Unsupervised Framework for Multi-Task Restoration of Natural Images under Adverse Weather Conditions

Margarita N. Favorskaya [1], Dmitriy N. Natalenko [1]

[1] Reshetnev Siberian State University of Science and Technology, 31 Krasnoyarsky Rabochy ave., Krasnoyarsk, 660037 Russian Federation - favorskaya@sibsau.ru, dmitriy.natalenko@mail.ru

Commission II, WG II/8

Keywords: Adverse Weather Condition, Image Restoration, Unsupervised Learning, Multi-Task Restoration.

Abstract

Adverse weather conditions can be an obstacle to recognizing animals in images captured by camera traps. Obviously, recognizing night images becomes more difficult when artifacts such as snow, rain, fog, or haze appear during shooting. Typically, the model is trained to remove any artifact, and very rarely two imposed meteorological artifacts, such as snow and rain. The diversity of deep neural models indicates interest to this problem, especially using unsupervised learning when there are no paired images – without and with artifact. The aim of this study is to develop a generalized framework for natural image restoration under adverse weather conditions based on mutual GAN training to simultaneously generate clean image and improve artifact mask. The core of image restoration depends on the physical features of a particular meteorological phenomenon and can be selected for different weather conditions. Analysis of images included in the dataset captured by camera traps in Ergaki National Park, Russia, shows that the most common artifacts are snow in winter and fog in summer and autumn. These artifacts were given special attention when building the framework. Additionally, the CSD, Rain100L, and O-Hazy datasets were utilized to evaluate the effectiveness of the proposed method under various adverse weather conditions. This comprehensive approach ensures that the framework is robust and adaptable to different types of artifacts encountered in real-world scenarios.

1. Introduction

Image quality degradation caused by adverse weather conditions can severely reduce the accuracy of object detection and recognition due to the appearance of additional meteorological artifacts in cluttered urban or natural surveillance scenes. Such artifacts can be streaks, drops or whitening the entire image. Most existing removal methods focus on learning the mapping of a degraded image to its clean content using pairs of clean and degrade images. Based on this assumption, the methods decompose the degraded image into a clean image and a mask with visual distortions. However, in many scenarios such image pairs are not available, leading to the need for unsupervised learning.

Typically, the study of this problem focuses on developing the following methods (in order of decreasing interest): rain streak removal, raindrop removal, snowflake removal, snow streak removal, haze removal, and fog removal. Most methods process the degraded image in the spatial domain (Wen et al., 2024; Park et al., 2024), but streak removal methods often use the frequency domain under the assumption that streaks are characterized by high-frequency components (Hsu and Chang, 2023). However, images captured by camera traps, which are located more often in the northern forest than in the open spaces, have a different distribution of distortions. First of all, we meet with snow streaks, rain streaks and fog. Therefore, these artifacts are considered in more detail.

Methods for eliminating meteorological effects can be divided into two main types: those based on the physical model and those based on image enhancement methods. Physical-based methods prevailed until the mid-2010s, but in the era of deep learning, research has shifted to the image enhancement methods. Physical methods require multiple images, whereas currently a single degraded image can be successfully recovered. Nevertheless, single task image processing such as image rain removal, image snow removal, and dehazing, often including additional denoising, deblurring and super-resolution, does not match real world scenarios (Xiao et al., 2025).

It is worth noting that image restoration methods are mainly divided into end-to-end supervised methods and unsupervised generative methods based on prior data. Unsupervised generative methods are more suitable for real-world image restoration, which explains rapid development of such methods at present.

In this study, we consider a single-image multi-task framework including snow removal, rain removal and defogging using three kernels. The snow removal and rain removal kernels have a similar GAN-based architecture with the idea of enhancing the extracted map with weather artifact. For image defogging, a GAN-based kernel is developed that uses a Retinex module to improve image contrast and a texture restoration module to recover image texture details.

The organization of the paper is as follows. Section 2 introduces various methods for restoring images degraded by adverse weather conditions. Problem statement is formulated in Section 3. The proposed framework is described in detail in Section 4. The experimental results are discussed in Section 5, and finally Section 6 concludes the paper.

