

Article

Multi-Scale Adaptive Feature Network Drainage Pipe Image Dehazing Method Based on Multiple Attention

Ce Li ^{1,*} , Zhengyan Tang ¹, Jingyi Qiao ¹, Chi Su ^{2,3} and Feng Yang ¹¹ Computer Science and Technology, China University of Mining & Technology, Beijing 100083, China² Beijing Kingsoft Cloud Network Technology Co., Ltd., Beijing 100089, China³ SmartMore Co., Ltd., Beijing 100102, China

* Correspondence: celi@cumtb.edu.cn

Abstract: Drainage pipes are a critical component of urban infrastructure, and their safety and proper functioning are vital. However, haze problems caused by humid environments and temperature differences seriously affect the quality and detection accuracy of drainage pipe images. Traditional repair methods are difficult to meet the requirements when dealing with complex underground environments. To solve this problem, we researched and proposed a dehazing method for drainage pipe images based on multi-attention multi-scale adaptive feature networks. By designing multiple attention and adaptive modules, the network is able to capture global features with multi-scale resolution in complex underground environments, thereby achieving end-to-end dehazing processing. In addition, we also constructed a large drainage pipe dataset containing tens of thousands of clear/hazy image pairs of drainage pipes for network training and testing. Experimental results show that our network exhibits excellent dehazing performance in various complex underground environments, especially in the real scene of urban underground drainage pipes. The contributions of this paper are mainly reflected in the following aspects: first, a novel multi-scale adaptive feature network based on multiple attention is proposed to effectively solve the problem of dehazing drainage pipe images; second, a large-scale drainage pipe data is constructed. The collection provides valuable resources for related research work; finally, the effectiveness and superiority of the proposed method are verified through experiments, and it provides an efficient solution for dehazing work in scenes such as urban underground drainage pipes.



Citation: Li, C.; Tang, Z.; Qiao, J.; Su, C.; Yang, F. Multi-Scale Adaptive Feature Network Drainage Pipe Image Dehazing Method Based on Multiple Attention. *Electronics* **2024**, *13*, 1406. <https://doi.org/10.3390/electronics13071406>

Academic Editor: Beiwen Li

Received: 18 February 2024

Revised: 2 April 2024

Accepted: 3 April 2024

Published: 8 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: drainage pipe; image dehazing; neural networks; multiple attention; multi-scale adaptation

1. Introduction

Drainage pipes play a vital role in the safety and normal operation of urban infrastructure. Even minor failures can lead to significant consequences. Therefore, regular inspection of drainage pipes is crucial to detect and repair potential problems in a timely manner to avoid potential catastrophic outcomes [1]. However, the humid environment and temperature differences inside the drainage pipe can lead to the generation of haze, which significantly impacts image quality and detection accuracy.

Haze is produced when water vapor combines with particles in the air to form suspended water droplets, reducing the contrast and clarity of images. In such scenarios, the application of image dehazing algorithms becomes crucial. Dehazing algorithms can effectively enhance image clarity, leading to more accurate and efficient detection of drainage pipes.

At present, image dehazing methods are mainly divided into two categories: image enhancement-based and image restoration-based. Methods based on image enhancement directly employ algorithms like histogram equalization, homomorphic filtering, and the Retinex algorithm [2–4]. However, these methods may sacrifice image details and introduce noise. The method based on image restoration focuses on the imaging mechanism of haze images and uses the atmospheric scattering model [5] to derive the original image.

In recent years, image dehazing methods based on deep learning have attracted much attention. These methods use deep convolutional neural networks [6] (Convolutional Neural Network, CNN) to learn dehazing models from large amounts of image data through supervised or un/semi-supervised learning methods. Compared to traditional methods, these approaches can automatically learn haze features in images, eliminating the need for manually designed prior knowledge.

Nevertheless, traditional convolution operations lack global awareness when processing images, making it challenging to capture long-range dependencies. In order to solve this problem, research in recent years has begun to introduce attention mechanisms [7] to enhance the performance of image dehazing networks. This can assist the network in better grasping the global information within the image and focusing more on crucial details and structures during feature learning.

In the drainage pipe environment, traditional dehazing algorithms face some challenges due to their closed nature and special working conditions. The first challenge is poor or low light conditions, making it difficult to accurately estimate atmospheric light conditions. Secondly, the haze distribution inside the drainage pipe is often uneven, which brings certain difficulties to the application of dehazing algorithms.

To overcome these challenges, we propose a dehazing method for drainage pipe images based on a multi-attention, multi-scale adaptive feature network. This algorithm combines low-level and high-level haze features, extracting abstract semantic features while retaining detailed information. At the same time, the introduction of the attention mechanism improves the performance of the model, the clarity of the image, and the generalization ability of the model. Compared with traditional methods, our algorithm can better restore the details and structure inside the drainage pipe, reduce the impact of haze, and improve the visibility and recognition accuracy of the image.

In the detection of satellite data images and natural disaster assessment, due to the influence of factors such as atmospheric conditions and weather, images often have problems such as haze and blur, which affects the accuracy and interpretation ability of the data, especially in the Indian Himalayas in areas where early natural disasters frequently occur [8]. The algorithm proposed in this article can improve the clarity and quality of satellite data images through image restoration technology and provide more reliable data support for early disaster monitoring and assessment.

We used a self-designed pipeline robot to collect pipeline images and pipeline depth images and constructed a data set named CDPD-55000 (CUMTB Drainage Pipe Defogging 55000), which has different fog concentrations and is designed to evaluate drainage pipelines' performance of image-defogging technology. On the basis of the data set, we proposed a drainage pipe image dehazing method based on a multi-attention, multi-scale adaptive feature network. Aiming at the problems of thick fog and poor image quality in the real underground drainage pipe environment by combining multiple attention force mechanisms and multi-scale feature fusion to achieve efficient and accurate defogging.

In Section 2, we introduce traditional defogging methods and neural network-based defogging algorithms. In Section 3, we first introduce the multi-scale adaptive feature network framework based on multi-attention and then introduce it in detail. Three innovative modules proposed for defogging work in complex environments are introduced. In Section 4, we introduce the collection and construction method of the data set and analyze and summarize the experimental results and model performance. In the real underground drainage pipe environment, we conducted extensive experimental verification, and the experimental results show that, compared with traditional methods, our method has significantly improved the dehazing effect. The processed image is clearer and has richer detailed information. It is achieved by optimizing the network structure and algorithm. While ensuring the dehazing effect, it also maintains high computing efficiency, providing a powerful basis for subsequent pipeline detection, maintenance and management.

2. Related Prerequisite Work Description

2.1. Dehazing Related Work

2.1.1. Atmospheric Scattering Model

When the image is affected by haze, the imaging process mainly includes two core links. First, light scatters when it passes through microscopic particles in the air. This process causes the light intensity to gradually weaken. The reason is that the light is hindered by particles during the propagation process, causing the propagation direction to change, thus affecting the light intensity. Secondly, scattered outdoor light will hit the camera lens. These two links together constitute the imaging process of hazy weather images.

2.1.2. Dark Channel

The dark channel is an image prior used to analyze and estimate the transmittance of hazy image scenes. In most outdoor natural images, at least in some local areas, there is a channel with very low pixel values, that is, the dark channel [9].

Due to the propagation and scattering effects of haze, the pixel intensity of object surfaces in hazed images will be reduced, resulting in smaller dark channel values. This is because the smaller dark channel value is primarily a result of the contribution of scattered light rather than light directly reflected by the object. After calculating the dark channel value for each pixel, a dark channel map is generated, where each pixel value represents the minimum original intensity value within its neighborhood. This plot provides information about scattering and occlusion in the image.

By utilizing the dark channel map, the transmittance of the image can be estimated, thus eliminating the haze effect and restoring the original scene. Transmittance estimation can be achieved by approximating the relationship between the original image and the dark channel value. Following the transmittance estimation, the haze effect can be removed, enhancing the clarity and quality of the image.

2.1.3. Max Contrast

In the dehazing algorithm, estimating the scattering and transmission components of the image helps to restore the original scene and make the image details clearer. Among them, the maximum contrast measures the degree of difference between pixel values in the image after dehazing. In image processing, increasing contrast can highlight details, make the image clearer, and facilitate subsequent analysis and processing.

2.1.4. Color Attenuation

In the dehazing algorithm, the color attenuation prior is based on the color information of the image, with the purpose of assisting the dehazing algorithm in restoring the clarity of the image. The core concept is that haze will lead to the reduction of image color information, and in a haze-free environment, the color distribution of the image should be richer and clearer. Therefore, this prior helps the dehazing algorithm to more accurately estimate the scattering and transmission components, thereby achieving more accurate dehazing effects.

As a constraint, it ensures the authenticity and vividness of colors during dehazing. By providing a more accurate estimate of the transmission component, the dehazing effect can be significantly improved. In practical applications, the color attenuation prior is combined with other prior knowledge (such as atmospheric light prior, dark channel prior, etc.) to further improve the performance of the dehazing algorithm.

