Sequential dual generative adversarial network for snowflake noise elimination

Yugang Li[1], Xianfu Bao[2], Zanxia Qiang[1,3], Guan Yang[1]

[1]School of Computer Science, Zhongyuan University of Technology, Zhengzhou, China
[2]School of Automation, Hangzhou Dianzi University, Hangzhou, China

[3]Qiangzanxia@163.com

Abstract. Aiming at the problem of snowflake occlusion and haze veil effect in the collected images by the vehicle-mounted camera for the influence of snowflakes and haze. The decoupling and double-supervised weather elimination network was proposed based on the sequential reconstruction of high-frequency and low-frequency components in background images. Decoupling the desnow and dehaze assignment to two sub-networks sequentially, the network filters the high-frequency feature vectors in the image background region by the eigenvalues of the convolution kernel in the spatial domain, and then the second subnetwork performs image coloring and edge fine-tuning tasks based on edge context features obtained by generative adversarial networks. The algorithm has tested on the SRRS-6000 dataset, which verified the effectiveness in the significant performance on noise removal. The Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) has reached 33.29dB/0.94, 32.8dB/0.9316, 30.13dB/0.93, 25.88dB/0.82 on Snow100K-S、Snow100K-M、Snow100K-L、I&O-Haze dataset, respectively. Experiments have shown that the decoupling and double supervised method have efficient snowflake and haze removal performance in image denoising tasks, and the adaptability of unmanned assistance systems under complex weather conditions has been enhanced.

Keywords: Snow Removal, Generative, Generative adversarial Network, Sequential Network.

1. Introduction

Outdoor monitoring systems often affect extreme weather interference and cause weather noise to remain in the video images, such as outdoor cameras and traffic monitoring systems. We have proposed many deep convolutional neural network models to solve the issue by removing weather noise from a single image [1-3]. These methods have verified that the combination of weather noise image and label background can effectively restore image background information. Although this type of algorithm performs well in a single type of weather, the network's generalization ability was still not ideal for the weather characteristics differenced from the real weather [4]. For the complexity of the actual scene, it is difficult to use a single type of weather model to restore the image background, such as haze often accompanies snow weather, and rain fog often accompanies rainfall. Moreover, there are many composite weather characteristics under real weather conditions, and the result is that the model trained under a single type of weather conditions is difficult to adapt in many complex weather scenarios. Therefore, taking the joint elimination of snowflakes and haze noise as an example, we attempt to repair the image background information from the perspective of separating edge generation from background restoration, rather than using the physical weather imaging model.

At present, many scholars have done a lot of research on single image enhancement to eliminate different types of weather noise respectively. However, it is difficult to construct a unified model for application for differences in weather's transparency, size, and distribution of different weather characteristics. In the scene where snow and haze coexist, it is difficult to adapt to a single type of physical snowfall or haze imaging model due to background occlusion and veil effect. In this paper, the high-frequency edge context feature and low-frequency color restoration of the image are separated, and the image background information is gradually restored from the perspective of the spectral and frequency domain and model decoupling. The weather image collected is composed of edge context information and image chromatic information, which corresponds to high-frequency components with rich details, and low-frequency regions in the spectrum domain. It's different from the two components of learning in that the components of the high-frequency spectrum fluctuate differently from those of the low-frequency spectrum. However, the weather noise elimination method can effectively deal with complex weather conditions from the perspective of gradual reconstruction. The contributions of this article are as follows: (1) A new solution is proposed for the background recovery of weather conditions of snowflake and haze mixed. (2) The weather feature elimination of the proposed model has achieved the same repairing performance as that of the State Of The Art (SOTA) model on the public and mixed dataset.

The remaining parts of this article are organized as follows. This paper briefly introduced the current work and its shortcomings in snow and haze elimination in Section 2. Section 3 described the proposed Snow and Haze Eliminate Network (SHEN) model in network structure and data flow. Section 4 provided the experimental result and analysis. Section 5 described the discussion and results for the network training and evaluation, and demonstrates its performance for multiple weather noises. In the end, Section 6 has given the extended application and conclusion of SHEN.

2. Related work

Deep learning has achieved good performance in various image processing problems, i.e., Dehaze, Derain, Desnow, and Style Transfer. Li [1]et al. proposed a novel single image dehazing framework DehazeFlow based on conditional regularization flow for a single image dehazing task. This method enables the model to sample multiple defogging results by learning the conditional distribution of haze images. In addition, the framework enhances the expression ability of a single network structure layer by using the attention-based coupling layer, converts natural images into potential spaces, and fuses them into features of pairs of data. N. Bharath Raj [2] et al. proposed an improved condition generation confrontation network to remove haze directly from the image without using depth information. The model uses the Tiramisu model as the generator network and uses a patch-based identifier to reduce artifacts in the output. At the same time, this model uses a mixed weighted loss function to train the model and achieves an excellent image defogging effect. Compared with the haze phenomenon, snowflake particles are difficult to remove snow due to their transparency, diversity of size, and shielding effect. Wei-Ting Chen [3] et. al remodeled the snowflake imaging model and proposed a new joint size and transparent perception snowflake elimination network JSTASR. It realizes multi-scale accommodation of snowflakes and realizes transparency perception of snowflakes. It has made remarkable improvements in eliminating snowflakes. Yun-Fu Liu [4] et. al. designed a multi-level network DesnowNet to deal with the removal of translucent and opaque snow particles. The model also distinguishes the transmittance and chromatic aberration properties of snowflake particles to accurately estimate the restored background. In addition, DesnowNet restores details covered by opaque snow by estimating snowflake complements for snowless images separately. This model realizes the snowflake elimination task with multi-scale and different transparency. These researches have shown excellent performance for snow and haze elimination, but few types of research adapt to snow and haze mixed scenes. There are still two challenges existence in dealing with the joint elimination of snowflakes and haze, namely how to adapt to the weather characteristics of different data distribution and the lacking

DOI: 10.54254/2753-8818/95/2024.21241

prior information in snowflake occlusion areas. This paper introduced two consecutive generative adversarial networks to reconstruct the image background. The first network is used to infer the edge context information of the occlusion area through the existing image background edge, and the second is used for local correction and color inpainting through the generative results of the first network. The literature's contributions are as follows: (1) Dual-GAN framework has been proposed to eliminate snow particles and haze and reconstruct clear image context from the perspective that the components of different frequencies bandwidths are gradually repaired. (2) The model has achieved good performance in snow and haze removal reached the SOTA model. (3) The model can be effectively extended to other types of composite weather, such as rain and fog weather.