2. Related Work

The recent research trends of image restoration under adverse weather conditions are based on U-Net models, Transformers and GANs. The U-Net model is a deep learning model originally developed for medical image segmentation tasks. An

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-2-W9-2025-71-2025 | © Author(s) 2025. CC BY 4.0 License.

encoder and a decoder learn the mapping relationship from input images to output segmentation maps using skip connections. An end-to-end deep residual haze removal network (DRHNet) was proposed to restore the haze-free and rain-free-images (Wang et al., 2020). This network included context-aware encoder, nonlinear transformation and haze decoder. Compared with traditional methods such as dark channel method, DRHNet is more complex, requires a large training dataset and consumes significant computational resources. Transformer can perform parallel calculations, while Vision Transformer (ViT) divides images into small blocks for parallel processing. These architectures actively use the attention mechanism. For example, a progressive network for image deraining based on detail scaling employed optimized transformer blocks for texture extraction (Huang et al., 2024). The CNN-ViT hybrid model based on wavelet transform was developed for image desnowing (Dai et al., 2024). GAN consisting of a generator and a discriminator uses adversarial learning to generate high-quality samples. It is mainly applied for both image generation and restoration. An end-to-end two stage conditional generative adversarial network applied a coarse-to-fine single image de-rained method to remove rain streaks (Wang et al., 2021). Two-stage generator produced more clear de-rained images with good visual quality and the discriminator enhanced the de-raining result. Conditional GAN (UnfairGAN) can effectively preserve essential details caused by heavy raindrops thanks to its advanced attention module and advanced activation function (Nguyen et al., 2022).

In the field of image restoration, a unified processing concept for image restoration under multi-weather conditions, called All-In-One, was introduced in 2020 and has since attracted wide attention. Let us consider several interesting approaches using U-Net, Transformer and GAN architectures.

Initially, the method for simultaneously learning multiple weather degradations assumed the translation into three domains: hazy, rain with veil and snow with veil (Patil et al., 2023). Then, a progressive multi-domain feature alignment module effectively extracted invariant cues across different weather conditions and the results were merged with a cascaded multi-head attention module. The practical advantage was that the proposed feature extractor can be used for real data even if it was trained using synthetic data. Another learning strategy based on a multi-teacher network and student network was proposed in (Chen et al., 2022). The two-stage knowledge distillation mechanism included knowledge collation and knowledge examination. Also multi-contrastive knowledge regularization function was embedded in the two-stage knowledge learning process. Three types of weather were investigated: haze, rain, and snow.

A transformer-based end-to-end model called as TransWeather compensated the rain, fog, and snow artifacts in corrupted image (Valanarasu et al., 2022). TransWeather had a specific transformer encoder with intra-patch transformer blocks to effectively remove non-significant weather degradations. The transformer decoder for restoring a clear image including small-scale weather degradations utilized learnable weather-type queries to identify different types of weather degradation. In the work (Luo et al., 2023), a neural network called "mixture-of-experts" first applied in natural language processing was used to process different types of blind adverse weather conditions. The proposed transformer-based weather-aware multi-scale mixture-of-experts framework included special key components such as weather-aware router, weather guidance fine-grained contrastive learning and multi-scale experts. Weather guidance fine-grained contrastive learning separated the weather (rain, raindrops, snow, and haze) and content information from the input image. An efficient histogram transformer (Histoformer) for unified adverse weather removal categorized pixel values that are similar in intensity but different in spatial location into histogram bins (Sun et al., 2024). To facilitate the extraction of complex features at both local and global scales, two types of histogram self-attention were utilized: bin-wise histogram reshaping and frequency-wise histogram reshaping. Weather-affected pixels and background pixels were divided into separate categories in descending order of intensity.

Recently, GAN-based image restoration methods have attracted significant attention in many field of image processing. The current topic is no exception. One of the first works was the GAN model, where the generator consisted of several task-specific encoders, each of which was associated with a specific adverse weather condition (Li et al., 2020). In addition, these authors introduced a series of tensor-based operations that encapsulate the basic physical principles of rain, fog, snow, and adherent raindrops. The discriminator simultaneously evaluated the correctness and classified the degradation type of the resorted image. Multiple weather translation GAN was a dual-purpose framework for simultaneously learning weather generation and removing artifacts from image data (Yang et al., 2023). The CycleGAN-based model consisted of four generators and four discriminators for haze, rain, snow, and clear weather conditions. Three GANs were used to generate three weather effects separately, and one GAN removed the weather effect.