2.1.5. Chromaticity Difference

Chromaticity difference refers to the difference in color between a pixel and its neighboring pixels. In color vision, it helps us distinguish object boundaries and details. In the training of deep learning models, chromaticity information is extracted as one of the important features to improve the quality of restoration, especially in terms of maintain-

ing color consistency and details. Chromaticity differences also play a key role in image dehazing. Haze changes the chromaticity of the image, so the chromaticity difference appears significantly different between hazy and haze-free images. Using chromaticity differences as features, we can better understand the impact of haze on images and design more effective dehazing algorithms. Chromaticity differences can not only be used alone but can also be combined with other features (such as dark channels, color attenuation, etc.) to extract more haze-related information. These features work together to help estimate the haze level more accurately and provide a basis for the design of more accurate dehazing algorithms.

2.2. Dehazing Related Work

Early image dehazing methods were mainly based on manually designed prior knowledge, such as dark channel prior (dark channel prior, DCP), color attenuation prior [10] (color attenuation prior, CAP) and haze line prior, [11] (Haze Line, HL), etc., use the haze imaging model to achieve dehazing by estimating the transmission map and atmospheric light value. The DCP (dark channel prior) method is based on the principle of dark channel prior, where the input hazy image is first processed to compute its dark channel image. The brightest pixels are selected from the dark channel image to estimate the value of atmospheric light. Then, the initial transmission rate map is calculated based on the dark channel image and the atmospheric light value. In order to improve the dehazing effect of the algorithm, refinement processing is usually applied to the initial transmission rate to eliminate block artifacts and retain more edge information. Finally, utilizing the atmospheric scattering model and combining the estimated atmospheric light with the refined transmission rate, the original hazy image is dehazed to obtain a clear haze-free image. He et al. [12] proposed an improved dark channel method incorporating color restoration using the atmospheric scattering model and a multi-scale Retinex strategy. While this method can restore clear images in specific scenes and has significantly contributed to the advancement of image dehazing technology, obtaining sufficient statistical information with manually designed prior knowledge remains challenging for handling real and complex haze scenes. Among them, the classic dark channel prior algorithm is simple and effective, but it is sensitive to specific scenes or atmospheric conditions, produces artifacts in some special cases, and is sensitive to the estimation of the initial atmospheric scattering rate, so it is not suitable for drainage pipe scenes.

Multi-sensor image dehazing algorithms, which incorporate data collected from multiple sensors such as optical cameras, infrared cameras, depth cameras, etc., estimate the degree of haze in images based on atmospheric scattering models and fuse the dehazing results from different sensors, fully leveraging the advantages of each sensor. However, practical usage faces challenges such as data asynchrony between sensors, difficulty in data fusion and calibration among different sensors, etc., thus imposing limitations on the practical application of multi-sensor dehazing methods.

Image decomposition techniques play a significant role in haze correction by effectively removing the influence of haze on images and improving their clarity and quality by decomposing the different components of the image. Hazy weather often causes edge blurring and information loss. Agrawal et al. [13] proposed an edge suppression method using cross-projection tensors for gradient field transformation, exhibiting strong edge protection capability in image processing. It preserves edge information while removing haze, thereby avoiding detail loss caused by excessive smoothing. This method is suitable for scenarios requiring high image quality, such as traffic monitoring and remote sensing image processing. Wu et al. [14] introduced a method using flight time imaging to decompose global light transmission, effectively removing the influence of haze on images by modeling and decomposing global light transmission. By incorporating depth information of the scene and light transmission decomposition, haze effects on images can be more accurately estimated and corrected, suitable for scenes requiring depth information, such as autonomous driving and robotic navigation. Muhuri et al. [15] proposed a scatter power

decomposition method based on geodesic distance, mainly applied to compact polarimetric SAR data processing. Through scatter power decomposition, a better understanding of the characteristics of different components in the image can be achieved. Although this method has not been directly applied to haze correction, it holds significant importance in distinguishing between haze components and real scene components in haze correction.

With the performance breakthrough of deep learning in the field of computer vision, many image dehazing methods based on deep convolutional neural networks (CNN) have been proposed [16–19]. DehazeNet [20] and multi-scale convolutional neural network [21] estimate the transmittance by learning the mapping relationship between hazy images and their projections and restoring haze-free according to the atmospheric scattering model image. The dense pyramid dehazing network [22] uses subnetworks to estimate transmittance and atmospheric light values and generates haze-free images through generative adversarial network training. Li [23] combined transmittance and atmospheric light into one variable and proposed a new network structure (all-in-one dehazing, AOD-Net) to estimate this variable, eliminating the error of training atmospheric light and transmittance separately. A better restoration effect was achieved, but the restoration result was overall darker. Liu [24] proposed a residual network structure to estimate the transmittance based on hazy images and their assumptions or prior information and then obtain the restoration results. Qian [25] proposed a new dehazing convolutional neural network (CNN)—CIASM-Net. The network model includes a color feature extraction sub-network and a deep dehazing sub-network and uses multi-scale convolution to estimate the transmission rate, and then obtain the recovery result. Chen [26] proposed a method of image enhancement using generative adversarial networks (GANs). This method learns from input images to enhance images in a natural and effective way by training a generator and discriminator network. This method is able to extract more details from ordinary photos and improve the brightness and contrast of the image. However, this method may have the disadvantages of a certain model training complexity and the need for a large amount of training data. Cycle-Dehaze [27] is a direct end-to-end network that generates haze-free images directly from hazy input images. These methods mainly adopt general network architectures (e.g., DenseNet [28], Dilated Network [29], Grid Network). However, due to the lack of light or low light in the urban underground pipeline environment, which is dark and humid and has severe water mist, the details of the collected video images are blurred. These methods are not optimized for the lack of detailed information in pipeline scenes, resulting in poor dehazing effects in drainage pipes. The advantages and disadvantages of each algorithm can be seen in Table 1, where \ denotes that the method is not used for pipeline scene and has no detailed results.

In order to solve this problem, this paper proposes a multi-scale drainage pipe dehazing network based on an attention mechanism. The network utilizes the “channel + spatial” attention mechanism to improve the dehazing ability and restore the detailed information of the image. This novel network architecture allows us to capture more detailed image feature representations by integrating both channel and spatial attention mechanisms. The attention mechanism is more adaptive and flexible compared to traditional convolution. It dynamically focuses on detailed information at specific locations or channels. This capability empowers the attention mechanism to deliver enhanced performance and effectiveness for complex image processing and machine learning tasks.

In summary, within the multi-scale drainage pipe dehazing network employing the attention mechanism for the drainage pipe image dehazing task, this attention mechanism aids the model in focusing more effectively on crucial detail areas and extracting clear features, consequently improving haze removal and restoring image quality.

Table 1. Literature comparison table.

Ref.	Methods/Techniques	Main Contributions/Features	Application Effect in Drainage Pipes	PSNR/SSIM
[11]	Combining dark channel prior with multi-scale Retinex strategy	Improved dark channel prior dehazing effect and enhanced color recovery	Performs well in some scenarios but is sensitive to certain atmospheric conditions	27.74/0.88
[12]	Cross-projected tensor gradient field transformation	Powerful edge protection capabilities to avoid loss of details	Suitable for scenarios that require high edge information	\
[13]	Global light transport decomposition for time-of-flight imaging	Improve the dehazing effect by obtaining depth information	Suitable for scenes that require depth information	\
[14]	Scattering power decomposition based on geodesic distance	Understand the characteristics of image components and distinguish haze from real scenes	Currently not directly applied to haze correction, but has potential value	\
[17]	Dense pyramid defogging network	Estimating transmittance and atmospheric light values using subnetworks	Produces haze-free images with higher quality, but may be more computationally complex	\
[18]	AOD-Net	Combine transmittance and atmospheric light into a single variable to simplify the training process	The restoration result is overall darker and needs further adjustment	24.14/0.92
[20]	DehazeNet	Learn the mapping relationship between foggy images and projection images	Improved defogging effect, but may be limited by training data	23.16/0.82
[21]	Multi-scale convolutional neural network	Improve dehazing performance through multi-scale features	Suitable for a variety of scenarios, but may require further optimization	21.32/0.85
[22]	Image enhancement using generative adversarial networks	Ability to extract more details from input images and improve brightness and contrast	Suitable for a variety of image enhancement tasks, but may require specific optimization in dehazing tasks	21.56/0.86
[23]	Direct end-to-end network	Generate haze-free images directly from hazy images	Simplified dehazing process, but may need to be optimized for pipeline scenarios	30.16/0.93
[25]	CIASM-Net	Including color feature extraction and deep dehazing sub-network	Improving transmittance estimation accuracy through multi-scale convolution	21.26/0.85
Our research team	Multi-scale adaptive feature network based on multiple attention	Utilize multiple attention mechanisms and multi-scale feature fusion to improve the dehazing effect	Achieve superior dehazing and preserve detail in sewer environments	39.87/0.98

3. Main Network Framework

3.1. Overall Structure

Existing research has shown that using multi-scale information in image restoration can significantly improve the performance of the model. In view of this, we propose a method that combines U-Net [30–34] architecture, multiple attention mechanisms and multi-scale networks to fully utilize the multi-scale information of images and features to achieve image dehazing so as to facilitate the subsequent use of YOLO, etc. network for advanced visual detection tasks [35].

