)$$

where LineAcc indicates the similarity between the slopes of these two lines and n is the number of uniformly sampled points in each line. (x_{g_i}, y_{g_i}) and (x_{d_i}, y_{d_i}) indicate the coordinates of the corresponding points in the reference and distorted images.

4.2. Image Restoration for Distortion of Object in Water-Air Cross-Media

4.2.1. Comparative Experiments on Public Datasets. This dataset is an air-water vision dataset produced by Li et al. [39] based on the ImageNet dataset, where a computer monitor is placed under a glass jar which is filled with about 13 cm of water, a small stirring pump is used to keep the water in motion, and a large dataset of distorted and undistorted image pairs is constructed by capturing images of ImageNet displayed under the water surface. For experiments on the ImageNet dataset, we selected 8250 images, divided the dataset according to 7:2:1, and selected 5775 images for training and 825 images for testing.

For the cross-media distorted image restoration experiments, in order to evaluate the performance of the proposed algorithm in conventional scenarios, the algorithm of this paper is compared with the more advanced methods proposed by the existing AWTFFNet [40], UnfairGAN [41], LPF [42], and CARNet [43]. The mean square error, peak signal-to-noise ratio, structural similarity, and time for image restoration of this paper's algorithm and the compared algorithms are shown in Figure 8.

From the visual effect Figure 8, it can be seen that different algorithms have repaired the distorted parts of the image, but the correction effect for the geometric lines is not very obvious, and the algorithm proposed in this paper is better than the other algorithms for the geometric line correction. The blurring of the letters in the figure is not obvious, but from the results, the algorithm proposed in this paper can vaguely see the letters, proving that the performance of the algorithm proposed in this paper is better than other algorithms on the ImageNet dataset.

It can be seen from Table 1 that, on the MSE and PSNR metrics, UnfairGAN performs the best, while LPF performs the worst. The superiority of the UnfairGAN maintains the consistency and coherence of the generated image in the case

of image clarity, which makes the repaired region and the surrounding real image more integrated and matched. The ability of the network to generate high-quality, high-detail images is enhanced by using a complex network structure and employing a specific loss function during the training process. On SSIM metric, our method performs the best, while LPF performs the worst. On the LineAcc metric, our method performs the best while LPF performs the worst. From the results of these metrics, our proposed method shows better performance overall. Our method achieves better results in terms of images quality (PSNR and SSIM) and line accuracy (LineAcc) and is able to better recover the details and geometric lines of the water-air cross-media images relative to other methods. However, it is also important to note that although our method shows better results in some metrics, there is still some room for improvement. For example, our method is slightly inferior to UnfairGAN in the MSE and PSNR metrics, which may be caused by the less than perfect processing of certain details in our method. Therefore, for further improvement of image quality and accuracy, we can continue to optimize and improve our method.

4.2.2. Doing Comparative Experiments on Self-Made Datasets

(1) Analysis of the Results of Experimental Comparisons in Routine Situations. For the cross-media aberration image restoration experiments, the algorithm of this paper is compared with other algorithms in order to evaluate the performance of the proposed algorithm in conventional scenarios. The MSE, PSNR, SSIM, and LineAcc of image restoration of this paper's algorithm and the compared algorithms are shown in Figure 9.

From the visual effect (Figure 9), it can be seen that this paper's self-made datasets repair works better, and other algorithms may be effective in some specific cases but may not generalize well for complex scenes. The proposed method performs better in terms of sharpness and contrast, retains more details of the water-air cross-media scene, and better guides the distorted images to be repaired at sharp boundaries or regions, further proving the effectiveness of the proposed algorithm.

From Table 2, it can be seen that UnfairGAN and our proposed method perform the best and LPF performs the worst in terms of MSE metric. This indicates that UnfairGAN and our method can reduce the error in reconstructing the images and better recover the details in the water-air cross-media images. However, the LPF uses the lucky-patch search strategy based on copying and pasting of image content, which relies on a large amount of image data. So, it does not work well in conditions where the number of datasets is limited. In terms of PSNR and SSIM metrics, our method performs the best while LPF performs the worst. The higher PSNR and SSIM values indicate that our method is relatively closer to the original image in terms of reconstructed image quality and can maintain good structural similarity. In the LineAcc metric, our method performs the best while LPF performs the worst. This indicates that our

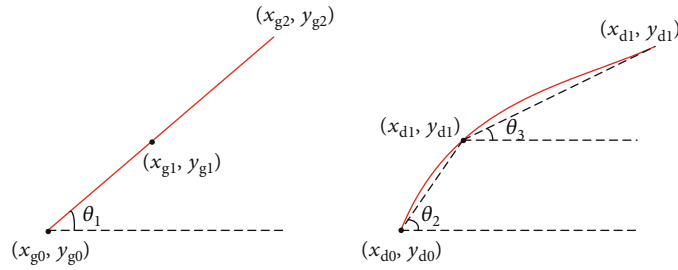


FIGURE 7: Schematic diagram of the LineAcc algorithm.