Numerous fully supervised methods, including those discussed, require the paired images – distorted and clean images. While they achieved good results using synthetic datasets, they suffer from poor generalization when applied to real scenes. To improve the generalization ability, semi-supervised and unsupervised methods are recommended. This requires further exploration of learning paradigms, architectures and experiments with real unpaired images under adverse weather conditions.

3. Problem Statement

A commonly used mathematical model of a rainy or snowy input image involves decomposing it into two components: a clean image and an imposed rain or snow masks:

$$I_r = I_{cl} + S_r \; ; \; I_s = I_{cl} + S_s \,, \tag{1}$$

where I_{cl} = clean image
I_r, I_s = input image of rain or snow
S_r, S_s = intensity fluctuations caused by rain and snow

Intensity fluctuations S_r and S_s can have different shapes, such as rain streaks, raindrops, snowflakes or snow streaks. Usually they are defined by mask creation, which is not an easy task, especially in unsupervised learning. To obtain a clean image we need to subtract the masked regions from the input image and then reconstruct these regions into the clean image:

$$I_{cl} = f_{rain}\left(I_r - S_r\right) ; \; I_{cl} = f_{snow}\left(I_s - S_s\right) , \tag{2}$$

where f_{rain} = rainy image restoration function
f_{snow} = snowy image restoration function

A view of functions f_{rain} and f_{snow} can be different, for example based on inpainting or contrastive learning in the spatial domain or high-frequency analysis in the transform domain.
The nature of fog and haze is different from the nature of rain and snow, so images resulting from fog and haze should be processed in another way. There are two main methods for image defogging: one based on the atmospheric scattering theory with better texture extraction and the other based on Retinex theory with improved contrast but loss of texture details. Since we do not need detailed texture extraction in the animal recognition task, the second approach was chosen. We do not have paired images – a foggy image and a foggy-free image, which is also in favour of Retinex.

The Retinex model assumes that an image is decomposed into two components: reflection and illumination:

$$I_f = R \cdot L, \tag{3}$$

where I_f = foggy image
R = reflection component, desired recovered image I_{cl}
L = illumination component

The Retinex algorithm assumes applying Gaussian convolution to the illumination map of the foggy image, which results in artifacts in the recovered image. In work (Galdran et al., 2018), this drawback was overcome using inverted intensities based on the following equation:

$$f_{fog}\left(I_f\right) = 255 - Retinex\left(255 - I_f\right), \tag{4}$$

where f_{fog} = fog removal function, RGB component composition

This approach was initially applied for image dehazing (Galdran et al., 2018) rather than image defogging due to high fog concentration and the inability to estimate the Gaussian coefficients. The Retinex network (RetNet) proposed in (Liu et al., 2019) helped solve the problem of eliminating fogging with the good visualization.

4. Proposed Framework

The proposed training strategy is presented in Figure 1. The main idea is that Desnowing, Deraining and Defogging modules learn to remove the corresponding artifact by learning to generate distorted image. This idea was originally proposed (Wang et al., 2024) as a good solution for rain streak removal for unpaired images, i.e., unsupervised learning. In this study, we extend this idea to snow and fog removal, with a view to future implementation of the All-In-One learning strategy. The architectures of Desnowing and Deraining modules are similar. The first branch is trained on snow images, and the second branch is trained on snow-free images, with the parameters of both branches shared during training. The Mask queue plays a role of a semi-dictionary when not only a single mask can be imposed to a synthetic snow-free image, but all available masks are applied to select the best option.

First, a real snowy image I_s is fed on the GENERATOR desnowing network $G_{s\text{-}f}$ and transformed into a synthesised snow-free image $J_{s\text{-}f}$. Then the real snowy image I_s and the synthetic J_s are fed to DISCRIMINATOR snowing D_s. The architectures of generators desnowing and snowing are similar but they solve direct and inverse problems. It is evident that a single image enhancement is a mathematically ill-posed problem when recovering image J from image I. Second, mask is extracted and placed it into the queue. Third, the generated snow-free image $J_{s\text{-}f}$ and the mask are fed to the GENERATOR snowing network G'_s to produce the synthesised snowy image J'_s. Then the synthetic $J_{s\text{-}f}$ and $J'_{s\text{-}f}$ are fed to DISCRIMINATOR desnowing network $D'_{s\text{-}f}$. The generators are trained using the patch-wise contrastive learning method proposed in the work (Park et al., 2020).