As depicted in Figure 1, the method comprises three main components: the encoder module, enhanced decoder module, and feature recovery module. The encoder module focuses on extracting valuable feature information from the input image, progressively reducing the image size through multiple convolutional layers to capture its high-level features. The enhanced decoder module uses skip connections in the U-Net structure to combine the output feature map of the encoder with the input feature map of the decoder. In the decoder, we apply the multi-scale feature fusion module to enhance the details and contrast of the image. The role of the feature recovery module is to transfer the output feature map of the encoder to the decoder through skip connections, restoring the lost high-level feature information. This allows the decoder to more effectively leverage the context information provided by the encoder to generate dehazed images.

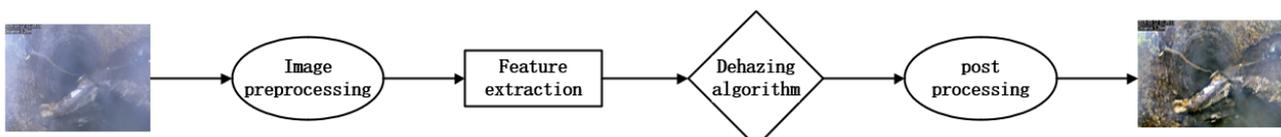


Figure 1. Dehazing process flow chart.

To better manage the details and contextual information of images, we introduce multi-scale attention modules in both the encoder and decoder. These modules include a channel attention mechanism and a multi-scale spatial attention mechanism, replacing traditional convolution operations. The channel attention mechanism can automatically learn the importance weights between channels, while the multi-scale spatial attention mechanism can adaptively capture image features at various scales.

To better handle the details and contextual information of images, we introduce multiple attention modules in the encoder and decoder. As shown in Figure 2, these modules consist of a channel attention mechanism and a multi-scale spatial attention mechanism, replacing traditional convolution operations. The channel attention mechanism can automatically learn the importance weights between channels, while the multi-scale spatial attention mechanism can adaptively capture image features at different scales.

3.2. Multiple Attention Module

In a convolutional neural network, the input data usually consists of multiple channels, each channel representing a different feature of the image. The channel attention mechanism aims to dynamically learn the weight of each channel so that the network can adaptively focus on the feature information of different channels.

When multi-scale feature information needs to be fused, we also need a mechanism to share information within a feature tensor. As shown in Figure 3, this paper uses an improved channel attention mechanism to replace traditional convolution. The input feature map $F_{in} \in R^{C \times H \times W}$, where C is the number of channels, H and W are the height and width of the feature map respectively. Convolution-deconvolution operations are first performed for preliminary feature extraction and spatial recovery to obtain the feature map F_1 , which helps the network learn the image. The semantic features increase the resolution of the feature map to more precisely locate and restore the boundary information and detailed information of the object. Then, through global average pooling, the feature map of each channel is compressed into a single value, which represents the global characteristics

of this channel. Using the learned global features, a small neural network is introduced to learn the weight of each channel. These weights represent the relative importance of each channel. They will be used to weight the input feature map to produce a channel attention-weighted feature map F_2 .

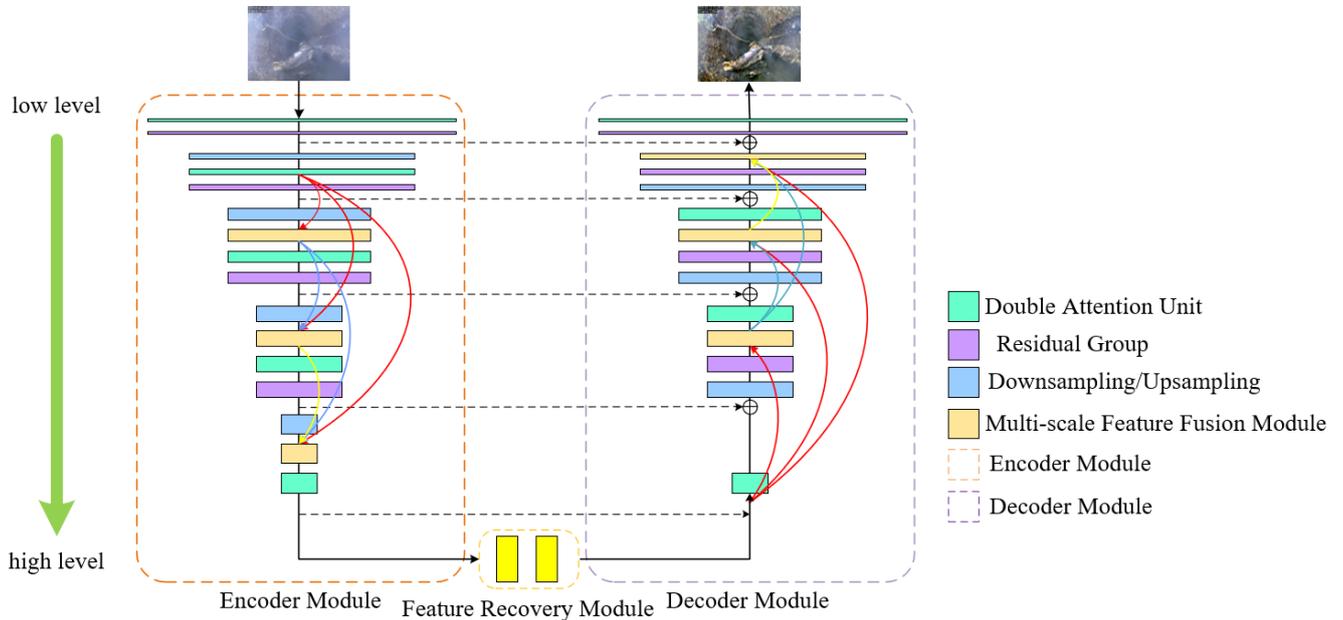


Figure 2. Network structure.

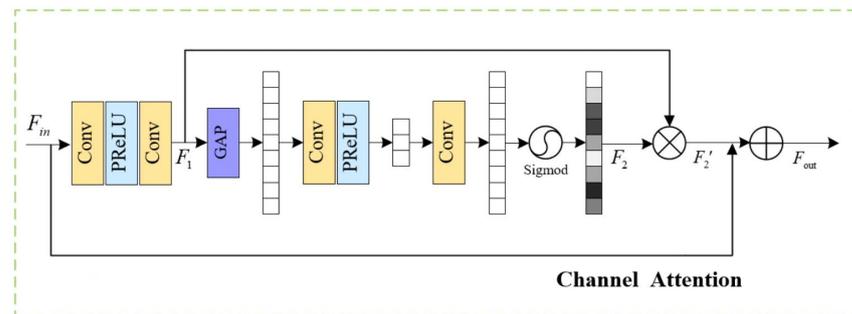


Figure 3. Channel Attention.

We also combine channel attention with the residual structure to improve the performance and learning ability of the network and use the skip connection of the residual structure to alleviate the vanishing gradient problem. The channel attention dynamically focuses on the feature information of different channels, enhancing the feature expression ability. In this way, the network can better adapt to features of different categories and complexity, improving the model’s performance capabilities. Between F_1 and F_2 we used a residual structure; after the residual connection, the module learns the weight of each channel and applies these weights to each channel of the feature map. In the output of the original feature map F_{in} and F_2' , a residual structure was also used, which resulted in F_{out} .

To heighten the awareness of the importance of various spatial locations in the input data and concentrate on image areas crucial to the dehazing task, we introduce an enhanced “channel + space” dual attention mechanism. The spatial attention (SA) branch is crafted to leverage the spatial correlation of convolutional features. Its primary objective is to produce a spatial attention map and utilize it to reweight the input features M . Through the incorporation of a spatial attention mechanism, the model can bolster its capability to discern the significance of diverse spatial locations in the image.

In the SA branch, first, the input feature M is subjected to two operations along the channel dimension, namely global average pooling and max pooling. In this way, the average and maximum values of the features in the channel dimension can be extracted separately and connected together to form a feature map $f \in R^{H \times W \times 2}$. Next, after a convolution operation and a sigmoid activation function, the feature map f is converted into a spatial attention map $\overset{\Delta}{f} \in R^{H \times W \times 2}$. $\overset{\Delta}{f}$ emphasizes task-critical image regions by learning a set of weights to weighted combinations of different spatial locations in the feature M . This weighting operation can increase the model’s attention to important locations in the image and help the model better understand and process the input data. Finally, by performing an element-wise multiplication operation on the spatial attention map $\overset{\Delta}{f}$ and the input feature M , each position of the feature M will be adjusted by the corresponding spatial attention weight. This process of re-adjusting feature weights can be viewed as a rescaling of the input features M in order to better utilize these weighted features in subsequent model levels.

3.3. Multiple Attention Module

To address the issues of spatial information loss during the downsampling process and the absence of connections between non-adjacent level features in the U-Net architecture’s encoder, we introduce a multi-scale feature fusion module as shown in Figure 4. This module effectively fuses features from different levels to compensate for the missing spatial information in upper layers and fully harnesses the features of non-adjacent layers.