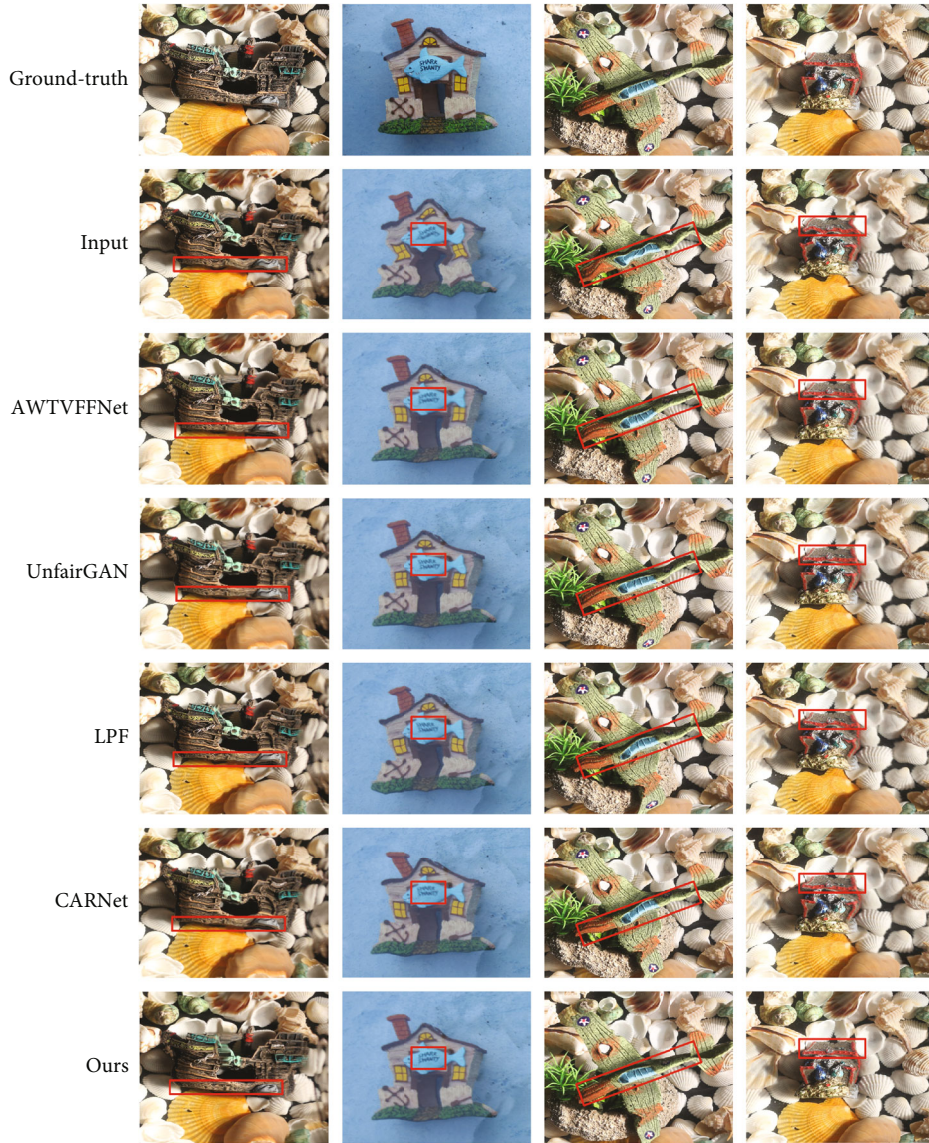


FIGURE 8: Comparison experimental effect diagram of ImageNet dataset.

method can better recover geometric lines in water-air cross-media images and improve line accuracy. Taken together, our method achieves better results in several metrics and can effectively solve the distortion problem in water-air cross-media images. However, further research and improvement of the method are still needed to further enhance the performance and adaptability.

(2) *Analysis of the Results of Comparative Experiments in Slightly Turbid Water Environments.* Due to the presence of a large amount of sediment in the scene where the water-air is located, the larger particles of sand will sink to the bottom of the water due to gravity when the water surface fluctuates relatively small, and a small portion of gravel will be suspended in the water resulting in slightly turbid

TABLE 1: Quantitative evaluation of comparison experiments on ImageNet dataset.

Method	MSE	PSNR	SSIM	LineAcc
AWTVFFNet	828.42	23.4936	0.5572	32.483
UnfairGAN	786.54	23.6897	0.5837	31.856
LPF	855.19	21.1439	0.4795	31.054
CARNet	831.47	22.2634	0.5271	32.758
Ours	799.56	23.5741	0.5906	33.125

Bolded black texts are the best data.

water quality. In order to further evaluate the performance of the proposed algorithm under turbid water, the algorithm in this paper is compared with other algorithms. The MSE, PSNR, SSIM, and LineAcc of image restoration of this paper’s algorithm and the compared algorithms are shown in Figure 10.

As can be seen from the visual effect (Figure 10), although the image restoration is carried out in the case of mildly turbid water quality, the contour boundaries in the image can be seen more clearly. The geometric line attention correction module proposed in this paper still shows better results, proving the effectiveness of the proposed algorithm in the turbid water quality and water surface fluctuation environment.

As can be seen from Table 3, the UnfairGAN and our proposed method perform better on the MSE metric, while LPF performs the worst. In terms of PSNR and SSIM metrics, UnfairGAN and our method perform better, while LPF performs the worst. In LineAcc metric, our method performs better while LPF performs worst. We note that the AWTVFFNet method is in a relatively stable state and the PSNR and SSIM metrics are relatively good due to the fact that the AWTVFFNet network performs adaptive weight assignment based on the structural features of different regions in the image by introducing anisotropic weighting technique. This helps in better processing of details and textures in the image. However, its parameters need to be carefully selected and tuned in order to be applicable to water-air cross-media scenarios. This indicates that our method can better recover geometric lines in water-air cross-media images and improve line accuracy. In the case of low water turbidity and slight blurring of the images, the algorithms in this paper demonstrate the remarkable performance of their image restoration. Taken together, our method shows better results on several metrics and can effectively solve the distortion problem in water-air cross-media images.

(3) *Analysis of the Results of Comparative Experiments in Environments with High Water Turbidity.* Due to the presence of water surface fluctuations in the scene where the water-air is located, larger particles of sand will be rolled up and suspended in the water, thus forming turbid suspended matter in the water, resulting in the water-air cross-media pictures taken appearing to have a high degree of turbidity. In order to further evaluate the restoration performance of the proposed algorithm under water turbidity, the algorithm in this paper is compared with other algo-

rithms. The MSE, PSNR, SSIM, and LineAcc of image restoration of this paper’s algorithm and the comparison algorithms are shown in Figure 11.

As can be seen in the visual effect (Figure 11), there are some challenges in dealing with the complex image restoration task. The poor restoration effect of the image is due to the turbidity of the water causing the extracted features to be blurred, the extracted geometric line information is weak, and the effect restoration is not obvious for the heavy shadow part. However, from the last line of the effect image, it can be seen that the algorithm proposed in this paper is clearer than other algorithms in repairing the boundary of the line parts, which proves the effectiveness of the proposed algorithm in the turbid water environment.