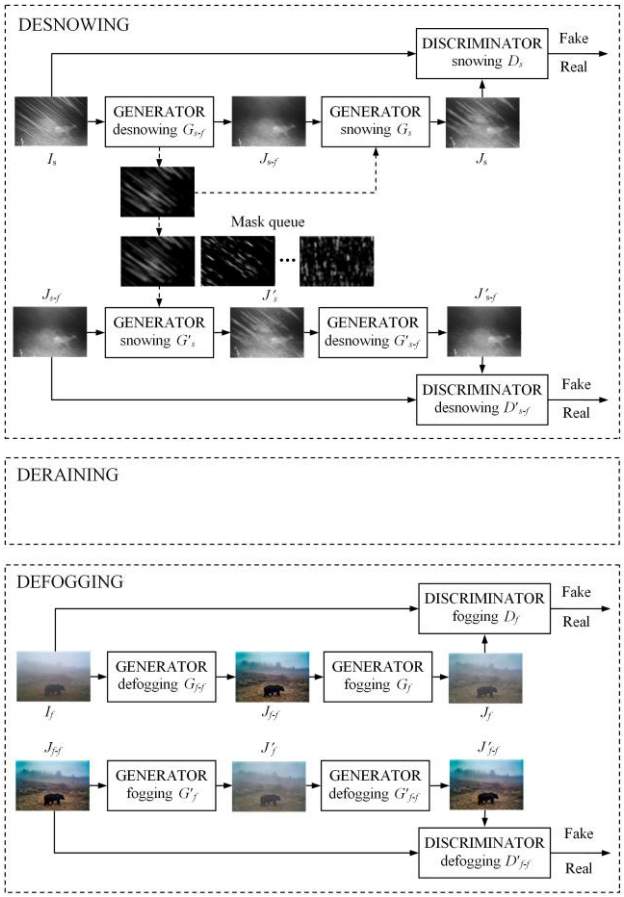

Figure 1. The proposed training strategy.

The architecture of the defogging module is the same, but the architectures of the generators are fundamentally different. The GENERATOR defogging network $G_{f\text{-}f}$ creates the synthesised foggy-free image $J_{f\text{-}f}$ using the inverted real input image I_f, which is fed to the RetNet (Retina network), and then re-inverts the output result. At the same time, the GENERATOR fogging network G_f uses the atmospheric scattering model (Liu et al., 2019):

$$J'(x) = J(x) \cdot T(x) + A(x)\left(1 - T(x)\right), \tag{5}$$
$$T(x) = e^{-\beta d(x)}$$

where $J'(x)$ = the synthesised foggy image
$J(x)$ = the synthesised foggy-free image
$T(x)$ = the transmission map as the portion of light which reaches the camera from the object
$A(x)$ = the atmospheric light
x = the image point
β = the scattering coefficient of the atmosphere

$d(x)$ = the depth from scene point to the camera

The two networks included in the GENERATOR fogging sub-module – AtmLightNet and TranNet have an encoder-decoder architectures and are trained to simulate the unknown parameters $A(x)$ and $T(x)$, respectively.

Loss functions for the Desnowing, Deraining and Defogging modules are composite functions. The loss functions for Desnowing and Deraining modules are identical and include adversarial loss, identity mapping loss, cycle consistency loss, and patchwise contrast learning loss. The $L_{GAN}(\cdot)$ adversarial losses for two GANs included in the Desnowing module have a view:

$$\begin{aligned} L_{GAN_s}\left(G_s, D_s\right) &= \mathrm{E}_{I_s \sim P(I_s)}\left[\log D_s\left(I_s\right)\right] \\ &+ \mathrm{E}_{J_s \sim P(J_s)}\left[\log\left(1 - D_s\left(G_s\left(I_s\right)\right)\right)\right] \\ L_{GAN_{s-f}}\left(G'_{s-f}, D'_{s-f}\right) &= \mathrm{E}_{J_{s-f} \sim P\left(J_{s-f}\right)}\left[\log D'_{s-f}\left(J'_{s-f}\right)\right] \\ &+ \mathrm{E}_{J_{s-f} \sim P\left(J_{s-f}\right)}\left[\log\left(1 - D'_{s-f}\left(G'_{s-f}\left(J'_{s-f}\right)\right)\right)\right] \end{aligned}, \quad (6)$$

where P = data distribution
$G_{(\cdot)}$ and $D_{(\cdot)}$ = GENERATOR and DISCRIMINATOR
I_s, $J_{(\cdot)}$ = original snowy image and synthesised image