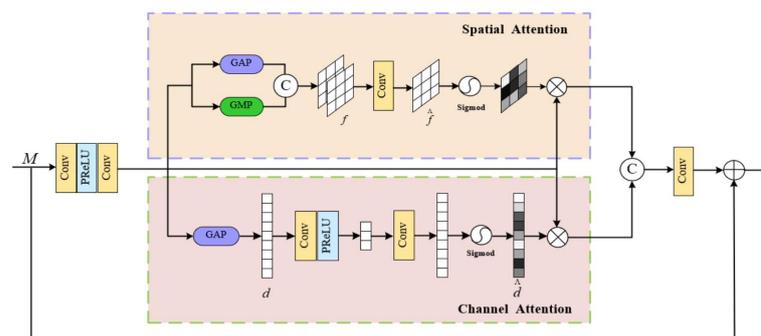


Figure 4. Multiple Attention Module.

The multi-scale feature fusion module (MSFFM) employs an error feedback mechanism to further enhance the features of the current level and is applied to both the encoder and decoder. This module preserves the spatial information of high-resolution features and utilizes non-adjacent features for effective image dehazing. Specifically, two MSFFM modules are introduced at each level—one before the residual group of the encoder and the other after the multi-scale enhancement decoder of the decoder. In the encoder/decoder, the enhanced FMFA output is linked to all subsequent MSFFMs for feature fusion. In this manner, within the U-Net architecture, we can effectively address the issues of missing spatial information and inadequate feature connections in non-adjacent layers during downsampling, thereby enhancing the performance of image dehazing.

Each level of the decoder has an MSFFM module, which is defined by Equation (1). In Equation (1), j^n represents the enhanced features of the current level of the decoder, \tilde{j}^n represents the fusion features obtained through feature fusion, and L represents the number of levels of the network, which represents the fusion features output by all previous $\{j^L, j^{L-1}, \dots, j^{n+1}\}$ level MSFFM modules in the decoder.

$$\tilde{j}^n = D_{de}^n(j^n, \{j^L, j^{L-1}, \dots, j^{n+1}\}) \tag{1}$$

The update process of the MSFFM module is as follows:

1. Calculate the difference e_t^n between $j_t^n (j_0^n = j^n)$ and \tilde{j}^{L-t} at the t -th iteration using Equation (2). Among them, p_t^n represents the projection operator, which downsamples the enhanced feature j_t^n to the same dimension as the fused feature \tilde{j}^{L-t} ;

$$e_t^n = p_t^n(j_t^n) - \tilde{j}^{L-t} \tag{2}$$

2. Use the back-projection difference j_t^n to update and calculate according to Equation (3). Among them, q_t^n represents the back-projection operator, which upsamples the difference e_t^n of the previous iteration to the same dimension as the fused feature j_t^n ;

$$j_{t+1}^n = q_t^n(e_t^n) + j_t^n \tag{3}$$

3. All previous fusion features $\{\tilde{j}^L, \tilde{j}^{L-1}, \dots, \tilde{j}^{n+1}\}$ are iteratively processed to obtain the final fusion features \tilde{j}^n .

However, during this process, the network is unknown to p_t^n and q_t^n . Inspired by super-resolution deep back-projection networks, we use convolutional/deconvolutional layers to learn corresponding downsampling or upsampling operations. To reduce the number of parameters, we use $(L - n - t)$ stacks of convolution and deconvolution layers with a stride of 2 to implement downsampling or upsampling operations. The structure of the MSFFM module is shown in Figure 5, which depicts the last multi-scale feature fusion module in the enhanced decoder module.

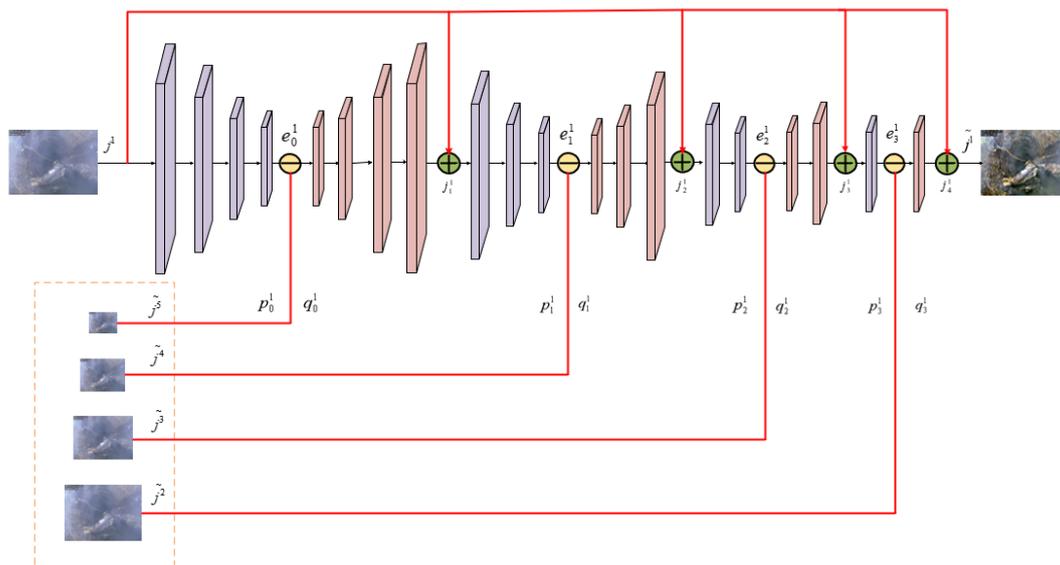


Figure 5. Multiple Attention Module.

4. Similarly, the MSFFM module at the encoder level can be defined using Equation (4):

$$\tilde{i}^n = D_{en}^n(i^n, \{i^1, i^2, \dots, i^{n-1}\}) \tag{4}$$

In Equation (4), i^n represents the latent features of the n -level encoder, and $\{i^1, i^2, \dots, i^{n-1}\}$ represents the fused features output by all previous $(n-1)$ -level MSFFM modules in the encoder. The architecture D_{en}^n of the MSFFM module at the encoder level is the same as the D_{en}^{L-n} at the $(L - n)$ -level of the decoder, except that the positions of the downsampling operation p_t^n and the upsampling operation q_t^n need to be interchanged.

3.4. Multi-Scale Feature Enhancement Module

In order to gradually improve the features in the feature recovery module in the decoder, we introduce a multi-scale feature enhancement module in the decoder of the network. It helps recover high-level visual features from low-level pixel-level information, thereby improving image quality and clarity. The structure of the multi-scale enhancement module is shown in Figure 6.

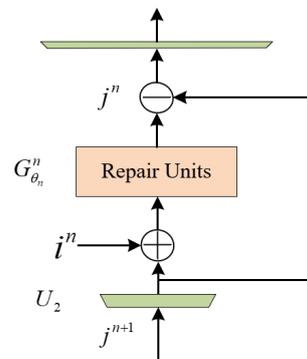


Figure 6. Multi-scale Feature Enhancement Module.

In the n -th level multi-scale enhancement module, the feature map obtained from the previous level is first upsampled. Then, the upsampled feature map is added to the feature map obtained by the corresponding encoder at the same level for enhancement. Next, the enhanced feature map is sent to the repair unit for processing, and the upsampled feature map is subtracted from the repair result. The output obtained in this way is the enhanced feature of the n -th-level multi-scale enhancement module, as shown in Equation (5):

$$j^n = G_{\theta_n}^n(i^n + U_2(j^{n+1})) - U_2(j^{n+1}) \quad (5)$$

In Equation (5), we upsample the feature map using an upsampling operation U_2 with a scale factor of 2. Then, use a symbol to represent the enhanced feature $i^n + U_2(j^{n+1})$. $G_{\theta_n}^n$ represents the trainable repair unit with parameters θ_n in the n th layer, which uses the residual group in the encoder to implement each repair unit.

The structure of the residual group is shown in Figure 7. The residual group consists of three residual blocks, which contain two convolutional layers with a convolution kernel size of 3×3 and a stride of 1. First, a convolution operation is performed on the features of the input residual block, then the PReLU activation function is applied, and then another convolution operation is performed. At the same time, the residual group processes the residual structure on itself. Finally, the convolution results are added to the skip-connected input features.

At the final layer of the decoder, we employ a convolutional layer to reconstruct the estimated haze-free image from the last feature map. This convolutional layer outputs the haze-free image of the drainage pipe.

This section primarily elaborates on the three modules: multi-attention, multi-scale feature fusion, and multi-scale feature enhancement. The model embraces a hierarchical design concept, incorporating multiple attentions in each layer and executing feature fusion across different levels. Each level can gradually expand the receptive field and achieve up- and downsampling, thereby obtaining feature maps of four different scales.

By utilizing multiple attention modules, the network can adaptively learn crucial information within the data, thereby enhancing the model's performance. Furthermore, the multi-scale feature fusion module aids in effectively fusing feature information of diverse scales, enhancing the model's expressive capability. The multi-scale feature enhancement module employs the fused multi-scale features to progressively restore higher-level features through other modules in the decoder. Simultaneously, each level incorporates the concept of a residual group. The residual group increases the model's depth, enabling it to better

capture the essential characteristics of the data during the training process. This mitigates the issue of gradient disappearance, thereby enhancing the model's generalization ability.