From Table 4, it can be seen that LPF performs the worst on the MSE metric, while UnfairGAN and our proposed method perform better. In terms of PSNR and SSIM metrics, our method performs relatively well, while LPF performs the worst. In LineAcc metric, our method performs better while LPF performs worst. In the environment of large fluctuations in the water surface and high turbidity of the water, the content of sediment is high, the color of the images changes, and the color contrast is weakened, which leads to poor results when extracting features. Taken together, although our method is degraded in the image quality index, our correction module still reflects its superiority. In the future, we will focus on the image quality aspect in order to effectively solve the distortion problem of water-air cross-media images.

Under high turbidity conditions, “UnfairGAN” performs well on all metrics except “LineAcc.” To better illustrate the value of the method proposed in this paper, we provide a comparison of the training time with other methods.

As seen in Figure 12, the method proposed in this paper shows better performance in terms of training time. This is due to the utilization of skip connections, which provide additional paths for gradient propagation. This helps in training the deep network more effectively, enhancing training efficiency and stability.

4.2.3. *Ablation.* An ablation experiment is conducted to verify the effectiveness of geometric line attention correction module (GACM) and blur artifact elimination module (BAEM) in the algorithm proposed in this paper. The experiment is performed on a self-made water-air cross-media image dataset. U-Net is used as the baseline for the ablation experiment. The experimental visual results are shown in Figures 13, and the results are shown in Table 5.

As the geometric line attention correction module can correct and enhance the geometric lines in the image, it makes the geometric structure in the image more obvious and accurate, improves the quality and visual effect of the image, and makes the image easier to understand and analyze. The image blur and artifact removal nodule not only removes blur and artifacts in the image by combining different scales of image information but also removes the residuals in the boundary area brought about by the geometric line attention correction module, which improves the clarity and details of the image, making the image clearer, sharper,

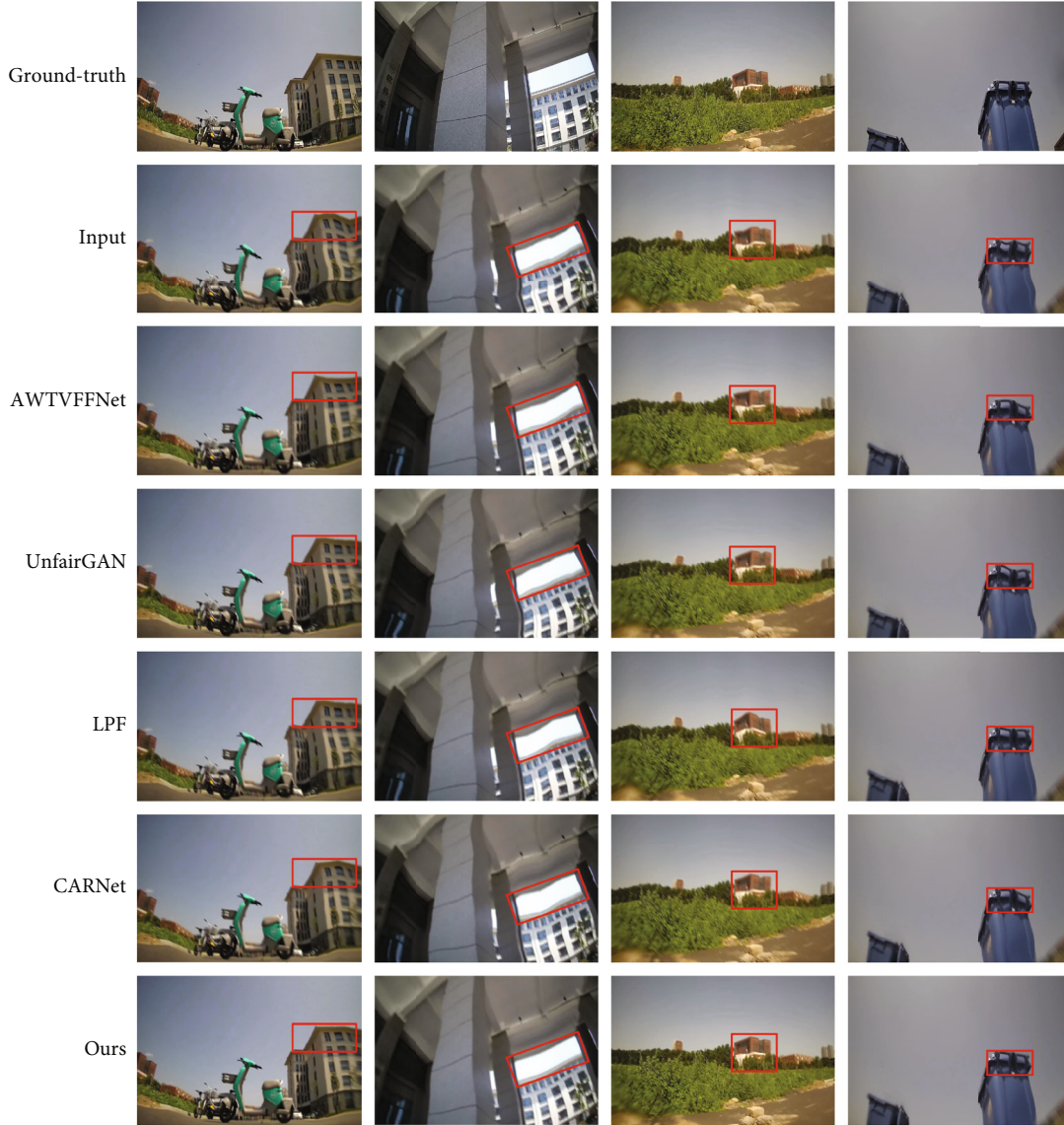


FIGURE 9: The effect of the image in the regular case.

TABLE 2: Quantitative evaluation of images in conventional situations.

Method	MSE	PSNR	SSIM	LineAcc
AWTVFFNet	829.56	23.4243	0.5811	32.115
UnfairGAN	789.15	23.6927	0.5895	31.554
LPF	843.91	21.9415	0.4878	31.428
CARNet	834.02	22.7726	0.5094	32.927
Ours	796.77	23.9278	0.5949	33.030

Bolded black texts are the best data.

and easier to observe and analyze. The combination of the two not only improves the evaluation indexes but also enables the repair of more complex scenes.