The $L_{idt}(\cdot)$ identity mapping loss optimises the two generators in order to keep the colour unchanged (Taigman et al., 2016):

$$\begin{aligned} L_{idt_s}\left(G_s\right) &= \mathrm{E}_{J_s \sim P(J_s)}\left[\left\|G_s\left(J_s, M\right) - J_s\right\|_1\right] \\ L_{idt_{s-f}}\left(G'_{s-f}\right) &= \mathrm{E}_{J'_s \sim P(J'_s)}\left[\left\|G'_{s-f}\left(J'_s\right) - J'_s\right\|_1\right] \end{aligned}, \quad (7)$$

where $\|\cdot\|_1$= L1 norm of regularization
M = mask

Minimizing cycle consistency loss (from CycleGAN) helps to achieve similarity between the recovered image and the original snow-free image:

$$\begin{aligned} L_{cyc_s}\left(G_{s-f}, G_s\right) &= \mathrm{E}_{I_s \sim P(I_s)}\left[\left\|G_s\left(G_{s-f}\left(I_s\right), M\right) - I_s\right\|_1\right] \\ L_{cyc_{s-f}}\left(G'_s, G'_{s-f}\right) &= \mathrm{E}_{J_{s-f} \sim P\left(J_{s-f}\right)}\left[\left\|G'_{s-f}\left(G'_s\left(J_{s-f}, M\right)\right) - J_{s-f}\right\|_1\right] \end{aligned}. \quad (8)$$

The patch-wise contrast learning loss called PatchNCE (Park et al., 2020) allows to generate images with more realistic details and textures. The total loss function for the Desnowing module is summarized as follows:

$$\begin{aligned} L_s\left(G_{(\cdot)}, D_{(\cdot)}\right) &= L_{GAN_s}\left(G_s, D_s\right) + L_{GAN_{s-f}}\left(G'_{s-f}, D'_{s-f}\right) \\ &+ \alpha_s\left(L_{idt_s}\left(G_s\right) + L_{idt_{s-f}}\left(G'_{s-f}\right)\right) \\ &+ \beta_s\left(L_{cyc_s}\left(G_{s-f}, G_s\right) + L_{cyc_{s-f}}\left(G'_s, G'_{s-f}\right)\right) \\ &+ \gamma_s L_{PatchNCE}\left(G_{s-f}\right) \end{aligned}, \quad (9)$$

where α_s, β_s, γ_s = hyperparameters for training the Desnowing module

The total loss function for the Defogging module includes adversarial loss, identity mapping loss and cycle consistency loss and has a view:

$$\begin{aligned} L_f\left(G_{(\cdot)}, D_{(\cdot)}\right) &= L_{GAN_f}\left(G_f, D_f\right) + L_{GAN_{f-f}}\left(G'_{f-f}, D'_{f-f}\right) \\ &+ \alpha_f\left(L_{idt_f}\left(G_f\right) + L_{idt_{f-f}}\left(G'_{f-f}\right)\right) \\ &+ \beta_f\left(L_{cyc_f}\left(G_{f-f}, G_f\right) + L_{cyc_{f-f}}\left(G'_f, G'_{f-f}\right)\right) \end{aligned}, \quad (10)$$

where α_f, β_f = hyperparameters for training the Defogging module

All loss components for the Defogging module are calculated using Equations 6-8, but without the M mask in Equations 7-8, which are defined the identity mapping loss and cycle consistency loss.

5. Experimental Results

Experiments were conducted using a dataset collected from camera traps in Ergaki Park, Russia and images found on the Internet. The effectiveness of our framework was evaluated using quantitative metrics as well as qualitative visual comparisons.