3. Weather noise elimination network

The main architecture of the proposed Snow and Haze Elimination Network (SHEN) has shown in figure 1. The network structure consists of two subnets responsible for different functions: (1) Background edge context reconstruction stage: the first generative adversarial network generated the edge context information from the weather image and its Canny edge feature, namely Edge Context Network (ECN). Then the weather image and the generated edge context feature are concatenated in channel dimension to fed into the second generative adversarial network, namely Color Repairing Network (CRN). (2) Color information repairing and fine-tuning stage: the concatenated feature was fed into the CRN for color inpainting and local context adjustment. More detail for SHEN was described in the following subsections 3.1 and 3.2.

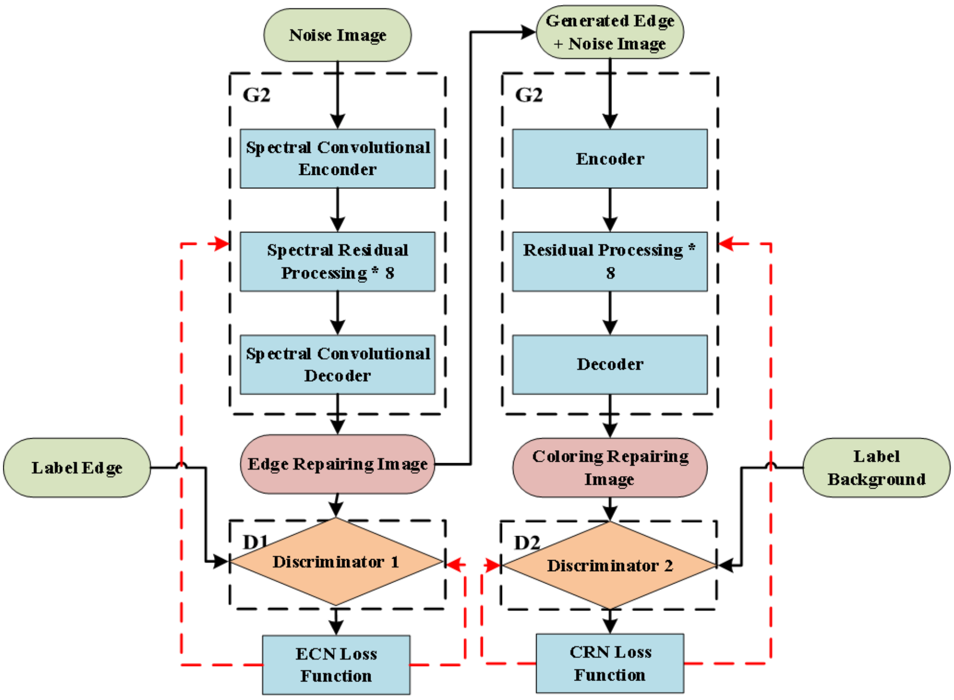


Figure 1. Proposed SHEN model. All noise images (Stream X) and Non-noise images (stream Y) from the selected training databases. The ECN network structure consists of G1 and D1. The CRN network structure consists of G2 and D2.

3.1. Decoupled network

The generator structure of SHEN has shown in figure 2, which is mainly composed of generator G1 of ECN and generator G2 of CRN. The network structure parameters of generators have been illustrated in figure 2. The ECN has a similar network structure to CRN. It is worth noting that the convolution layer structure is widely used to normalize convolution kernel parameters by spectral feature normalization

DOI: 10.54254/2753-8818/95/2024.21241

in ECN network structure. The spectral normalization to make ECN converge faster than CRN and to ensure that ECN meets the Lipschitz constraints. The ECN are used to learn high-frequency context information in images through the guidance of the Canny edge feature of noise image. The edge context feature, donate as X_{edge}, is extracted from the input weather image. Then, the CRN network reconstructed the image background information, which donatesX_{color}, according to the obtained edge context information. In addition to the generator part, the SHEN model also included the discriminators and VGG-16 pre-training network to extract the semantic feature of the image to provide feature matching loss. The convolution layer of the discriminator and VGG-16 is mainly used to extract convolution features from the generated background and label the background, providing feedback gradient loss for the generator.