As evident in Table 5, the addition of GACM and BAEM to the baseline individually resulted in significant improve-

ments across all indicators compared to the baseline. However, the highest quality metrics were obtained when both GACM and BAEM were added together. This can be attributed to the geometric line correction module, which corrects images by reducing the distance between positive and negative sample features. Additionally, the multiscale fusion single U-Net model effectively handles image features at different scales through the introduction of multiscale information and fusion operations, leading to better artifact removal. The combination of these two techniques not only enhances the evaluation indexes but also enables the repair of more complex water-air cross-media scenes.

5. Discussion

In the case of water-air transboundary situations, refraction and reflection of light propagating between water and air result in deformation of the image. This poses a challenge

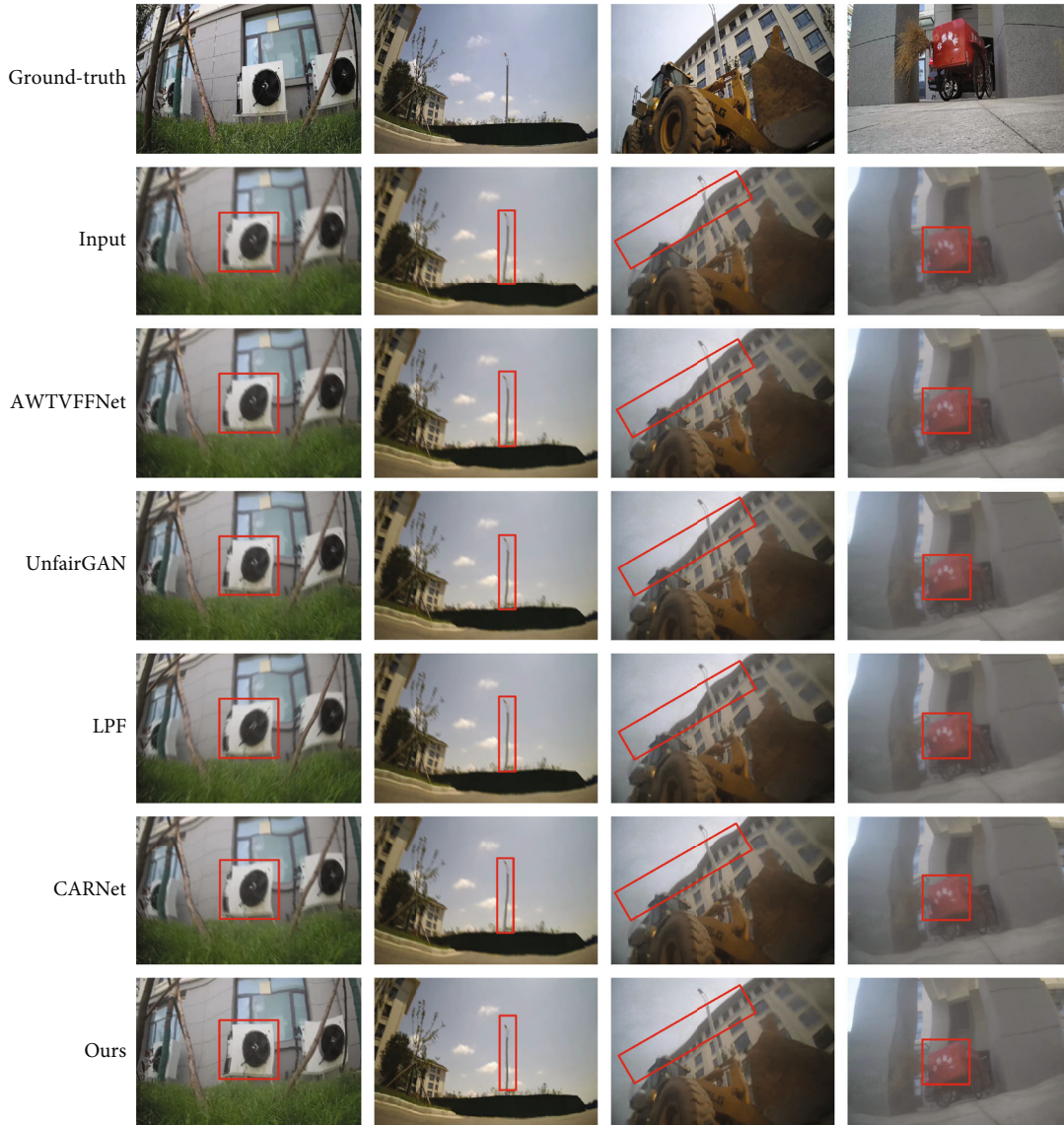


FIGURE 10: The effect of the image in a slightly turbid water environment.

TABLE 3: Quantitative evaluation of images in slightly turbid water environments.

Method	MSE	PSNR	SSIM	LineAcc
AWTVFFNet	834.83	22.5275	0.5736	30.412
UnfairGAN	795.24	22.4501	0.5844	30.807
LPF	857.62	21.7234	0.4778	30.051
CARNet	848.40	22.0536	0.5054	31.199
Ours	801.71	22.5967	0.5868	31.274

Bolded black texts are the best data.

for the exploration of the marine environment. In such cases, the image of the object over water is usually reduced and decreased, which means that the size of the object is smaller and its position is decreased compared to that in the water. This is due to the high refractive index of water, which bends the light as it exits the water, making the

object appear smaller. In addition, the refraction of light causes the position of the object to appear to drop. However, it should be noted that factors such as water ripples and distortions are often present at the water-air interface, which can further affect the degree of image distortion. At the same time, light propagation in water is also affected by factors such as water turbidity and immersion depth. For object-distorted images under water-air cross-media, researchers can analyze and study them by using optical principles and image processing techniques to realize the correction and correction of distorted images [44]. These researches are of great significance to the fields of underwater observation, underwater photography, and underwater image processing. Therefore, researchers need to comprehensively consider the influencing factors in different media in order to seek effective image restoration methods to improve the accuracy of recognizing and recovering object features in the marine environment.