It is worth noting that about 10% of the images include various weather artifacts such as snow, rain, and fog. Therefore, additional datasets were used to create synthetic images: CSD dataset for snowfall (Chen et al., 2021), Rain100L dataset for rain (Yang et al., 2017), and O-Hazy dataset for fog (Ancuti et al., 2018). During the pre-training of the GAN models, all images were divided into positive and negative examples. The goal of using positive examples is to minimize the distance between their representations in the model's feature space. This helps the model understand which elements remain unchanged despite the added weather effects. For negative examples, the goal is to maximize the distance between their representations. This allows the model to learn to distinguish what exactly makes images different and to understand how weather conditions affect visual features.

To evaluate the obtained results on synthetic image sets, we used two quality assessment metrics: PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). PSNR is a widely used metric for assessing the quality of recovered images compared to their original counterparts. It measures the peak error between the two images, providing an indication of the maximum possible signal-to-noise ratio. The PSNR is expressed in dB and is calculated using the mean squared error (MSE) between the original and processed images. A higher PSNR value indicates better quality, as it suggests that the recovered image is closer to the original, with less distortion.

SSIM is a perceptual metric that quantifies the similarity between two images based on structural information. Unlike PSNR, which relies solely on pixel-wise differences, SSIM considers changes in structural patterns, luminance, and contrast. The metric assesses the correlation between local patterns of pixel intensities in the original and processed images, providing a more holistic evaluation of perceived visual quality. The SSIM value ranges from –1 to +1, where a value of 1 indicates perfect structural similarity. This makes SSIM particularly useful for applications where human perception of quality is critical.

To evaluate the obtained results on real images set, we used three no-reference metrics: BRISQUE, NIQE and Entropy. BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) is a metric designed to evaluate the quality of images without the need for references. It analyzes local patterns and statistics of images to determine how distorted or degraded they are. In the evaluation process, each image is divided into small blocks, for which statistical features are calculated. The extracted features are compared with a model based on statistical data.

NIQE (Naturalness Image Quality Evaluator) is another metric for evaluating the quality of images without using references. It measures the naturalness of images based on statistical models trained on a set of natural images. In the evaluation process, various characteristics such as texture and contrast are calculated for each image. The extracted features are still compared with certain statistical models.

Entropy is a measure of the uncertainty or complexity of an image. It is used to evaluate the diversity and information content of an image. In the evaluation process, all images are converted to grayscale to simplify the analysis. For each brightness level, the probability of its appearance in the image is calculated. After that, using Shannon's formula, the entropy value is calculated using Equation 11:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i), \quad (11)$$

where $H(X)$ = the entropy of the random variable
$p(x_i)$ = the probability of occurrence of the event x
n = the total number of possible events

The average objective values for the synthetic images and for the real images with weather artifacts and the corresponding recovered images are presented in Tables 1 and 2, respectively.

Dataset	Image Type	PSNR, dB ↑	SSIM ↑
CSD	Snowy images	14.27	0.70
	Recovered images	18.94	0.83
Rain100L	Rainy images	25.52	0.81
	Recovered images	30.87	0.89
O-Hazy	Foggy images	13.68	0.59
	Recovered images	17.12	0.62

Table 1. Comparative evaluation of synthetic images with weather artifacts and recovered images

Image Type	BRISQUE ↓	NIQE ↓	Entropy↓
Snowy images	40.0	4.8	7.8
Recovered images	32.5	4.1	7.2
Rainy images	38.5	4.2	7.0
Recovered images	29.8	3.7	6.8
Foggy images	42.0	4.6	7.6
Recovered images	35.2	4.5	7.5

Table 2. Comparative evaluation of real images with weather artifacts and recovered images

The PSNR and SSIM values show that the quality of the recovered images from the CSD and O-Hazy datasets is slightly lower compared to the images from the Rain100L dataset. This indicates that the model is able to preserve the structural integrity of rain-affected images, resulting in a more visually coherent reconstruction. In the case of snow and foggy images, the values indicate the complexity of the artifacts present. However, the results are clearly positive, as most of the artifacts are either eliminated or masked quite well.

The BRISQUE metric values indicate that the quality of rainy and snowy images improved significantly compared to foggy images. This observation suggests that the model demonstrates superior performance in removing rain artifacts, likely due to the relatively less complex nature of rain-induced distortions compared to those caused by fog, which often results in more complex visual challenges.