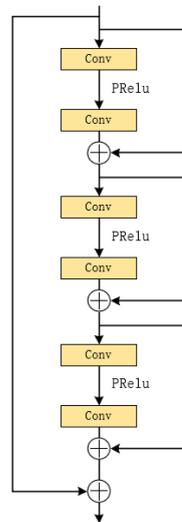


Figure 7. Residual Group Structure.

The hierarchical design model proposed in this section, with the aid of the multi-attention module, multi-scale feature fusion module, and multi-scale feature enhancement module, gradually expands the receptive field. It achieves the fusion and enhancement of up and downsampling features, obtaining feature maps at four different scales. This design provides the model with significant performance advantages and enhances its robustness.

4. Discussion and Analysis of Experimental Results

4.1. Drainage Pipe Data Set

This article employs a pipe robot equipped with a binocular stereo depth camera (Intel RealSense Depth Camera D435i) for collecting image data of drainage pipes. As shown in Figure 8, the pipeline robot, independently designed and developed by our research team, operates smoothly in various complex environments and is controlled by a traction rope. The camera primarily captures color images and corresponding depth images inside the pipeline. The used D435i camera is equipped with four cameras, comprising a left infrared camera, an infrared dot projector, a right infrared camera, and a 20-megapixel RGB camera. This camera employs binocular vision technology. Utilizing image data generated by the left and right infrared cameras, combined with the built-in depth processor, the depth value of each pixel can be calculated to generate a depth image. Unlike RGBD cameras, the D435i's infrared camera can provide reliable depth measurements under low-light conditions, a crucial factor for data collection in drainage pipe scenes.

Throughout the data collection process, we selected five distinct areas in Beijing for field collection. These real urban underground drainage pipe data aim to validate the effectiveness of our proposed dehazing algorithm, ensuring its capability to produce satisfactory dehazing effects in future real-world scenarios. Besides their academic research value, these data can serve as a reference for the management and maintenance of drainage pipelines.

Acquiring both hazy and clear images of the same scene poses a significant challenge in most cases. This is because the acquisition of hazy images and clear images needs to consider many factors, such as weather conditions, light intensity, and camera parameters. To solve this problem, we use an atmospheric scattering model to generate hazy sky images with different concentrations. The primary innovation of this method involves utilizing a pipeline robot to capture clear images and simulate haze effects of different concentrations through an atmospheric scattering model.

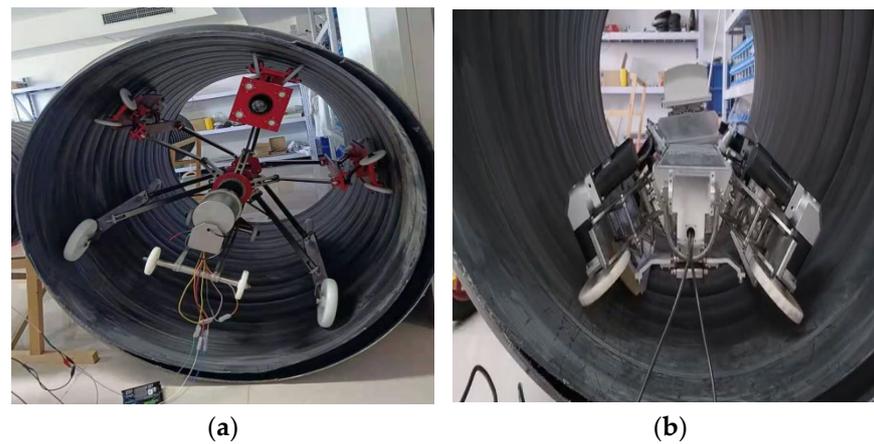


Figure 8. (a) Self-developed drainage pipe robot; (b) Working diagram of drainage pipe robot simulation data collection.

Pipe robots are autonomous devices capable of maneuvering in confined spaces and equipped with high-definition cameras to capture clear images. Initially, these images are utilized to train deep learning models, extracting key features in the scene. Subsequently, the trained model is employed to generate hazy images.

The process of generating hazy images can be divided into the following steps:

1. Collect clear images: Use the high-definition camera mounted on the pipeline robot to capture clear images in various scenarios.
2. Train the deep learning model: Use the collected clear images to train the deep learning model and extract scene features.
3. Generate hazy images: Based on the trained model and atmospheric scattering model, simulate the haze effects of different concentrations to generate hazy images, as shown in Figure 9.
4. Analysis and evaluation: Perform quality assessment and analysis on the generated hazy images and explore the impact of different haze concentrations on image quality.

By following the above steps, we are able to generate hazy images with various concentrations in diverse scenarios, thereby providing essential data support for subsequent research. This approach introduces a fresh perspective to hazy image generation and contributes to enhancing drivers' safety during hazy weather conditions. Simultaneously, the outcomes of this research are anticipated to find applications in other fields, including computer graphics, virtual reality, and more.

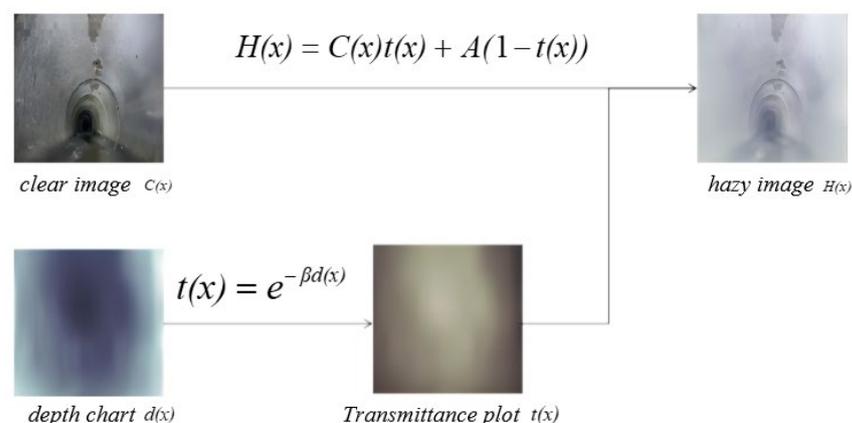


Figure 9. The Process of Generating Hazy Images.

To facilitate adaptive dehazing training for drainage pipe images, we curated a data set comprising hazed pipe images with varying haze concentrations. We fixed the global

atmospheric light at a value of 1 and randomly selected the scattering coefficient within the range of 0.6 to 1.5, resulting in the synthesis of 10 hazy images for each clear image. This methodology resulted in a data set comprising clear pipeline images and their corresponding hazy counterparts, with each clear image having 10 variations generated with different scattering coefficients. Using this approach, we utilized 5750 clear images collected by the pipeline robot to generate corresponding hazy images, adjusting them to a size of 448×448 pixels. Out of these, 55,000 pairs of hazy and clear images are allocated for training purposes, while 2500 pairs are designated for testing. Finally, we constructed the drainage pipe dehazing data set, as shown in Figure 10.

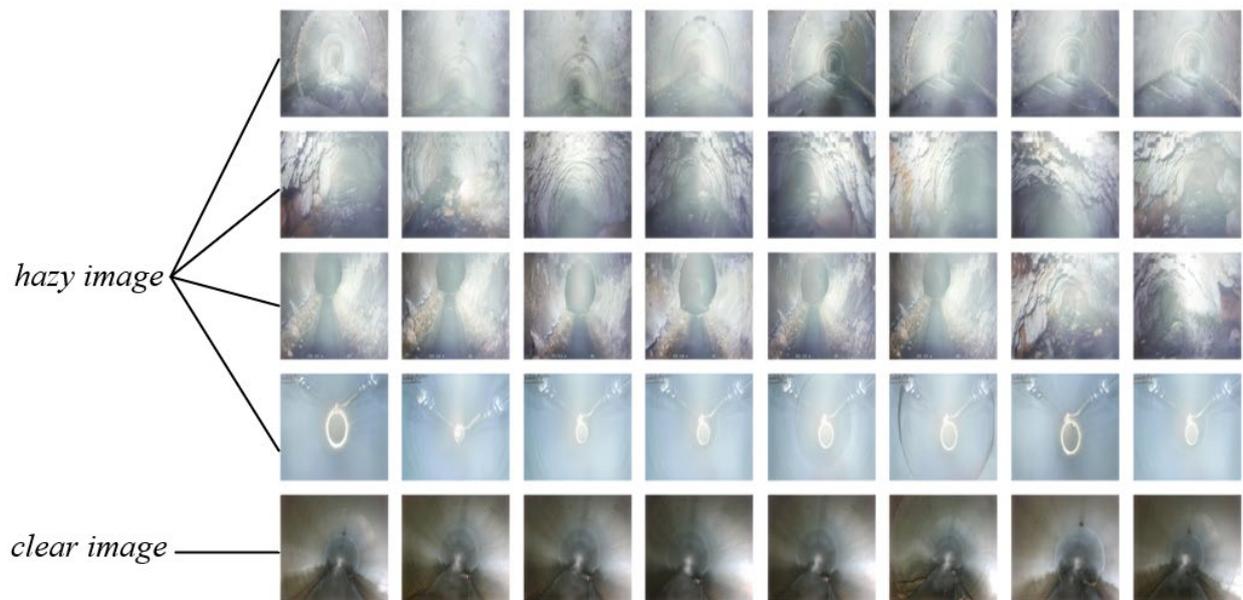


Figure 10. Drainage Pipe Dehazing Data Set.

