All in One Bad Weather Removal using Architectural Search

Ruoteng Li[1], Robby T. Tan[1,2], and Loong-Fah Cheong[1]

[1]National University of Singapore
[2]Yale-NUS College

Abstract

Many methods have set state-of-the-art performance on restoring images degraded by bad weather such as rain, haze, fog, and snow, however they are designed specifically to handle one type of degradation. In this paper, we propose a method that can handle multiple bad weather degradations: rain, fog, snow and adherent raindrops using a single network. To achieve this, we first design a generator with multiple task-specific encoders, each of which is associated with a particular bad weather degradation type. We utilize a neural architecture search to optimally process the image features extracted from all encoders. Subsequently, to convert degraded image features to clean background features, we introduce a series of tensor-based operations encapsulating the underlying physics principles behind the formation of rain, fog, snow and adherent raindrops. These operations serve as the basic building blocks for our architectural search. Finally, our discriminator simultaneously assesses the correctness and classifies the degradation type of the restored image. We design a novel adversarial learning scheme that only backpropagates the loss of a degradation type to the respective task-specific encoder. Despite being designed to handle different types of bad weather, extensive experiments demonstrate that our method performs competitively to the individual and dedicated state-of-the-art image restoration methods.

1. Introduction

Bad weather image restoration problem has been studied intensively in the research fields of image processing and computer vision; examples include deraining [20, 18, 59, 8, 53, 30, 50, 62, 36, 4, 44, 29], dehazing/defogging [47, 3, 1, 57, 7, 13, 26, 43], desnowing [44, 37], and adherent raindrops removal [41, 42], etc.. Most of these works focus only on single weather types and propose dedicated solutions [29, 43, 41]. While they can attain excellent performance, they may not yield optimal results on other types of bad weather degradations, since the factors that cause the degradations in other types are not carefully considered. As a result, real outdoor systems would have to decide and switch between a series of bad weather image restoration algorithms.

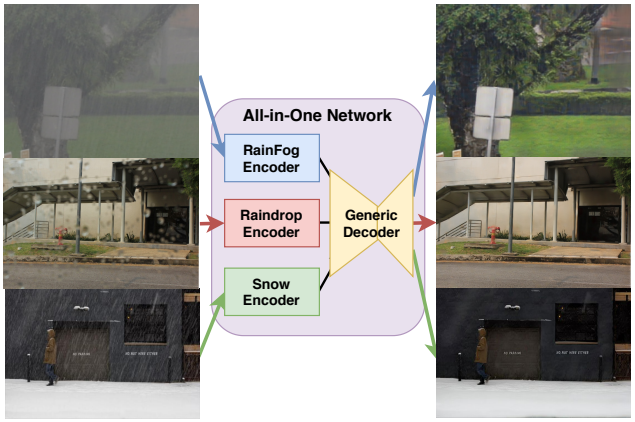


Figure 1: High-level view of our network, with different types of bad weather images as input and the respective clean images as output. The proposed method is able to process multiple types of bad weather images using the same set of weights/parameters.

In this paper, we develop a single network-based method to deal with many types of bad weather phenomena including rain, fog, snow and adherent raindrop. It is worth noting that a few recent studies attempt to recover multiple degradation problems [39, 6]. However, none of them can deal with multiple degradations with solely one set of pretrained weights. To achieve our goal, we need to consider a few factors related to our problem.

First, different bad weathers are formed based on different physical principles, which means the degraded images do not share the same characteristics. In order to yield the

†This work is supported by MOE2019-T2-1-130.

optimal performance, we need to design the network according to the underlying physics principles.

Second, bad weather image restoration can be considered as a many-to-one feature mapping problem, i.e., image features from different bad weather domains (rain, fog, raindrop, snow) are transformed to clean image features by a set of network parameters (multiple encoders), after which the clean features are transformed to the clean natural images (one decoder). Hence, it is critical to find a proper way to process features from multiple domains and subject them to further appropriate operations. This motivates us to design an architectural-search approach that automatically finds an optimal network architecture for the aforementioned task. The basic building blocks for our network-search module are made up of a series of fundamental operations that can convert degraded image features to clean features based on the physics characteristics of bad weather image degradation.