The NIQE results further support the findings of the BRISQUE metric, showing that rainy images tend to look more natural than their snowy and foggy counterparts. This trend highlights the model's success in effectively restoring visual characteristics in rainy conditions, where preserving naturalness is paramount.

High entropy values for foggy and snowy images imply that these images exhibit greater variety and complexity, which can complicate the restoration process. The presence of numerous artifacts and distortions in these images presents significant challenges for accurate restoration. In contrast, the lower entropy values associated with rainy images suggest that they were processed more successfully, resulting in a clearer and more coherent visual output.

Visual examples of the synthetic and real snowy images are presented in Figure 2. Visual evaluation of both synthetic and real snow images shows good results. Monochrome backgrounds provide the best visual results of snowfall removal. Large snow-covered areas and areas with complex backgrounds present the greatest challenge.

a) b) c)

Figure 2. Visual comparison of synthetic (1st and 2nd rows) and real (3rd row) snow images: a) ground images; b) snow images; c) recovered images.

Visual examples of synthetic and real rainy images are presented in Figure 3. Visual evaluation of both synthetic and real rainy images also shows good results. The most difficult ones are those with heavy rain.

a) b) c)

Figure 3. Visual comparison of synthetic (1st and 2nd rows) and real (3rd row) rainy images: a) ground images; b) rainy images; c) recovered images.

Visual examples of synthetic and real foggy images are presented in Figure 4. Visual evaluation of both synthetic and real images with fog shows good results. Lighting correction allows to preserve image contrast.

a) b) c)

Figure 4. Visual comparison of synthetic (1st and 2nd rows) and real (3rd row) foggy images: a) ground images; b) foggy images; c) recovered images.

Visual presentation of the resulting images confirms the data obtained using the objective metric. The resulting images indeed have significantly fewer meteorological artifacts, making it easier to distinguish the desired object against a complex background. Unfortunately, not all generated areas have the same clarity, but even in this case, the images retain sufficient detail. Additional use of other image processing methods can help preserve even more useful information.

Overall, the results demonstrate that the proposed framework excels in processing images affected by various weather effects. The processed images exhibit a high level of quality with notable preservation of detail and satisfactory removal of weather artifacts. These findings highlight the potential of framework to enhance image restoration tasks across various adverse weather conditions, paving the way for future applications in fields such as image processing, surveillance, and environmental monitoring.

6. Conclusions

In this study, we developed an unsupervised framework for restoring natural images affected by adverse weather conditions, focusing on the removal of snow, rain, and fog using a GAN-based architecture. Our framework demonstrates significant performance according to reference and no-reference metrics. Improvements in both image quality and object visibility are observed. Its modular design provides adaptability to various weather conditions, making it suitable for applications like wildlife monitoring and environmental studies. Each component, especially the mask queue, is crucial for improving restoration quality. Looking ahead, we plan to extend our framework to tackle more complex weather scenarios and explore its practical applications in diverse settings. Overall, our research significantly contributes to the field of unsupervised image restoration by offering a robust solution for enhancing image quality in challenging environments.

References

Ancuti C.O., Ancuti C., Timofte R., De Vleeschouwer C., 2018: O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. *IEEE Conference on Computer Vision and Pattern Recognition, NTIRE Workshop,* Salt Lake City, Utah, USA, *arXiv preprint arXiv:1804.05101*.

Chen, W.-T., Fang, H.-Y., Hsieh, C.-L., Tsai, C.-C., Chen, I., Ding, J.-J., 2021: All snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss. *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4196–4205.

Chen, W.-T., Huang, Z.-K., Tsai, C.-C., Yang, H.-H., Ding, J.-J., Kuo, S.-Y., 2022: Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: Toward a unified model. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, New Orleans, LA, USA, pp.17653-17662.

Dai, X., Zhou, Y., Qiu, X., Tang, H., Tan, T., Zhang, Q., Tong, T., 2024: WaveFrSnow: Comprehensive perception wavelet transform frequency separation transformer for image snow removal. *Digital Signal Processing* 155, 104715.1-104715.13.