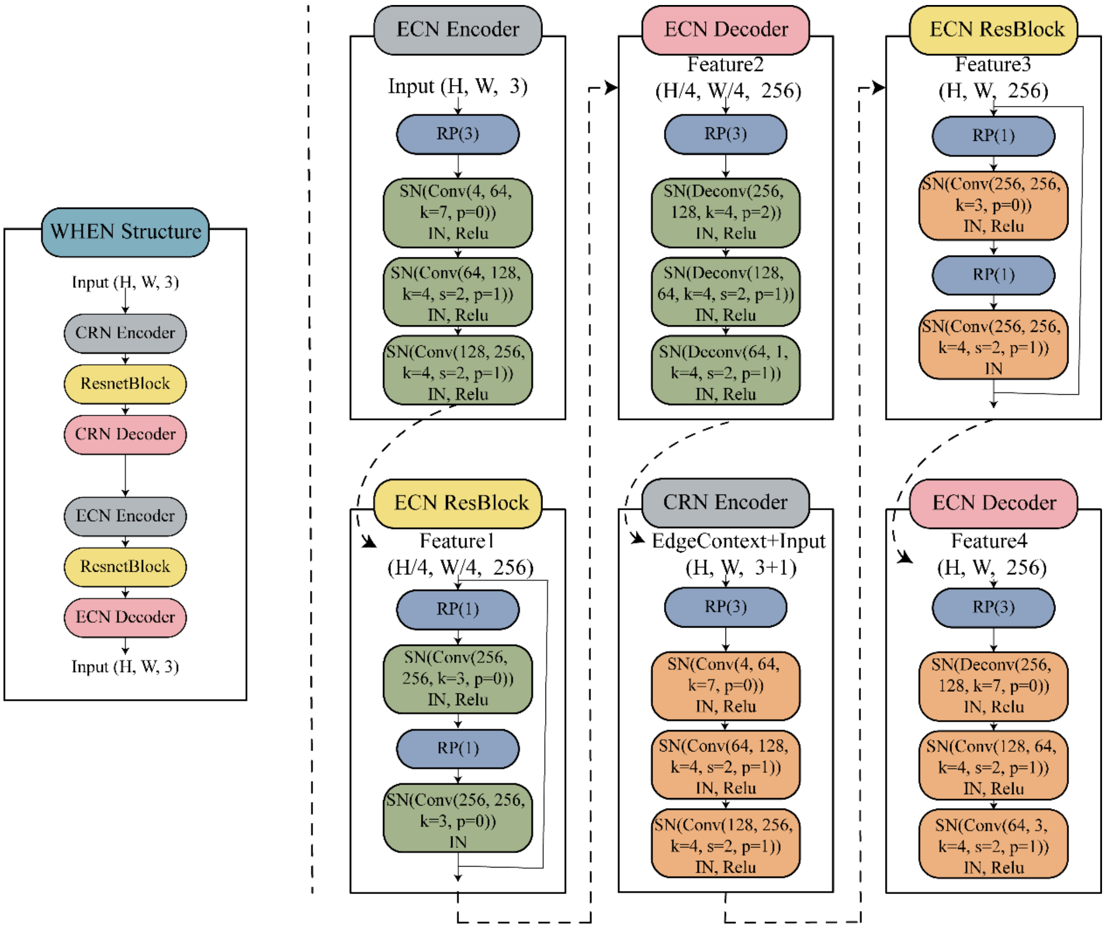


Figure 2. Proposed SHEN model. All noise image (Stream X) and Non-noise image (stream Y) X-ray images from the selected databases. Dx and Dy data flow are translated by the Dual GAN to deal with the limitation problems.

Decoupling background reconstruction tasks to the subtasks of reconstructing background edge context and color inpainting stepwise, we use sequential GAN to implement the above two subtasks respectively. ECN to learn the mapping function F1 from noisy image (X) to edge context feature (Z), and CRN to learn the mapping function F2 from the noise image (X) and edge context feature (Z) to clean background (Y). Therefore, the snow and haze image and corresponding label background are applied in the ECN to generate the edge context feature (Z), in which the high-frequency components are filtered by the singular value of the convolution kernel. The second generative adversarial network used the generated edge context feature and noise image to generate the clean background (Y). The network data flow chart has shown in figure 3. ECN is responsible for generating edge context features,

DOI: 10.54254/2753-8818/95/2024.21241

namely high spectral response characteristics of background, by learning weather images and its Canny edge. ECN uses normalized spectral convolution for convolution kernel, extracting the image's high-frequency spectral feature into high-dimensional feature space. And then ECN's generator uses the spectral convoluted residual structure for high-frequency screening and obtains edge context features through decoding and dimensionality compression. The CRN accepts the edge context features and the noise image with channel concatenation, and processes feature through CRN coding to high-dimensional space. And then CRN model carries out residual processing and decoding with dimension reduction successively, and finally obtains the background image eliminated weather noise. ECN discriminator D1 and CRN discriminator D2 were used to supervising the processes respectively and feedbacked gradient loss according to the discriminant results.

As a low-level semantic feature, the edge feature is easy to learn and inference with great help to image detail repair to occluded areas. There are differences in losses between ECN and CRN. The model uses two discriminators and adopts gradient truncation in ECN's end and joint cross-entropy loss to ensure consistency of network training, which can greatly reduce the difficulty of network convergence. There are four data flow paths in the network, namely G1- > D1, G1- > G2- > D2, D2- > G2, D1- > G1. The first two occurred in the forward reasoning process of the network training process, and the second two occurred in the transmission process of the gradient direction of the network. All original weather images and generated Canny edge are been fed to the ECN and CRN to infer the edge context feature, and repainting background. The discriminator and VGG-16 combine the generated edge context features with the repaired background, and compare with the canny edge and label background to generate the loss gradient of the ECN and CRN network structure.

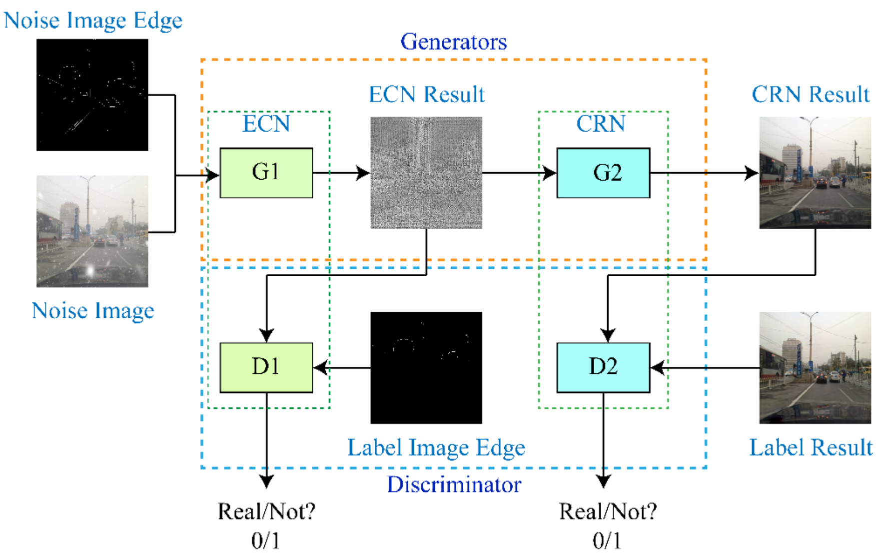


Figure 3. Our proposed model with generator adversarial loss for edge context generation and colorization.