4.2. Implementation Process

The model embraces a hierarchical design concept, meticulously crafting four levels. Each level incorporates a multi-attention module, a multi-scale feature fusion module, and a multi-scale feature enhancement module. The hierarchical design not only enlarges the model's receptive field but also ensures the accuracy and efficiency of information transmission through meticulous layer-by-layer processing. At each level, multiple modules work together to extract and fuse multi-scale feature information layer by layer, further improving the model's adaptability to complex scenes.

To enhance the model's convergence speed and stability, we employ a 3×3 filter size in all convolutional and deconvolutional layers. This setting can capture more local information while ensuring spatial resolution. Moreover, we substitute the traditional Relu activation function with the PRelu activation function. The PRelu activation function better simulates the actual distribution of data during the training process, expediting the model's convergence and enhancing its stability.

In order to measure the actual effect of the model, we use the mean square error (MSE) as the loss function. As a classic loss function, the mean square error can accurately measure the difference between the network output and the real image. By minimizing the mean square error, we can ensure that the gap between the dehazed image output by the model and the real clear image gradually narrows, thereby achieving a more accurate dehazing effect. The selection of this loss function ensures the reliability and accuracy of model training.

During the entire training process, we set 200 epochs for model training. Within each epoch, the model traverses the entire data set once and updates the network parameters. We set the batch size to 16. Additionally, the initial learning rate was set to 10^{-4} and adjusted

with a decay rate of 0.75 after every 10 epochs. This strategy of dynamically adjusting the learning rate aids the model in converging quickly and stably at various training stages.

All experimental operations are conducted on the NVIDIA GeForce RTX 3060 GPU, ensuring the computational speed and stability of the model. Since every time an image is read from the disk, I/O operations are required, which may affect training efficiency. Therefore, in order to increase data reading speed and avoid I/O bottlenecks, we save the raw image data into HDF5 files. The HDF5 file format is highly flexible and scalable and can meet our needs for quickly traversing and accessing data during the training process. This optimization of data storage not only enhances the training efficiency of the model but also facilitates subsequent data analysis and processing.

4.3. Performance Analysis

In order to verify the effectiveness of the algorithm proposed in this article, we use a self-built drainage pipe data set to train and test the dehazing network model instead of using public image data sets, such as RESIDE and NYU-depth2 data sets. For assessing the performance of the dehazing algorithm, we employed the following evaluation metrics: Peak Signal-to-Noise Ratio (PSNR) [36,37] and Structural Similarity (SSIM) [38,39].

Peak signal-to-noise ratio is a commonly used metric for image quality, measured in dB. The higher the value, the better the image quality. Namely, the closer the dehazed image aligns with the original clear image in detail, the lower the image distortion. In general, a PSNR higher than 40 dB implies that the image quality closely approximates the original image; a range of 30–40 dB suggests that the distortion loss of image quality is within an acceptable range; a range of 20–30 dB indicates relatively poor image quality; and below 20 dB signifies severe image distortion. Specifically, in our experiments, we used the PSNR indicator to compare the DCP algorithm, DehazeNet algorithm, AOD-Net algorithm, MSBDN algorithm and the algorithm proposed in this paper on the test data set. Performance of dehazing the hazy images of the pipeline.

Structural similarity is an index that measures the similarity of structural information before and after image processing. It considers the three dimensions of brightness, contrast and structure. The value range of this index is [0, 1]. The closer the value is to 1, the more similar the images are, indicating better image quality after dehazing. In our experiments, we employed the SSIM indicator to compare the performance of the DCP algorithm, DehazeNet algorithm, AOD-Net algorithm, MSBDN algorithm, and the algorithm proposed in this paper for dehazing hazy images of drainage pipes on the test data set.

As shown in Table 2, the dehazing algorithm presented in this paper outperforms others in terms of peak signal-to-noise ratio, surpassing the MSBDN algorithm by 6.37 dB. Compared to other methods, there is an improvement in structural similarity. In comparison to the MSBDN, the structural similarity (SSIM) value of our proposed method has increased by 0.062.

Table 2. Analysis of the dehazing quality of five methods from the perspective of PSNR and SSIM.

Method	DCP	DehazeNet	AOD-Net	MSBDN	Ours
PSNR	25.74	27.86	28.98	33.50	39.87
SSIM	0.759	0.797	0.859	0.926	0.988

This paper proposes a novel network structure that integrates channel and spatial attention. Building upon the original MSBDN, this structure combines multiple attention mechanisms with multi-scale feature fusion and enhancement modules, resulting in a more intricate and hierarchical network. Experimental results demonstrate that the incorporation of this multiple attention mechanism has led to a 5~20% improvement in both PSNR and SSIM indicators.

The ablation experiment part is designed with reference to MBTFCN. Conducting ablation research can help us understand the impact of each module on network performance, thereby verifying the effectiveness of the network design. MSBDN has achieved good

results in image reconstruction tasks, but it ignores the importance of spatial information and channel information. To overcome this limitation, we introduce the multiple attention module (DAU). By focusing on the correlation between different channels, DAU enables the model to better capture the channel characteristics of the image, thereby improving the reconstruction quality. Secondly, in order to better preserve image details, we also integrated the MSFFM structure into the model.

According to Table 3, after adding the DAU attention mechanism, the network performance has been enhanced to a certain extent. Especially when processing images with complex backgrounds and thick fog, the attention mechanism can automatically focus on key areas in the image and reduce background interference, thus improving classification accuracy. When processing images with complex textures and color changes, the introduction of DAU enables the model to better retain the detailed information of the original image. After adding the MSFFM module, the model also improved in PSNR and SSIM indicators. This proves that the MSFFM structure can effectively fuse features of different scales and enhance the representation ability of the model. MSFFM performs well in processing image structures at different scales, especially in processing large-size images or images containing elements of multiple scales. When DAU and MSFFM are introduced at the same time, MSBDN + CAB + MSFFM achieves the best performance in all indicators. This shows that there is a complementary effect between DAU and MSFFM, which jointly improves the reconstruction ability of the model. It is demonstrated that the model performs well in handling various complex scenes and image types, providing an efficient and reliable solution for image reconstruction tasks.

Table 3. Ablation Experiment.

Method	MSBDN	MSBDN + DAU	MSBDN + MSFFM	Ours
PSNR	33.50	34.28	38.86	39.87
SSIM	0.926	0.954	0.985	0.988

Through ablation experiments, we verify the effectiveness of the multiple attention module (DAU) and the multi-scale feature fusion module (MSFFM) in image reconstruction tasks. The introduction of these two modules not only improves the performance of the model but also enhances the model's ability to handle complex scenes and image types.

We also analyze the computation time of the proposed network and compare it with previous methods. As shown in Table 4, We recorded the time required to process the same set of drainage pipe images using different dehazing algorithms. In order to ensure the fairness of the experiment, the same hardware configuration and software environment were used, and the size and resolution of the images were consistent. Our network performs well in terms of computation time compared to previous methods. Although the network adopts multi-scale adaptive modules and multiple attention mechanisms, these designs do not significantly increase the computational burden. On the contrary, due to the optimization of the network structure and the efficient implementation of the algorithm, the time consumption of our method when processing each image is comparable to some traditional methods and even faster in some cases.

Table 4. Average run-time (in seconds) on test images.

Method	DarkChannel	DehazeNet	AOD-Net	MSCNN	MSBDN	Ours
Average Time	26.90	7.02	3.23	1.34	1.10	1.05

The analysis of these objective evaluation indicators shows that the defogging algorithm we proposed performs well in dehazing drainage pipe images, has a high peak signal-to-noise ratio and structural similarity, and has certain advantages in computational efficiency, proving the effectiveness and feasibility of the algorithm.

4.4. Subjective Analysis

On the drainage pipe data set, we conducted comparative experiments on multiple dehazing algorithms to evaluate the performance of each algorithm in practical applications. Figures 11–13 depict the effects of thin haze, thick haze, and uneven haze, respectively. Through comparative analysis, we found that various dehazing methods have achieved certain results in removing haze, but they also have certain problems.