Third, most existing discriminators in GAN-based approaches are trained to judge whether the restored images are real or not. However, it does not provide error signals for the generative network to differentiate the images into different degradation types. The encoders may not be able to update their learnable parameters based on its own assessment of degradation type independently. To solve this problem, we propose a multi-class auxiliary discriminator that can classify the image degradation type and judge the correctness of the restored image simultaneously. In addition, unlike other existing GAN-based methods, our network has multiple feature encoders, each of which corresponds to a particular degradation type. When we backpropagate the discriminative loss, the network propagates only the loss to the respective encoder based on the classified results. Thus, only the corresponding encoder will update the parameters based on the adversarial loss; and, all the other encoders will not be affected.

We summarise the contributions of our method as follow:

1. We propose an all-in-one bad weather removal method that can deal with multiple bad weather conditions (rain streaks, rain veiling effect, snow, and adherent raindrop) in one network.

2. We propose a neural architecture search technique to find the best architecture for processing the features using different weather encoders. A series of fundamental operations that result in features invariant to bad weather are introduced. These fundamental operations form the basic building blocks for the search.

3. We propose a novel end-to-end learning scheme that can handle multiple bad weather image restoration tasks. The key idea is to let the errors of the discriminative loss backpropagate into a specific encoder, in accordance to the type of the bad weather input.

2. Related Works

Deep learning based solutions have achieved promising performance in various image processing problems such as denoising [58], image completion [16], super-resolution [21], deblurring [45], style transfer [9], etc. This is also true for bad weather restoration or image enhancement, such as dehazing [47, 3, 1, 57, 7, 13, 26, 43], removal of raindrop and dirt [20, 18, 59, 8, 53, 30, 50, 62, 36, 4], of moderate rain [44, 29], and of heavy rain [25, 54]. These recent works have all shown the superiority of deep neural network models to conventional methods.

Rain Removal Kang et al.'s [20] is the first work to introduce single image deraining method that decomposes an input image into its low frequency and high-frequency components using bilateral filter. Recent state-of-the-art rain removal strategies are dominated by deep neural networks. Fu et al.'s [8] develop a deep CNN to extract discriminative features from the high frequency component of the rain image. Yang et al. [54] design a multi-task deep learning architecture that learns the location and intensity of rain streaks simultaneously. Li et al.'s [25] propose a network that addresses the rain streaks and rain veiling effects prevalent in heavy rain scenes. This method not only proposes a residue decomposition step, but also elegantly integrates the physics-based rain model and adversarial learning to achieve state-of-the-art performance. It jointly learns the physics parameters of heavy rain, including streak intensity, transmission, atmospheric light and utilizes generative adversarial network to bridge the domain gaps between the proposed rain model and real rain.

Raindrop Removal There are a number of methods that detect and remove raindrops from single image based on traditional hand-crafted features [52, 55]. Eigen et al.'s [5] train a CNN with pairs of raindrop-degraded images and the corresponding raindrop-free images. Its network is a fairly shallow model that only has 3 convolutional layers. While the method works, particularly for relatively sparse and small droplets as well as dirt, the result tends to be blurry. Qian et al. [41] use attention maps in a GAN network that successfully removes raindrop from single image. However, the main drawback of this approach is the attention maps that are inherently difficult to obtain. The automatically computed attention map ground truth often results in poor quality. Quan et al.'s [42] further explore the generation of attention maps based on the mathematical description of the shape of raindrops. It combines the raindrop attention maps and detected raindrop edges to obtain state-of-the-art performance of single image raindrop removal.

Snow Removal [2, 46] use HOG techniques to capture characteristics of snow flakes for snow removal from single images. Xu et al. [51] utilize color assumptions to model the falling snow particles. In contrast to these hand-crafted features that capture partial characteristic of snow