3.2. Loss function

There are two discriminators in SHEN, namely discriminator D1 of the ECN network and discriminator D2 of the CRN network. The overall optimization objective function of the SHEN network is shown in Formula 1.

$$min_{G_1,G_2D_1,D_2}\begin{pmatrix} E_{x,z\sim P_{\text{data}}}\left[log\,D_1(z) + log\,D_2(x,z)\right] + \\ E_{z,y\sim P_{\text{latent}}(x)}\Big[log\big(1 - D_1(G_1(x))\big) + \\ log\Big(1 - D_2\big(G_2(x, G_1(x))\big)\Big)\Big] \end{pmatrix} \tag{1}$$

DOI: 10.54254/2753-8818/95/2024.21241

where, variable represents the input weather image matrix, and z represents the generated edge context feature matrix by ECN. G_1 and G_2 represent the generator of ECN and CRN, respectively. D_1 and D_2 represent the discriminator of ECN and CRN for providing cross-entropy loss for input sample and generated sample.

The two items of the ECN's loss function are shown in Formula (2, 3). The total generative loss function of ECN is shown in Formula (4):

$$L_1 = -\sum_{i=1}^{N} \frac{w_n}{N_i} \begin{pmatrix} log\left(D_1\left(\left[I_{\text{LabelEdge}}, I_{\text{Label}}\right]\right)\right) + \\ log\left(1 - D_1([C_{\text{ECN}}, C_{\text{CRN}}])\right) \\ + log\left(D_1(I_{\text{label}})\right) + \\ log\left(1 - D_1(C_{\text{ECN}} \cdot)\right) \end{pmatrix} \tag{2}$$

$$L_2 = \sum_{i=1}^{N} \frac{1}{N} \left\| \begin{matrix} D_1^{(i)}\left[C_{\text{labelEdge}}, I_{\text{label}}\right] - \\ D_1^{(i)}\left[C_{\text{ECN}}, I_{\text{pred}}\right] \end{matrix} \right\|_1 \tag{3}$$

$$L_{ECN} = \lambda_1 L_1 + \lambda_2 L_2 \tag{4}$$

Where represents the weight matrix of convolution layers in ECN and represents the number of convolution layers. N_i represents the number of parameters of the convolution weight matrix at each layer. $C_{labelEdge}$ can be obtained by applying the Canny edge detection algorithm on the label imageI_{Label}. C_{ECN} stands for edge context features generated by ECN, andC_{CRN} stands for restored background generated by CRN's generator. λ_1 and stand for weight factors, setting as 10 and 1, respectively. $i$indicate the number of the activation map of the D_2convolution layers.

The gradient truncation was applied in the generator's junction point between the CRN and ECN network, which can effectively realize the decoupling of the model. The loss function L_{CRN} of CRN is a specific definition shown in Formula. (5-7):

$$L_3 = -\sum_{i=1}^{N} \frac{w_n}{N} \begin{pmatrix} \frac{1}{2} log\left(D_2(I_{\text{label}})\right) + \\ \frac{1}{2} log\left(1 - D_2(C_{\text{CRN}})\right) \end{pmatrix} \tag{5}$$

$$L_4 = \sum_{i=0}^{i=5} \sum_{j=1}^{i=N} \frac{w_{ij}}{N_{ij}} \left\| \phi_i\left(I_{\text{label}}^i\right) - \phi_i\left(C_{\text{CRN}}^i\right) \right\|_1 \tag{6}$$

$$L_5 = \sum_{i=1}^{N} \frac{w_n}{N} \left\| G_i^{\phi}\left(I_{\text{label}}^i\right) - G_i^{\phi}\left(C_{\text{CRN}}^i\right) \right\|_1 \tag{7}$$

Where measures the weather removal performance through calculating the cross-entropy loss between repairing the background and label image I_{lable}. L_4 has been applied to measure the Euclidean distance between the activation feature map I_{label} and the feature activation map generated by the pre-trained VGG network. The variable represents the number of the pre-trained convolutional network layers. And stands for the activation map processed by the i-th convolution layers in the pre-trained VGG network. The value of style loss has measured the covariance difference between the and the I_{label}. The variable G_i^{ϕ} is the Gram matrix of the corresponding feature map through computing the convolutional activation feature map. Thus, the total generative loss expression of the CRN model have shown in Formula (8):

$$L_{CRN} = \lambda_{ECN} L_2 + \lambda_{\text{CRN}} L_3 + \lambda_P L_4 + \lambda_S L_5 \tag{8}$$

Where, the weight factor has set as λ_{ECN} =1, λ_{CRN} =2, λ_P =0.5, λ_S =240, respectively. The sum of losses was applied to the CRN's training and gradient background. The weight coefficient of the Gram matrix was added to promote CRN's learning of color information.