Galdran, A., Alvarez-Gila, A., Bria, A., Vazquez-Corra,l J., Bertalmo, M., 2018: On the duality between Retinex and image dehazing. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Salt Lake City, UT, USA, vol. 3, pp. 8212-8221.

Huang, J., Tang, Z., He, X., Jun Zhou, Zhou, D., Chen, C.Y.-C., 2024: Progressive network based on detail scaling and texture extraction: A more general framework for image deraining. *Neurocomputing* 568, 127066.1-127066.11.

Hsu, W.-Y., Chang, M.-C., 2023: Recurrent wavelet structure-preserving residual network for single image deraining. *Pattern Recognition* 137, 109294.1-109294.17.

Li, R., Tan, R.T., Cheong, L.-F., 2020: All in one bad weather removal using architectural search. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, WA, USA, pp. 3175-3185.

Liu, W., Yao, R., Qiu, G., 2019: A physics based generative adversarial network for single image defogging. *Image and Vision Computing* 92, 103815.1-103815.15.

Luo, Y., Zhao, R., Wei, X., Chen, J., Lu, Y., Xie, S., Wang, T., Xiong, R., Lu, M., Zhang, S., 2023: MoWE: Mixture of weather experts for multiple adverse weather removal. *arXiv preprintarXiv:2303.13739*, pp. 1-12.

Nguyen, D.M., Le, T.P., Vo, D.M., Lee, S.-W., 2022: UnfairGAN: An enhanced generative adversarial network for raindrop removal from a single image. *Expert Systems With Applications* 210, 118232.1-118232.12.

Park, T., Efros, A.A., Zhang, R., Zhu, J.-Y., 2020: Contrastive learning for unpaired image-to-image translation. *In: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision – ECCV 2020*, LNCS, vol. 12354, Springer, Cham, pp. 319–345.

Park, S.Y., Park, T.H., Eom, I.K., 2024: Iterative image rain removal network using consecutive residual long short-term memory. *Neurocomputing* 589, 127752.1-127752.1.

Patil, P.W., Gupta, S., Rana, S., Venkatesh, S., Murala, S., 2023: Multi-weather image restoration via domain translation. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Paris, France, pp. 21696-21705.

Sun, S., Ren, W., Gao, X., Wang, R., Cao, X., 2025: Restoring images in adverse weather conditions via histogram transformer. *In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024*. LNCS, Springer, Cham, vol. 15080, pp. 111-129.

Taigman, Y., Polyak, A., Wolf, L., 2016: Unsupervised cross-domain image generation., *arXiv preprint arXiv:1611.02200*, pp. 1-14.

Valanarasu, J.M.J., Yasarla, R., Patel, V.M., 2022: TransWeather: Transformer-based restoration of images degraded by adverse weather conditions. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, New Orleans, LA, USA, pp. 2353-2363.

Wang, C., Li, Z., Wu, J., Fan, H., Xiao, G., Zhang, H., 2020: Deep residual haze network for image dehazing and deraining. *IEEE Access*, 8, 9488-9500.

Wang, J., Gai, S., Huang, X., Zhang, H., 2021: From coarse to fine: A two stage conditional generative adversarial network for single image rain removal. *Digital Signal Processing* 111, 102985.1-102985.12.

Wang, P., Wang, P., Chen, M., Lau, R.W.H., 2024: Mask-DerainGAN: Learning to remove rain streaks by learning to generate rainy images. *Pattern Recognition* 156, 110840.1-110840.10.

Wen, Y., Gao, T., Zhang, J., Zhang, K., Chen, T., 2024: From heavy rain removal to detail restoration: A faster and better network. *Pattern Recognition* 148, 110205.1-110205.9.

Xiao, H., Liu, S., Zuo, K., Xu, H., Cai, Y., Liu, T., Yang, Z., 2025: Multiple adverse weather image restoration: A review, *Neurocomputing* 618, 129044.1-129044.20.

Yang, H., Carballo, A., Zhang, Y., Takeda, K., 2023: Framework for generation and removal of multiple types of adverse weather from driving scene images. *Sensors* 23(3), 1548.1-1548.16.

Yang, W., Tan R.T., Feng J., Liu J., Guo Z., Yan S., 2017: Deep joint rain detection and removal from a single image. *arXiv preprint arXiv:1609.07769*, pp. 1685–1694.