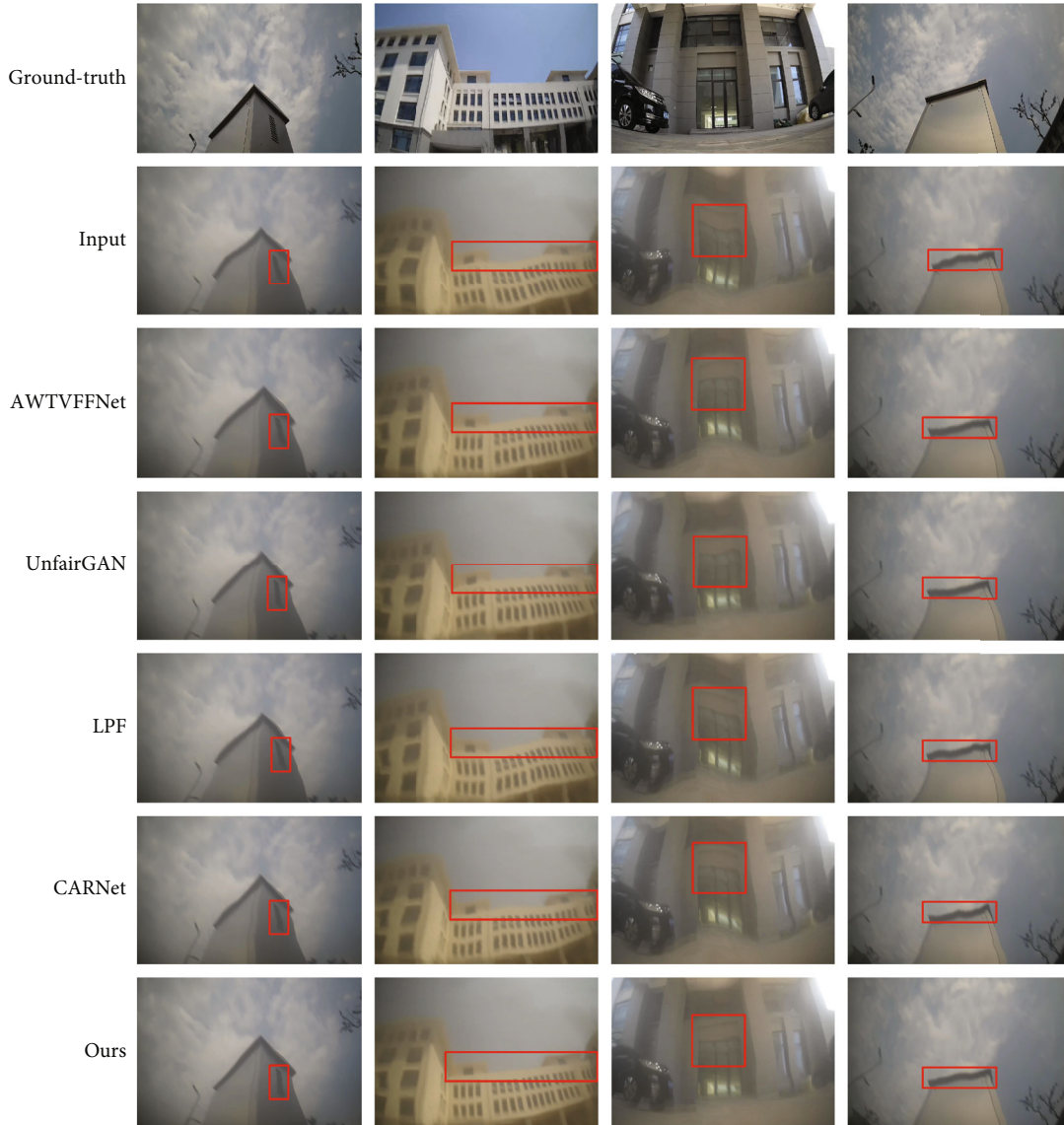


FIGURE 11: The effect of the image under the environment of high turbidity of water quality.

TABLE 4: Quantitative evaluation of images in environments with high water turbidity.

Method	MSE	PSNR	SSIM	LineAcc
AWTVFFNet	889.52	22.2069	0.5724	30.370
UnfairGAN	813.73	22.2519	0.5826	30.715
LPF	905.91	21.2177	0.4437	28.127
CARNet	866.47	21.9323	0.5028	30.913
Ours	818.24	22.2154	0.5819	30.974

Bolded black texts are the best data.

Image restoration of object aberrations in water-air cross-media is a complex and challenging task. To solve this problem, researchers have carried out a series of works. First, they analyzed and understood the aberration mechanism of light in the water-air cross-media process through optical principles and physical models. This helps to reveal the

causes and characteristics of image aberrations and provides a theoretical basis for subsequent restoration methods. Second, the researchers utilize image processing techniques and computer vision algorithms to realize the restoration of the object distorted image. This includes removing or correcting deformations and distortions in the image and restoring the true shape and position of the object. Commonly used restoration methods include image alignment based on feature point matching [42, 45], deformation modeling and correction [46], and filtering algorithms to remove water surface fluctuations. These methods are aimed at transforming and filtering the image according to the distortion features, making the restored image reflect the appearance of the original object more accurately. In addition, to improve the restoration results, researchers often utilize advanced machine learning and deep learning techniques. They train neural network models to recognize and restore distorted object images by using large-scale training

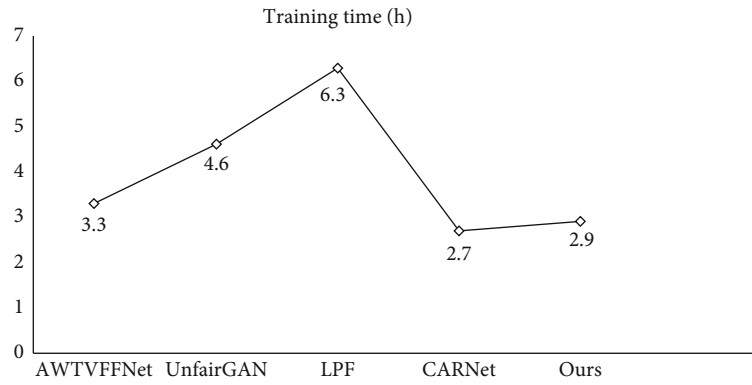


FIGURE 12: Comparison figure of training time at high turbidity.