DOI: 10.54254/2753-8818/95/2024.21241

4. Algorithm introduction

The input noise image contains multiple frequency components, and the use of a discriminator makes the convolution kernel parameters in the generator unable to be effectively adjusted. It is easy to cause fluctuations in GAN training, especially in snowflake and haze removal tasks. By using the model decoupling method, different feedback gradients are transmitted to different sub-generators, which is conducive to the smooth adjustment of the convolution kernel in the generator. ECN initially enhances and perceives the detailed spectral feature with high frequency through singular value calculation and normalization of convolution kernels. And then, CRN reconstructs the image coloration and supplements the low-frequency part in the background. Step-by-step learning of image components can effectively reduce the difficulty of multitasking and accelerate model convergence.

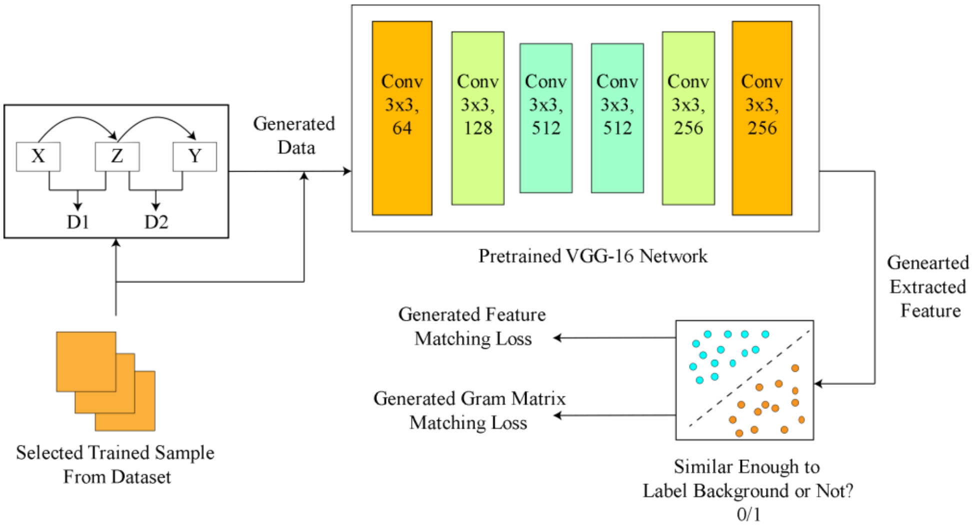


Figure 4. The Proposed SHEN model architecture for weather noise elimination.

The gradient loss generation process of pretrained VGG-16 is shown in figure 4. Data samples are screened from the training dataset and then sent into the SHEN network for training. The data samples generated by SHEN include the Canny edge of weather image, the Canny edge of label image, the edge context feature C_{ECN} and the repair background C_{CRN}. The resulting samples are then sent to the VGG-16 network to calculate feature matching loss and Gram matrix matching loss. Feature matching loss is a routinely performed process that is used to increase generalizability. However, the GANs could offer a novel approach to feature matching. Therefore, the Vgg-16 is used in the network to distinguish the error between generated feature and the label sample. The main flow of the whole process of model training has shown in figure 5. The network model first receives the noise image I_{Noise} as the sampling of noise space X, and obtains the noise image's edge through the Canny edge algorithm. Then, after the size adjustment and normalization to the interval [-1,1], the image is been sent to the ECN network for edge context feature inference. The noise image then combined the edge context featuresC_{ECN}. And the combined feature was sent to the CRN network for background reconstruction for weather noise elimination. Finally, the gradient loss feedback process of the network begins.

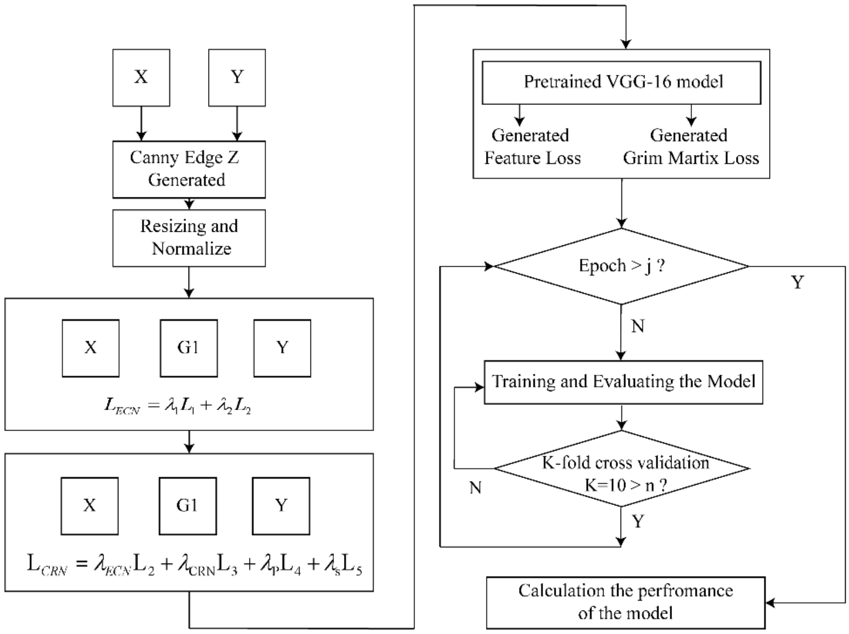


Figure 5. Flowchart of the proposed algorithm

5. Experiment

5.1. Experimental configuration

Our experiments are based on the Ubuntu 21.04 system, in which the CUDA version is 11.2, the CUDNN version is 8.2, the Pytorch version is 1.6.0, and the GPU version is RTX-5000, and the host memory is 128GB. The network model's sample inference size was 512×512, and the batch size setting was 2. The optimizers used in the training process of the ECN and CRN network was the ADAM optimizer. The initial learning rate was set to 0.02. The two learning rates were set to the same parameter and the attenuation rate was 0.90, each 1000 round is updated once, and the total training epochs of network training was 2500.