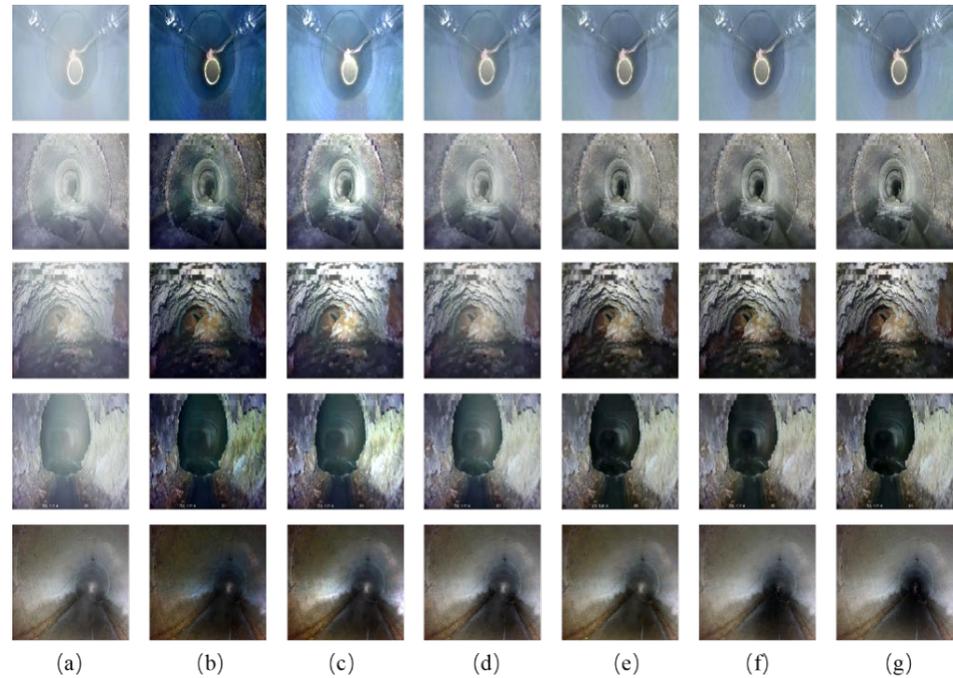


Figure 11. The visual effect of haze dehazing in the drainage pipe data set using the adaptive dehazing algorithm: (a) haze image, (b) DCP, (c) DehazeNet, (d) DCPDN, (e) MSBDN, (f) Ours, (g) haze-free image.

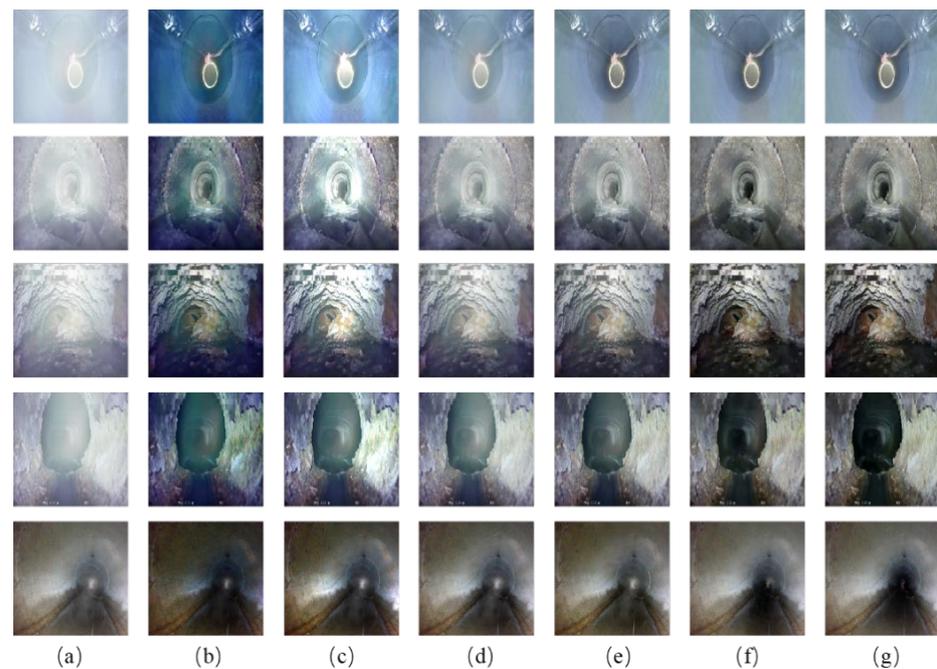


Figure 12. Visual effect of adaptive dehazing algorithm on dense haze in drainage pipe data set: (a) Dense haze image, (b) DCP, (c) DehazeNet, (d) DCPDN, (e) MSBDN, (f) Ours, (g) Haze-free image.

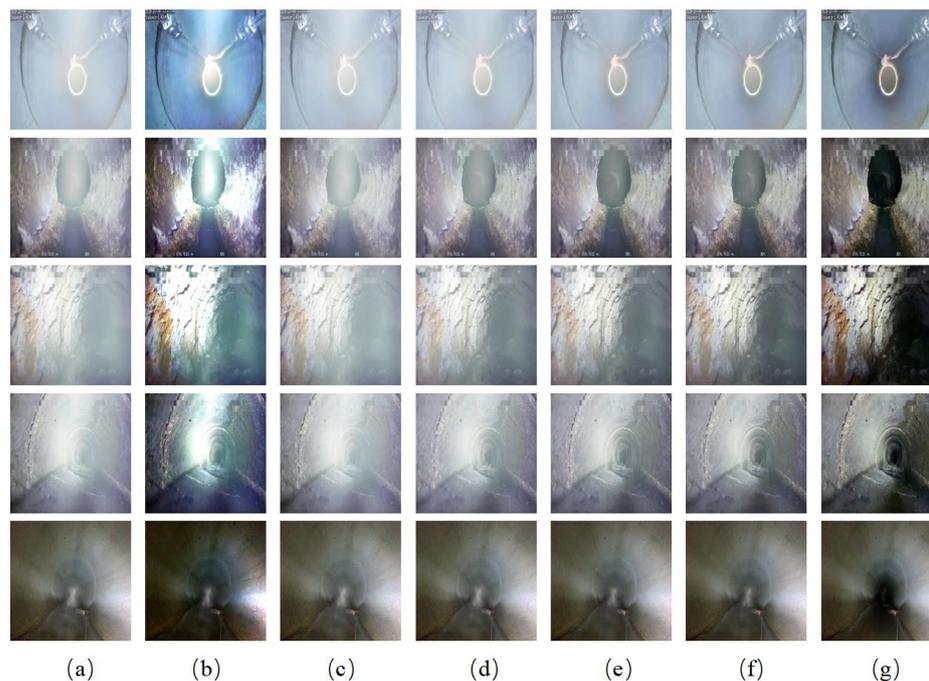


Figure 13. Visual effect of dehazing in uneven concentration haze data set: (a) Uneven concentration haze image, (b) DCP, (c) DehazeNet, (d) DCPDN, (e) MSBDN, (f) Ours, (g) None haze image.

While the DCP dehazing algorithm successfully eliminates haze, it tends to reduce the overall brightness and color saturation of the image, resulting in a darker appearance. In addition, the DCP algorithm is prone to color distortion when processing complex scenes. After applying the DehazeNet, AOD-Net, and DCPDN algorithms to remove haze, some residual haze remains in the image. Particularly when handling thick haze scenes, the dehazing effect is less than ideal. At the same time, the overall brightness of the image after haze removal by these three algorithms has improved, but excessive enhancement will lead to the loss of image details.

By comparison, the multi-scale adaptive dehazing algorithm based on multiple attention mechanisms proposed in this paper exhibits superior performance in handling thin haze, thick haze, and uneven haze scenes. The algorithm can produce more realistic and natural haze-free images while preserving details and texture information in the image. In addition, this algorithm avoids image artifacts and color distortion problems in the process of eliminating haze.

Comprehensive analysis reveals that the multi-scale adaptive dehazing algorithm based on multiple attention mechanisms outperforms other comparative algorithms on the drainage pipe data set.

5. Conclusions

Drainage pipe systems play a crucial role, and the environmental conditions of drainage pipelines are harsh, including darkness, moisture, heavy haze, etc. These factors have caused great troubles in the image collection of imaging equipment. The decline in image quality and loss of detailed information has brought many difficulties to the detection, maintenance and management of drainage pipelines. To solve this problem, we deeply study the dehazing technology of drainage pipe images and propose a multi-scale adaptive dehazing network based on multiple attention.

We have carefully designed a multi-scale adaptive module. The core idea of this module is to use multiple attention mechanisms to accurately capture global features at multi-scale resolutions. The design of multiple attention enables the network to achieve end-to-end defogging of drainage pipe images in the presence of uneven fog concentration and complex environments. The advantage of this method is not only its powerful dehazing

ability but also its ability to retain the detailed information in the image to the greatest extent while dehazing, making the processed image clearer, more natural, and closer to real-world vision.

In order to fully verify the performance of the dehazing network proposed in this article, we carefully constructed a huge data set, which contains tens of thousands of clear/foggy image pairs of drainage pipes. The large scale and rich diversity of the data set provide strong support for network training. By training and testing on the data set, the network can fully learn the image features in different scenes and different fog concentrations and has strong generalization capabilities.

During the experiment, we compared the proposed multi-scale adaptive dehazing network based on multiple attention with various mainstream dehazing algorithms. Experimental results show that our network has significant advantages in dehazing effect, showing excellent performance in terms of thoroughness of dehazing, retention of details, and processing speed.

More importantly, our dehazing network shows great potential in practical applications. In the specific scene of urban underground drainage pipes, image acquisition becomes extremely difficult because the internal environment of the pipes is dark, humid, and often accompanied by fog. By applying the defogging network proposed in this article, the fog can be effectively removed, and the clarity of the image restored, which provides great convenience for pipeline detection, maintenance and management. The network can also be expanded to other similar scenarios, such as image restoration in dark and humid environments such as tunnels and warehouses, which has broad market prospects and application value.