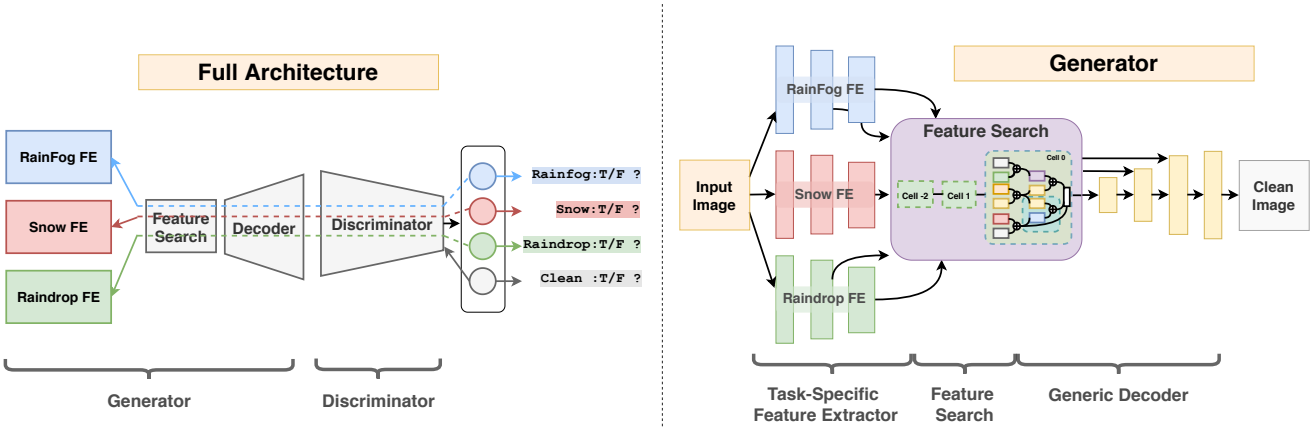


Figure 2: **Left:** Full architecture of the proposed network. The dotted lines indicate the back-propagation paths of the adversarial loss from the discriminator. The discriminator classifies the degradation type and also determines whether the input image is real or fake. The classification error of a particular degradation type is only used to update the corresponding encoder assigned to this type in the generative network. The loss from one degradation type is only propagated to update the corresponding feature encoder. **Right:** The detailed structure of the generator of the proposed network. In our experiment, we set the number of cells to 3 (*cell*-2, *cell*-1 and *cell* 0). **FE** stands for Feature Extractor.

flakes and streaks, Li et al.'s [24] encode snow flakes or rain streaks using an online multi-scale convolutional sparse coding model.

Neural Architecture Search (NAS) Neural Architecture Search aims at automatically designing neural network architectures to achieve optimal performance, while minimizing human hours and efforts. Early works like [63, 19, 11] directly construct the entire network and train it automatically with supervision from designing a reinforcement learning controller RNN. Many recent papers [32, 40] point out that searching the repeatable cell structure and fixing the network level structure are more effective and efficient. The PNAS method [32] proposed a progressive search that significantly reduces the computation cost. Our work is closely related to [34, 31] that further relax the network searching task into an end-to-end optimization problem.

3. Proposed Method

3.1. Problem Formulation

Different weather phenomena degrade images according to different physics principles. For example, a heavy rain image (where rain veiling effect, visually similar to fog/mist, is prevalent) is modelled as [25]:

$$\mathbf{I}(x) = \mathbf{t}(x)(\mathbf{J}(x) + \sum_i \mathbf{R}_i(x)) + (1 - \mathbf{t}(x))A, \quad (1)$$

where $\mathbf{I}(x)$ is the rain image at location x, $\mathbf{t}$ is the transmission map and A is the global atmospheric light of the scene. $\mathbf{R}_i$ represents the rain streaks at the i-th layer along the line of sight. An adherent raindrop image is modelled as [41] :

$$\mathbf{I}(x) = (1 - \mathbf{M}(x))\mathbf{J}(x) + \mathbf{K}(x), \quad (2)$$

where $\mathbf{I}$ is the colored raindrop image and $\mathbf{M}$ is the binary mask. $\mathbf{J}$ is the background image and $\mathbf{K}$ is the imagery brought about by the adherent raindrops, representing the blurred imagery formed the light reflected by the environment. Lastly, a snow image can be modelled as [37]:

$$\mathbf{I}(x) = \mathbf{z}\mathbf{S}(x) + \mathbf{J}(x)(1 - \mathbf{z}), \quad (3)$$

where $\mathbf{S}$ represents the snow flakes and $\mathbf{z}$ is a binary mask indicating the location of snow.