The comprehensive multi-type weather dataset, which is combined several types of public weather noise, was used in the experiment to train and evaluate the model's performance. The database consists of four types of weather datasets: (1) 16000 images and corresponding label images from SRRS Dataset [3]; (2) 16000 snow particle images and corresponding label background images from Snow100K-S [4]; (3) 16000 snow particle images and corresponding label background image from Snow100K-M [4]; (4) 16000 snow particle images and corresponding label background image from Snow100K-L [4]. (5) Two public datasets, namely I-Haze [5]and O-Haze [6], were used to test the model dehazing performance. Samples of weather noise and their label images have been to train the model in this experiment.

Table 1. Composition of desnow and dehaze dataset

Dataset	Trainset nums	Test set nums
SRRS-6000 [3]	16000	4000
Snow100K-S [4]	16600	4160
Snow100K-M [4]	16600	4160
Snow100K-L [4]	16600	4160
I-Haze [5]	—	35
O-Haze [6]	—	45
Total Number	65800	16560

DOI: 10.54254/2753-8818/95/2024.21241

5.2. Evaluation metrics

Evaluation of obtained background is an important step in evaluating model performance. Peak Signal-to-Noise Ratio (PSNR) [7] and Structural Similarity (SSIM) [7] are used for evaluating the performance of weather noise removal. The mathematical calculation formulas of the SSIM and PSNR are as follows:

$$\mathrm{SSIM(x, y)} = \frac{(2\mu_x\mu_y+\mathrm{c}_1)(2\sigma_{\mathrm{xy}}+\mathrm{c}_2)}{(\mu_x^2+\mu_y^2+\mathrm{c}_1)(\sigma_x^2+\sigma_y^2+\mathrm{c}_2)} \tag{9}$$

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{10}$$

$$\mathrm{PSNR}=10 \times \log_{10}\left(\frac{(2^n-1)^2}{MSE}\right) \tag{11}$$

where, μ_x represents the mean value of the image matrix x; μ_y represents the mean value of the image matrix y; σ_x^2 and σ_y^2 variance represents the sum of image matrices x and y; σ_{xy} represents the covariance of the image matrix x and the matrix y. The two variables stabilized the division result by increasing the value of the denominator in Eq. (3); $c_1 = (k_1L)^2, c_2 = (k_2L)^2$ which represents the variation range of image pixel values; sets k_1=0.01 and k_2=0.03. SSIM is used to measure the similarity of color value and pixel structure between generation and label image matrices. The value range of SSIM is to determine the ratio of similarity between the two sets of data struct ranging from 0 to 1. When the two sets of data are completely the same, 1 is taken, and if there is no structural similarity, 0 is taken. PSNR is the most common and widely used objective measurement method to evaluate the quality of images. However, many experimental results show that the scores of PSNR cannot be completely consistent with the visual quality seen by people. Those with higher PSNR may look worse than those with lower PSNR. Therefore, we use SSIM and PSNR to comprehensively evaluate the image quality of the generated background.

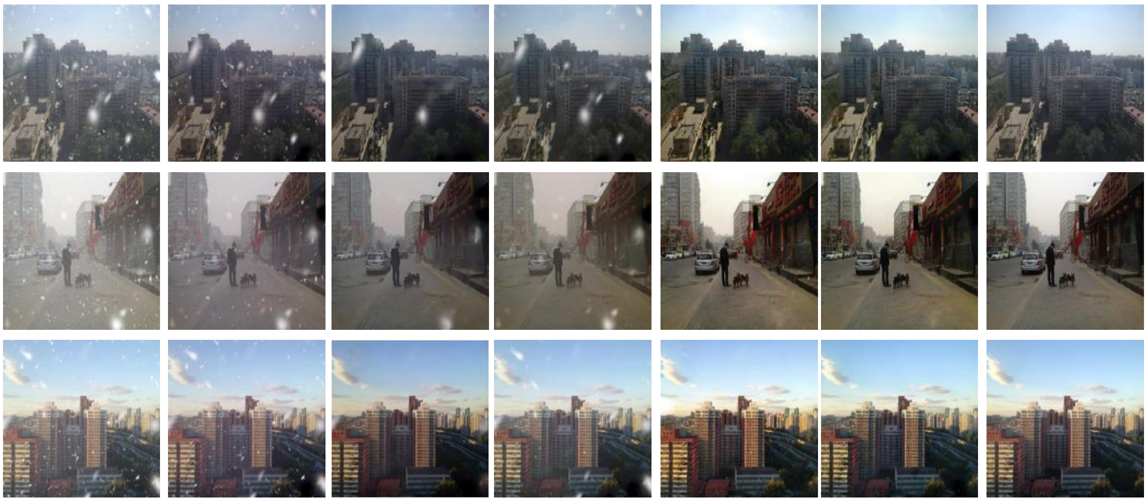
5.3. Ablation Experiment

To assess the effectiveness of decoupling generation of the image background, we conducted the ablation experiments on the network structure. We test whether the sequential network is used, whether the gradient truncation is performed, and whether the weather feature elimination performance under the network with double discriminators is used, respectively, through taking the network structure shown in figure 2 as a reference. These restored samples were generated by Network ablative structure for comparing each models' experimental results in the elimination of artifacts. The experimental results are shown in figure 6, which have showed the experimental performances for snow and haze removal between different ablation models. The structural settings of each ablation model have shown in table 2. The experimental image indicated that the SHEN framework can apply the application of snow and haze removal and obtain excellent background restoration performance. Figure 7 is a partial intermediate result generated in the network at initial stage of training, which contains the edge texture feature map generated by ECN and the background restoration map generated by CRN. In order to verify their characteristic response in frequency space, we give their Fourier response spectrum in the second line of figure 7. CRN conducts color rendering priority learning in the learning edge-intensive region, indicating that edge features have an obvious guiding function in color rendering through comparing the results of the first and third columns. In the visual display of the Fourier spectrum, it can be seen that the high and low-frequency allocation positions of the CRN's edge texture spectrum are roughly the same as the texture feature spectrum generated by ECN. However, the quantity of features learned is different, indicated that ECN has a better recognition performance on background's edge frequency, and can quickly guide the model to learn the high-frequency spectrum components. The regions generated by color features of the CRN's generation are all in the dense areas of the ECN's edge context feature, indicated that the edge texture generated by ECN plays a guiding role in the color reconstruction of

CRN’s generation, and also shows the stable convergence effectiveness of the SHEN sequence model on GAN training fluctuation.