In summary, the multi-scale adaptive defogging network based on multiple attention designed and proposed in this article not only has powerful defogging capabilities, but also can provide clear and natural visual effects while maintaining image details. We will continue to optimize and improve the multi-scale adaptive dehazing network based on multi-attention. The research scope will be expanded and the dehazing technology will be applied to more similar scenes. In addition to dark and humid environments such as urban underground drainage pipes, tunnels and warehouses, we can also explore the application of defogging technology in areas such as underwater image restoration and traffic monitoring in haze weather. In addition, we will also work on how to realize the real-time operation of the dehazing network, improve the processing speed, and quickly process large amounts of image data to meet the needs of practical applications.

Author Contributions: Conceptualization, C.L. and F.Y.; Data curation, Z.T. and J.Q.; Formal analysis, Z.T.; Funding acquisition, C.L. and F.Y.; Investigation, Z.T., J.Q. and C.S.; Methodology, C.L. and Z.T.; Project administration, C.L., C.S. and F.Y.; Resources, Z.T., J.Q. and C.S.; Supervision, C.L.; Validation, C.L. and Z.T.; Visualization, Z.T.; Writing—original draft, C.L. and Z.T.; Writing—review & editing, C.L. and Z.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is sponsored by Beijing Nova Program of Science and Technology (Z211100002121147, Z191100001119106), Beijing Municipal Natural Science Foundation (4202065), National Key Research and Development Program of China (2021YFC3090304), National Natural Science Foundation of China (62176260, 62076016, 61972016). Ce Li is the corresponding author.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to this project not being complete.

Conflicts of Interest: The authors declare no conflicts of interest. Author Chi Su was employed by the company Beijing Kingsoft Cloud Network Technology Co., Ltd. and SmartMore Co., Ltd., Beijing. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wang, D.; Tan, J.; Peng, S.; Zhong, Z.; Chen, G.; Li, G. Intelligent identification system of drainage pipelines defects based on deep learning model. *Bull. Surv. Mapp.* **2021**, *141*–145.
2. Wang, H.; Zhang, Y.; Shen, H.; Zhang, J. Review of image enhancement algorithms. *Chin. Opt.* **2017**, *10*, 438–448. [[CrossRef](#)]
3. Shi, Z.; Zhu, J.; Yang, W.; Wang, X.; Huang, H. Research on Cloud Processing Method of Solar Flare Area in Remote Sensing Images. *J. Beijing Electron. Sci. Technol. Inst.* **2018**, *26*, 46–52.
4. Li, X. Research and Application of Image dehazing and Enhancement Technology. Master's Thesis, Qilu University of Technology, Jinan, China, 2019.
5. He, K.; Sun, J.; Tang, X. Single Image Haze Removal Using Dark Channel Prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2341–2353.
6. Qin, X.; Wang, Z.; Bai, Y.; Xie, X.; Jia, H. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11908–11915.
7. Liu, X.; Ma, Y.; Shi, Z.; Chen, J. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 7313–7322.
8. Shugar, D.H.; Jacquemart, M.; Shean, D.; Bhushan, S.; Upadhyay, K.; Sattar, A.; Schwanghart, W.; McBride, S.; de Vries, M.V.W.; Mergili, M.; et al. A massive rock and ice avalanche caused the 2021 disaster at Chamoli, Indian Himalaya. *Science* **2021**, *373*, 300–306. [[CrossRef](#)] [[PubMed](#)]
9. Zhang, J.; Sun, X.; Chen, Y.; Duan, Y.; Wang, Y. Single-Image Dehazing Algorithm Based on Improved Cycle-Consistent Adversarial Network. *Electronics* **2023**, *12*, 2186. [[CrossRef](#)]
10. Zhu, Q.S.; Mai, J.M.; Shao, L. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior. *IEEE Trans. Image Process.* **2015**, *24*, 3522–3533. [[PubMed](#)]
11. Berman, D.; Treibitz, T.; Avidan, S. Single Image Dehazing Using Haze-Lines. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *3*, 720–734. [[CrossRef](#)]
12. He, T.; Li, C.; Liu, R.; Wang, X.; Sheng, L. Pipeline Image Dehazing Algorithm Based on Atmospheric Scattering Model and Multi-Scale Retinex Strategy. In Proceedings of the 2019 IEEE International Conference on Unmanned Systems and Artificial Intelligence (ICUSAI), Xi'an, China, 22–24 November 2019; pp. 120–124.
13. Agrawal, A.; Raskar, R.; Chellappa, R. Edge suppression by gradient field transformation using cross-projection tensors. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 2, pp. 2301–2308.
14. Wu, D.; Velten, A.; O'toole, M.; Masia, B.; Agrawal, A.; Dai, Q.; Raskar, R. Decomposing global light transport using time of flight imaging. *Int. J. Comput. Vis.* **2014**, *107*, 123–138. [[CrossRef](#)]
15. Muhuri, A.; Goita, K.; Magagi, R.; Wang, H. Geodesic distance based scattering power decomposition for compact polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 2004412. [[CrossRef](#)]
16. Yang, F.; Tang, S. Adaptive Tolerance Dehazing Algorithm Based on Dark Channel Prior. *Algorithms* **2020**, *13*, 45. [[CrossRef](#)]
17. Yang, J.; Yang, J.; Luo, L.; Wang, Y.; Wang, S.; Liu, J. Robust Visual Recognition in Poor Visibility Conditions: A Prior Knowledge-Guided Adversarial Learning Approach. *Electronics* **2023**, *12*, 3711. [[CrossRef](#)]
18. Tian, E.; Kim, J. Improved Vehicle Detection Using Weather Classification and Faster R-CNN with Dark Channel Prior. *Electronics* **2023**, *12*, 3022. [[CrossRef](#)]
19. Zheng, M.; Luo, W. Underwater image enhancement using improved CNN based dehazing. *Electronics* **2022**, *11*, 150. [[CrossRef](#)]
20. Cai, B.; Xu, X.; Jia, K.; Qing, C.; Tao, D. DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Trans. Image Process.* **2016**, *25*, 5187–5198. [[CrossRef](#)]
21. Ren, W.; Liu, S.; Zhang, H.; Pan, J.; Cao, X.; Yang, M.H. Single image dehazing via multi-scale convolutional neural networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Cham, Switzerland, 2016; pp. 154–169.
22. He, Z.; Patel, V.M. Densely Connected Pyramid Dehazing Network. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018.
23. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. Aod-net: All-in-one dehazing network. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4770–4778.
24. Liu, R.; Fan, X.; Hou, M.; Jiang, Z.; Luo, Z.; Zhang, L. Learning aggregated transmission propagation networks for haze removal and beyond. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *30*, 2973–2986. [[CrossRef](#)]
25. Qian, W.; Zhou, C.; Zhang, D. CIASM-Net: A novel convolutional neural network for dehazing image. In Proceedings of the 2020 5th International Conference on Computer and Communication Systems (ICCCS), Shanghai, China, 15–18 May 2020; pp. 329–333.
26. Ma, J.; Yu, W.; Chen, C.; Liang, P.; Guo, X.; Jiang, J. Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. *Inf. Fusion* **2020**, *62*, 110–120. [[CrossRef](#)]
27. Engin, D.; Genç, A.; Kemal Ekenel, H. Cycle-dehaze: Enhanced cycleGAN for single image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 825–833.
28. Zhu, Y.; Newsam, S. Densenet for dense flow. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 790–794.

29. Chen, D.; He, M.; Fan, Q.; Liao, J.; Zhang, L.; Hou, D.; Yuan, L.; Hua, G. Gated context aggregation network for image dehazing and deraining. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Waikoloa Village, HI, USA, 7–11 January 2019; pp. 1375–1383.
30. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings Part III 18. Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.
31. Shahin, A.I.; Aly, W.; Aly, S. MBTFCN: A novel modular fully convolutional network for MRI brain tumor multi-classification. *Expert Syst. Appl.* **2023**, *212*, 118776. [[CrossRef](#)]
32. Huang, H.; Lin, L.; Tong, R.; Hu, H.; Zhang, Q.; Iwamoto, Y.; Han, X.; Chen, Y.W.; Wu, J. Unet 3+: A full-scale connected unet for medical image segmentation. In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 1055–1059.
33. Li, X.; Chen, H.; Qi, X.; Dou, Q.; Fu, C.W.; Heng, P.A. H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans. Med. Imaging* **2018**, *37*, 2663–2674. [[CrossRef](#)]
34. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-unet: Unet-like pure transformer for medical image segmentation. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer Nature: Cham, Switzerland, 2022; pp. 205–218.
35. Wei, W.; Li, C.; Li, S.; Chen, Z.; Yang, F. SewerOD: A visual sewer disease detection dataset for machine learning. *J. Phys. Conf. Series* **2023**, *2646*, 012011. [[CrossRef](#)]
36. Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.
37. Sara, U.; Akter, M.; Uddin, M.S. Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study. *J. Comput. Commun.* **2019**, *7*, 8–18. [[CrossRef](#)]
38. Setiadi, D.R.I.M. PSNR vs SSIM: Imperceptibility quality assessment for image steganography. *Multimed. Tools Appl.* **2021**, *80*, 8423–8444. [[CrossRef](#)]
39. Wang, S.; Rehman, A.; Wang, Z.; Ma, S.; Gao, W. SSIM-motivated rate-distortion optimization for video coding. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *22*, 516–529. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.