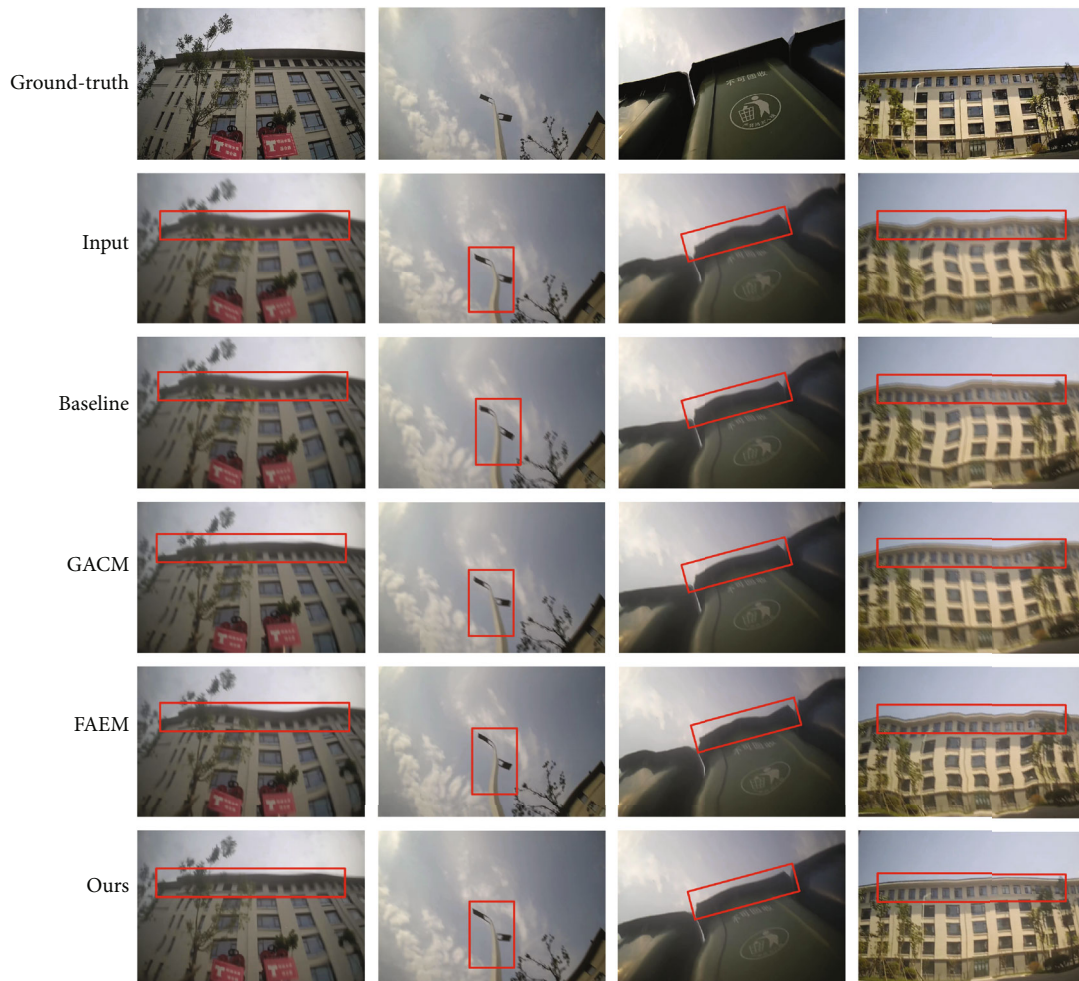


FIGURE 13: Experimental effect of ablation of distorted image in water-air cross-media.

datasets. This approach better captures features and details in the image and generates more accurate restoration results. Finally, researchers need to verify the effectiveness and accuracy of the restoration method through experimentation and evaluation. They will use simulated data or actual captured image data [47] to perform quantitative and qualitative evaluations to compare the image quality before and after restoration to assess the performance of the restoration method.

In the future, image restoration of object aberrations in water-air cross-media will continue to see development and improvement. With the application of deep learning technology, deep learning has made great progress in the field of image processing, which can automatically learn the features and patterns of images by learning from a large amount of training data. The application of deep learning techniques to image restoration of object aberrations in

TABLE 5: Quantitative evaluation of ablation experiments of aberrated images in water-air cross-media.

GACM	BAEM	MSE	PSNR	SSIM	LineAcc
		1046.38	20.1650	0.3985	29.072
√		854.27	21.9718	0.5714	31.825
	√	866.41	21.9157	0.5676	30.473
√	√	799.17	22.3546	0.5933	32.104

Bolded black texts are the best data.

water-air cross-media can further improve the accuracy and robustness of the restoration algorithms. Utilizing multi-modal data fusion (e.g., LiDAR, sonar, and optical images) in conjunction with sensor calibration techniques can provide more comprehensive and accurate object information. Methods for fusing multiple data sources can be used to further improve the restoration of object distortion images. The development of more accurate and adaptable physical models can better characterize the deformation of water-air cross-media. These models can more accurately capture and repair aberrations in images and provide a more reliable basis for restoration algorithms. Efficient real-time restoration algorithms are designed for the application needs in the fields of underwater observation and underwater photography. These algorithms should have the ability to process images quickly and perform repair operations quickly and accurately in real-time scenes. The adaptive restoration methods that can automatically adjust restoration strategies according to different water quality, water conditions, and observation environments were introduced. This can better adapt to the needs of object aberration image restoration in different environments and improve the adaptability and robustness of restoration. When studying scenes in water-air cross-media images, it is very important to consider air-water cross-media scenes at the same time. Such comparative and contrasting studies can help us to better understand the differences in optical properties and object representations between different media, providing us with a more comprehensive perspective, as well as more valuable information for research and applications in related fields. Computer vision, optical imaging, and deep learning are comprehensively applied to establish suitable models and algorithms to solve the problems of target object recognition, depth estimation, and localisation after image restoration. This will help improve the accuracy and reliability of underwater target detection and imaging and expand the research and application prospects in related fields. With the continuous progress of technology and the introduction of new methods, object distortion image restoration under water-air cross-media will face more opportunities and challenges. This will drive the improvement of restoration algorithms and provide more accurate and reliable image restoration tools for underwater observation, underwater photography, and underwater image processing.