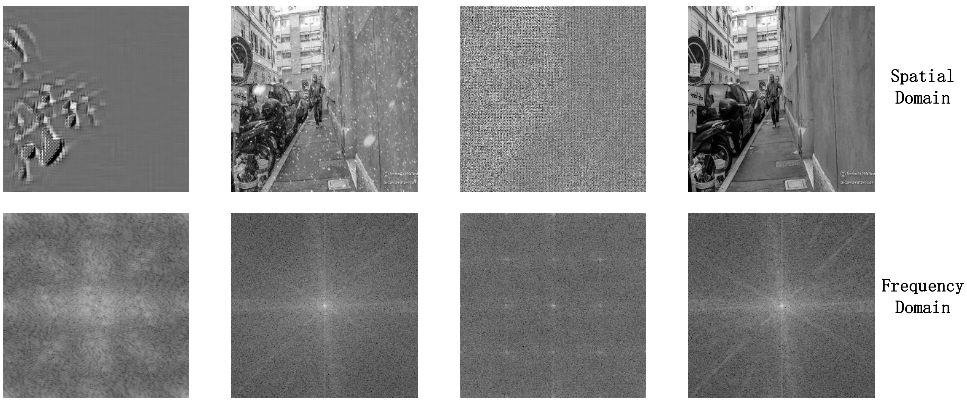
Table 2. Ablation Experiment of desnow and dehaze dataset

Structure	ECN			gradient truncation	CRN		
	Spectral Normalization of Encoder	Spectral Normalization of Encoder	Discriminator_1		Spectral Normalization of Encoder	Spectral Normalization of Decoder	Discriminator_2
Model_1	0	0	0	0	0	0	1
Model_2	0	0	1	0	0	0	1
Model_3	1	1	1	0	0	0	1
Model_4	0	0	1	1	0	0	1
Model_5	1	1	1	1	0	0	1



(a) Snow and Haze Image **(b) Model_1** **(c) Model_2** **(d) Model_3** **(e) Model_4** **(f) Model_5** **(g) Label Background**

Figure 6. Effect of ablation experiment

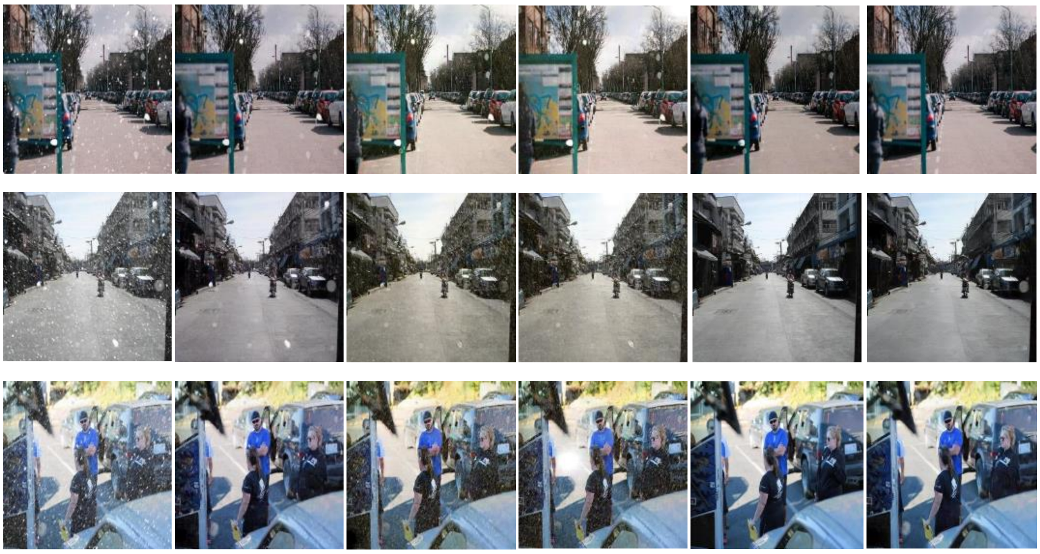


(a) Repairing Edge (b) Original Image (c) Inpainting Image (d) Label Image

Figure 7. Performance of the frequency domain

5.4. Model Comparison

The DerainNet [8], DehazeNet [9], GAN-Dehaze [2], DesnowNet [4], DuRN-S-P [10], MSBDN [11], and KDDN [12] models were used in the comparison experiment for testing the model‘s performance for snow and haze elimination in the Snow100K and O-Haze dataset. The comparison results are shown in table 3 by calculating PSNR and SSIM indexes. The results have shown that the PSNR index obtained by SHEN on dataset listed in table 3 is better than the comparison models, and the snow removal performance have shown in figure 8, and the dehaze performance on the O-Haze test set and the snow fog test performance in real scenes are shown in figure 9 and figure 10. The spectral normalizes the convolutional parameter kernel in ECN framework. Its generator of ECN satisfies the Lipschitz constraint [13], which ensured the training stability of the ECN network and provided stable prior information to the CRN network. The results of the SHEN repairing have shown that the occlusion area is better repaired, and artifacts and artificial traces are well eliminated shown as figure 8. The PSNR and SSIM obtained in the comparison experiment with other models on the test dataset have displayed in table 3. SHEN has achieved an average SSIM value of 32.84 and PSNR value of 0.93, respectively. The SSIM and PSNR of the test dataset have been explain the validity of the model that SHEN outperformed the compared SOTA models. Meanwhile, it illustrated that SHEN has the ability to perform both snow and fog removal for composite weather and achieved better performance than the SOTA model compared.



(a) Snow Image **(b) DesnowNet** **(c) JORDER** **(d) DuRN-S-P** **(e) Ours** **(g) Label Background**

Figure 8. Visual comparison with other models

Table 3. Comparison data with SOTA models

Method		Derain-Net	Dehaze-Net	GAN-Dehaze	DesnowNet	DuRN-S-P	MSBDN	KDDN	Ours
Snow100 K-S	PSNR	25.74	24.96	25.94	32.33	32.27	31.17	31.15	33.29
	SSIM	0.86	0.88	0.88	0.95	0.9497	0.93	0.9396	0.94
Snow100 K-M	PSNR	23.36	24.16	24.36	30.86	30.92	29.18	31.17	32.84
	SSIM	0.85	0.87	0.86	0.94	0.94	0.92	0.93	0.93
Snow100 K-L	PSNR	19.18	26.61	21.29	27.17	27.21	26.17	28.32	31.13
	SSIM	0.75	0.77	0.77	0.90	0.89	0.86	0.89	0.92
SRRS	PSNR	20.13	20.64	22.31	30.14	32.68	33.79	34.72	36.42
	SSIM	0.74	0.80	0.81	0.87	0.96	0.98	0.98	0.98
O+I-Haze	PSNR	15.49	19.62	22.31	16.73	18.32	21.23	19.39	25.88
	SSIM	0.51	0.59	0.74	0.52	0.61	0.71	0.59	0.82

DOI: 10.54254/2753-8818/95/2024.21241

Figure 9. O-Haze test set dehaze test

Figure 10. Real scene snow and haze elimination test

6. Conclusion

The paper proposed to decompose the image into edge context and color information and reconstruct the background information in stages for eliminating snow and haze features at one framework. For a variety of weather noise distribution differences, this paper gradually reconstructs the background edge and color method, the model decoupling, and double supervision method to solve. We use dual-generators to construct a new snow and haze elimination network, which connect two generative adversarial networks, and used the gradient truncation technique between the two generators. It can be used for the elimination of weather characteristics under complex weather and extreme weather conditions. The results of SHEN have been evaluated using datasets composed of Rain1400, SRRS, and Snow100K. The performance for snow and haze elimination has reached SSIM = 32.27 and PSNR = 0.98. In the future, it is necessary to verify the effectiveness of the model in more scenes requiring background enhancement and to evaluate the enhancement performance for different scenes. Then, this algorithm is applied to the scene of eliminating air pollution and multiple weather coexistence scenario.

References

[1] Li, H., Li, J., Zhao, D., and Xu, L., 'Dehazeflow: Multi-Scale Conditional Flow Network for Single Image Dehazing', in, Proceedings of the 29th ACM International Conference on Multimedia, (2021)

[2] Bharath Raj, N., 'Single Image Haze Removal Using a Generative Adversarial Network', arXiv e-prints, 2018, p. arXiv: 1810.09479.

[3] Chen, W.-T., Fang, H.-Y., Ding, J.-J., Tsai, C.-C., and Kuo, S.-Y., 'Jstasr: Joint Size and Transparency-Aware Snow Removal Algorithm Based on Modified Partial Convolution and Veiling Effect Removal', in, European Conference on Computer Vision, (Springer, 2020)

DOI: 10.54254/2753-8818/95/2024.21241

[4] Liu, Y.-F., Jaw, D.-W., Huang, S.-C., and Hwang, J.-N., ‘Desnownet: Context-Aware Deep Network for Snow Removal‘, IEEE Transactions on Image Processing, 2018, 27, (6), pp. 3064-3073.

[5] Ancuti, C., Ancuti, C.O., Timofte, R., and Vleeschouwer, C.D., ‘I-Haze: A Dehazing Benchmark with Real Hazy and Haze-Free Indoor Images‘, in, International Conference on Advanced Concepts for Intelligent Vision Systems, (Springer, 2018)

[6] Ancuti, C.O., Ancuti, C., Timofte, R., and De Vleeschouwer, C., ‘O-Haze: A Dehazing Benchmark with Real Hazy and Haze-Free Outdoor Images‘, in, Proceedings of the IEEE conference on computer vision and pattern recognition workshops, (2018)

[7] Hore, A. and Ziou, D., ‘Image Quality Metrics: Psnr Vs. Ssim‘, in, 2010 20th international conference on pattern recognition, (IEEE, 2010)

[8] Fu, X., Huang, J., Ding, X., Liao, Y., and Paisley, J., ‘Clearing the Skies: A Deep Network Architecture for Single-Image Rain Removal‘, IEEE Transactions on Image Processing, 2017, 26, (6), pp. 2944-2956.

[9] Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D., ‘Dehazenet: An End-to-End System for Single Image Haze Removal‘, IEEE Transactions on Image Processing, 2016, 25, (11), pp. 5187-5198.

[10] Liu, X., Suganuma, M., Sun, Z., and Okatani, T., ‘Dual Residual Networks Leveraging the Potential of Paired Operations for Image Restoration‘, in, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2019)

[11] Dong, H., Pan, J., Xiang, L., Hu, Z., Zhang, X., Wang, F., and Yang, M.-H., ‘Multi-Scale Boosted Dehazing Network with Dense Feature Fusion‘, in, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, (2020)

[12] Hong, M., Xie, Y., Li, C., and Qu, Y., ‘Distilling Image Dehazing with Heterogeneous Task Imitation‘, in, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2020)

[13] Qi, G.-J., ‘Loss-Sensitive Generative Adversarial Networks on Lipschitz Densities‘, International Journal of Computer Vision, 2020, 128, (5), pp. 1118-1140.