6. Conclusions

In amphibious robotic water-air cross-media reconnaissance, the complex environment, such as inhomogeneous

media and irregular fluid flow, can cause changes in optical conditions, leading to image distortions and aberrations. These factors can make it difficult for amphibious robots to detect and identify objects in the sea-land junction area. To address these challenges, we propose an aberration image restoration method for underwater and aerial cross-media images. First, we extract same-scale features from positive and negative samples of water-air cross-media images using convolutional combinations. We then perform feature fusion on feature points that show smaller differences between positive and negative samples at the same scale. For feature points with larger differences, we apply attention labeling. Additionally, we introduce an attention correction module for geometric lines, which corrects geometric lines in water-air cross-media images by comparing and sensing marked feature points with large differences. We utilize the line similarity in positive and negative samples for this correction. Finally, we employ a blurring artifact elimination module that uses multiscale fusion of individual U-Net information streams. This module eliminates artifacts caused by image blurring and geometric line correction. To evaluate the algorithm's feasibility, we conducted a comparison experiment using the public ImageNet dataset. The LineAcc metric showed an 8.6% improvement. We also conducted experiments in three different environments (conventional, slightly turbid water, and high turbidity water) using the homemade dataset WCDID. The results demonstrated that the method performed well under various water-air cross-media conditions, achieving LineAcc improvements of 10.3%, 7.5%, and 6.1% in the respective environments. Ablation experiments were also conducted to prove the significant effects of the proposed geometric line attention correction module and blur artifact elimination module on image restoration in water-air cross-media. This study provides an effective solution to image distortion in water-air cross-media, enhancing image quality and geometric accuracy. It holds significant practical value for amphibious robots in detecting and identifying objects at the land-sea interface area.

Data Availability

The data used to support the results of this study are available from the first author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the Henan Provincial Key Research and Development Program (231111220700), the Major Science and Technology Special Project of Henan Province (221100110500), and the Science and Technology Tackling Project of Henan Province (232102320338).

References

- [1] A. Sekkat, Y. Dupuis, V. Kumar et al., "SynWoodScape: synthetic surround-view fisheye camera dataset for autonomous driving," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8502–8509, 2022.
- [2] K. Zhao, C. Lin, K. Liao, S. Yang, and Y. Zhao, "Revisiting radial distortion rectification in polar-coordinates: a new and efficient learning perspective," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 6, pp. 3552–3560, 2022.
- [3] W. Fan, G. Xu, Q. Wu et al., "Azure Kinect calibration and parameter recommendation in different scenarios," *IEEE Sensors Journal*, vol. 22, no. 10, pp. 9733–9742, 2022.
- [4] Y. Zhang, X. Zhao, and D. Qian, "Learning-based distortion correction and feature detection for high precision and robust camera calibration," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10470–10477, 2022.
- [5] P. Xu, J. Zheng, X. Wang et al., "Design and implementation of lightweight AUV with multisensor aided for underwater intervention tasks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 12, pp. 5009–5013, 2022.
- [6] L. Zhou, W. Min, D. Lin, Q. Han, and R. Liu, "Detecting motion blurred vehicle logo in IoV using filter-DeblurGAN and VL-YOLO," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3604–3614, 2020.
- [7] P. Duan, Z. Wang, B. Shi, O. Cossairt, T. Huang, and A. K. Katsaggelos, "Guided event filtering: synergy between intensity images and neuromorphic events for high performance imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8261–8275, 2021.
- [8] B. Zheng, S. Yuan, C. Yan et al., "Learning frequency domain priors for image demoiring," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7705–7717, 2022.
- [9] C. Tao, Y. Meng, J. Li et al., "MSNet: multispectral semantic segmentation network for remote sensing images," *GIScience & Remote Sensing*, vol. 59, no. 1, pp. 1177–1198, 2022.
- [10] J. Qiu, S. Lai, G. Huang et al., "TriView-ParNet: parallel network for hybrid recognition of touching printed and handwritten strings based on feature fusion and three-view co-training," *Applied Intelligence*, vol. 53, no. 13, pp. 17015–17034, 2022.
- [11] B. Xie, Z. Yang, L. Yang, A. Wei, X. Weng, and B. Li, "AMMF: attention-based multi-phase multi-task fusion for small contour object 3D detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 1692–1707, 2022.
- [12] S. Qi, J. Jia, and L. Cai, "Underwater incomplete target recognition network via generating feature module," *International Journal of Distributed Sensor Networks*, vol. 2023, Article ID 5337454, 11 pages, 2023.
- [13] C. Wang, M. Shao, D. Meng, and W. Zuo, "Dual-pyramidal image inpainting with dynamic normalization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 9, pp. 5975–5988, 2022.
- [14] A. Chandrashekar, A. Handa, N. Shivakumar et al., "A deep learning pipeline to automate high-resolution arterial segmentation with or without intravenous contrast," *Annals of Surgery*, vol. 276, no. 6, pp. 1017–1027, 2020.
- [15] Z. Wang, N. Xu, B. Wang, Y. Liu, and S. Zhang, "Urban building extraction from high-resolution remote sensing imagery based on multi-scale recurrent conditional generative adversarial network," *GIScience & Remote Sensing*, vol. 59, no. 1, pp. 861–884, 2022.
- [16] X. Li, X. Xu, X. He, X. Wei, and H. Yang, "Intelligent crack detection method based on GM-ResNet," *Sensors*, vol. 23, no. 20, pp. 8369–8369, 2023.
- [17] D. K. Rakesh and P. K. Jana, "A general framework for class label specific mutual information feature selection method," *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 7996–8014, 2022.
- [18] W. Lambert, L. A. Cobus, J. Robin, M. Fink, and A. Aubry, "Ultrasound matrix imaging—part II: the distortion matrix for aberration correction over multiple isoplanatic patches," *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3921–3938, 2022.
- [19] M. Jian, X. Liu, H. Luo, X. Lu, H. Yu, and J. Dong, "Underwater image processing and analysis: a review," *Signal Processing: Image Communication*, vol. 91, article 116088, 2021.
- [20] Y. Gao, C. Lin, Y. Zhao, X. Wang, S. Wei, and Q. Huang, "3-D surround view for advanced driver assistance systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, pp. 320–328, 2018.
- [21] L. Cai, Y. Li, C. Chen, and H. Chai, "Dynamic multiscale feature fusion method for underwater target recognition," *Journal of Sensors*, vol. 2022, Article ID 8110695, 10 pages, 2022.
- [22] S. Mozaffari, F. Feroldi, F. Larocca et al., "Retinal imaging using adaptive optics optical coherence tomography with fast and accurate real-time tracking," *Biomedical Optics Express*, vol. 13, no. 11, pp. 5909–5925, 2022.
- [23] S. Yang, C. Lin, K. Liao, C. Zhang, and Y. Zhao, "Progressively complementary network for fisheye image rectification using appearance flow," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6348–6357, Nashville, TN, USA, 2021.
- [24] F. Xu, H. Wang, Z. Liu, and W. Chen, "Adaptive visual servoing for an underwater soft robot considering refraction effects," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 12, pp. 10575–10586, 2020.
- [25] B. Kowalski, A. Vyas, and A. Dubra, "Correction of non-uniform angular velocity and sub-pixel jitter in optical scanning," *Optics Express*, vol. 30, no. 1, pp. 112–124, 2021.
- [26] X. Yang, Z. Liu, H. Qiao et al., "Graph-based registration and blending for undersea image stitching," *Robotica*, vol. 38, no. 3, pp. 396–409, 2020.
- [27] L. Decker and Q. Zhang, "Quantifying and correcting residual azimuthal anisotropic moveout in image gathers using dynamic time warping," *Geophysics*, vol. 85, no. 5, pp. 071–082, 2020.
- [28] Y. Tian and S. G. Narasimhan, "Seeing through water: Image restoration using model-based tracking," in *2009 IEEE 12th International Conference on Computer Vision*, pp. 2303–2310, Kyoto, Japan, 2009.
- [29] M. Ljubenic, A. Artesani, S. Bonetti, and A. Traviglia, "Beam-shape effects and noise removal from THz time-domain images in reflection geometry in the 0.25–6 THz range," *IEEE Transactions on Terahertz Science and Technology*, vol. 12, no. 6, pp. 574–586, 2022.
- [30] M. Jian, Q. Qi, J. Dong, Y. Yin, and K.-M. Lam, "Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection," *Journal of Visual Communication and Image Representation*, vol. 53, pp. 31–41, 2018.
- [31] J. Lee, H. Son, J. Rim, S. Cho, and S. Lee, "Iterative filter adaptive network for single image defocus deblurring," in

- Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2034–2042, Nashville, TN, USA, 2021.
- [32] J. Liu, J. Shang, R. Liu, and X. Fan, “Attention-guided global-local adversarial learning for detail-preserving multi-exposure image fusion,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 8, pp. 5026–5040, 2022.
- [33] W. Zhao, S. Rong, T. Li, J. Feng, and B. He, “Enhancing underwater imagery via latent low-rank decomposition and image fusion,” *IEEE Journal of Oceanic Engineering*, vol. 48, no. 1, pp. 147–159, 2023.
- [34] Z. Jiang, Y. Zhang, D. Zou, J. Ren, J. Lv, and Y. Liu, “Learning event-based motion deblurring,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3320–3329, Seattle, WA, USA, 2020.
- [35] L. Cai, X. Qin, and T. Xu, “EHDC: enhanced dilated convolution framework for underwater blurred target recognition,” *Robotica*, vol. 41, no. 3, pp. 900–911, 2023.
- [36] D. Zhang, J. Shen, J. Zhou, E. Chen, and W. Zhang, “Dual-path joint correction network for underwater image enhancement,” *Optics Express*, vol. 30, no. 18, pp. 33412–33432, 2022.
- [37] P. Liu, C. Zhang, Q. Hao, G.-y. Wang, and H. Zheng, “Multi-attention DenseNet: a scattering medium imaging optimization framework for visual data pre-processing of autonomous driving systems,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 25396–25407, 2022.
- [38] F. Zhu, Z. Shu, P. Wang, H. Wang, H. Yan, and S. Liu, “Semi-supervised wide-angle portraits correction by multi-scale transformer,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 19689–19698, New Orleans, LA, USA, 2022.
- [39] T. Li, Q. Yang, S. Rong, L. Chen, and B. He, “Distorted underwater image reconstruction for an autonomous underwater vehicle based on a self-attention generative adversarial network,” *Applied Optics*, vol. 59, no. 32, pp. 10049–10060, 2020.
- [40] H. Qi, S.-L. Tan, and Z. Li, “Anisotropic weighted total variation feature fusion network for remote sensing image denoising,” *Remote Sensing*, vol. 14, no. 24, 2022.
- [41] D. M. Nguyen, T. P. Le, D. M. Vo, and S.-W. Lee, “UnfairGAN: an enhanced generative adversarial network for raindrop removal from a single image,” *Expert Systems with Applications*, vol. 210, article 118232, 2022.
- [42] B. Jian, C. Ma, D. Zhu, Q. Huang, and J. Ai, “Water-air interface imaging: recovering the images distorted by surface waves via an efficient registration algorithm,” *Entropy*, vol. 24, no. 12, pp. 1765–1785, 2022.
- [43] M. Liu, L. Tang, L. Fan, S. Zhong, H. Luo, and J. Peng, “CAR-Net: context-aware residual learning for JPEG-LS compressed remote sensing image restoration,” *Remote Sensing*, vol. 14, no. 24, p. 6318, 2022.
- [44] Z. Li, Z. Murez, D. Kriegman, R. Ramamoorthi, and M. Chandraker, “Learning to See Through Turbulent Water,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 512–520, Lake Tahoe, NV, USA, 2018.
- [45] C. Cai, H. Meng, R. Qiao, and F. Wang, “Water-air imaging: distorted image reconstruction based on a twice registration algorithm,” *Machine Vision and Applications*, vol. 32, no. 3, 2021.
- [46] T. Sun, Y. Tang, and Z. Zhang, “Structural information reconstruction of distorted underwater images using image registration,” *Applied Sciences*, vol. 10, no. 16, 2020.
- [47] R. Zhang, D. He, Y. Li, L. Huang, and X. Bao, “Synthetic imaging through wavy water surface with centroid evolution,” *Optics Express*, vol. 26, no. 20, pp. 26009–26019, 2018.