From the formulations of these different bad weather images, it is evident that these problems do not share the same intrinsic characteristics, which explains why a dedicated algorithm designed for one task does not work on the other tasks. To address this problem, we model the bad weather tasks with the following generic function:

$$\mathbf{J}(x) = \mathcal{F}(\mathbf{I}(x)), \quad (4)$$

where $\mathcal{F}$ represents an auto-encoder that maps degraded images to clean background images, and should embody the mentioned formulations such as Eq. (1)(2)(3). To realize this, we consider a network with multiple encoders:

$$\mathbf{J}(x) = \mathcal{G} \odot \mathcal{E}_\rho(\mathbf{I}_\rho(x)), \quad (5)$$

where $\mathcal{E}_\rho$ represents the encoder that takes in a degraded image $\mathbf{I}_\rho$ with respect to a degradation type ρ. $\mathcal{G}$ is the generic decoder that restores the input to a clean background image $\mathbf{J}$.

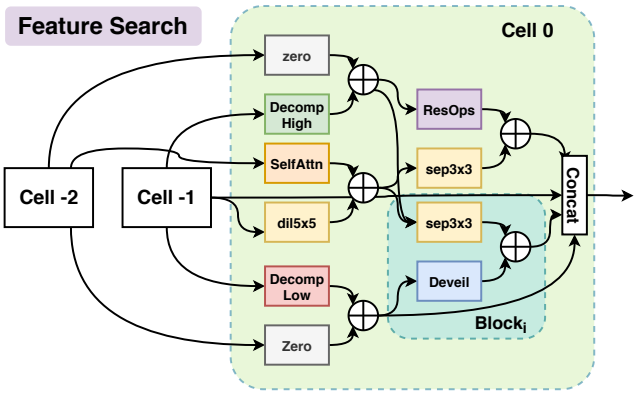


Figure 3: The detailed illustration of the feature-search block in our proposed network. This diagram shows a possible architecture layout after training. In the figure, *dil5x5* refers to dilated convolution with kernel 5×5. *sep3x3* refers to separable convolution with kernel 3×3.

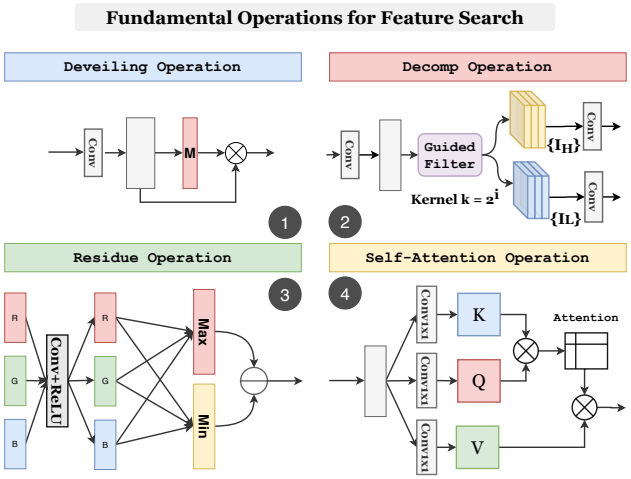


Figure 4: An detailed illustration of the 4 fundamental operations in the search component designed for bad weather restoration tasks.

3.2. Architectural Search Methodology

To effectively process features coming in from multiple encoders, we utilize an NAS technique at the ends of the multiple encoders. In this search module (details shown in Fig. 3), we follow the configuration of most of the recent NAS methods, which define a *cell* to be the smallest basic module that can be repeated multiple times to construct the entire network architecture. Therefore, the network search space comprises of both a network level search, which refers to finding the structure of connections between cells, and a cell level search, which explores the structure inside a cell.

3.2.1 Network Cell Architecture

Due to limited computing resources, we adopt the fundamental rule of [31] in designing the basic cell structure: a *cell* is a directed acyclic graph consisting of B blocks. Each *block* i in the l-th cell C^l is represented as a 5-tuple (I_1, I_2, O_1, O_2, H) two-to-one mapping structure, where $I_1, I_2 \in I_i^l$ are the selections of the input tensors; and $O_1, O_2 \in \mathcal{O}$ are the selections of the layer types applied to the corresponding input tensors. $H \in \mathcal{H}$ is the method used to combine the outputs of two layers O_1, O_2. The set of possible layer types $\mathcal{O}$ consists of the following ten operators:

- 3x3 separable conv
- 5x5 separable conv
- 3x3 dilated conv
- 5x5 dilated conv
- no/skip connection
- **Deveiling Ops**
- **Residue Ops**
- **Self-Attention Ops**
- **Decomposition Ops**

Beside the common convolutional operations such as dilated convolution, depthwise-separable convolution, we introduce new fundamental operations to deal with the bad weather image degradation according to the physics laws embedded in the formation of each degradation type, namely *decomposition operation, residue operation, self-attention operation, and deveiling operation* (shown in Fig. 4). In the following paragraphs, we describe the design and function of the new operations in details.

Deveiling Operation Most of the existing methods solve the problem according to the following model of fog/haze formation:

$$\mathbf{I}(x) = \mathbf{t}\mathbf{J}(x) + (\mathbf{1} - \mathbf{t})A. \tag{6}$$

The variables in this equation follow the same meaning in Eq. (1). Inspired by [23, 28], it is possible to learn a latent variable $\mathbf{M}$ that transforms the veiling images into clean background images:

$$\mathbf{J}(x) = \mathbf{M}(x) \odot \mathbf{I}(x), \tag{7}$$

where $\mathbf{M}(x) = (\mathbf{I}(x) + \mathbf{t}A - A)/\mathbf{t}\mathbf{I}(x)$, and is a learnable latent variable dependent on the input image $\mathbf{I}$. Thus, we apply one layer of $conv1 \times 1$ on the extracted image feature to estimate the latent variable $\mathbf{M}$, and multiply $\mathbf{M}$ with the extracted feature as shown in Fig. 4.1.

Decomposition Operation Image decomposition has been widely used in rain streaks removal and snow removal [20, 8, 25]. We consider the decomposition in a feature space as a fundamental operation that is effective for snow and rain removal. As shown in Fig. 4.2, we apply deep image-guided filters with a kernel family ranging in 2^k,

where k = 1,2,...,6. We use a $conv1 \times 1$ layer at the end of both the low- and high-frequency components to extract appropriate features for next layer.

Residue Operation Inspired by [28], we have implemented the residue channel operation (shown in Fig. 4.3 in the feature space. The residue channel [27] has been shown to be effective to remove rain from a single image.

Self-Attention Operation A few methods [41, 42] have demonstrated the advantages of using an attention map for removing adherent raindrops. In these methods, the raindrop attention maps are explicitly learned from the ground truths of raindrop masks. However, obtaining these raindrop masks is expensive [12] and any lack of quality may affect the performance. To overcome this problem, a self-attention mechanism [56, 48] has been applied to many image reconstruction methods [60, 33]. We consider this self-attention module as a fundamental block of operation that can be exploited by the NAS (shown in Fig. 4.4).

3.2.2 Architecture Search Space

Following the continuous relaxation described in [35], in each $block_i$, the output tensor T_i^l is connected to all the input tensor I_i^l through searched operation $O_{j\to i}$:

$$T_i^l = \sum_{H_j^l \in I_i^l} O_{j\to i}(T_j^l). \tag{8}$$

We approximate the step of searching best $O_{j\to i}$ with continuous relaxation, yielding $\bar{O}_{j\to i}$:

$$\bar{O}_{j\to i} = \sum_{O\in\mathcal{O}} \alpha_{j\to i} O_{j\to i}(T_j^l), \tag{9}$$

where $\sum \alpha_{j\to i} = 1$. In practice, this is implemented as a softmax operation. Therefore, the cell level architecture can be summarized as :

$$T^l = Cell(T^{l-1}, T^{l-2}; \alpha). \tag{10}$$

To maximize the potential of these fundamental knowledge competencies, we allow heterarchical connections between cells and blocks.

3.3. Categorical Adversarial Training

In the standard generative adversarial network (GAN) configuration, the generator G takes a random noise vector z and produces an image X_{fake}. The discriminator D takes a ground truth image and the output image of the generator to predict a probability distribution $P(S|X)$ over possible images sources [10]. During training process, the discriminator is trained to maximize the log-likelihood of the correct source:

$$\mathcal{L}_s = \mathbb{E}[\log P(S = real)|X_{real})] \tag{11}$$
$$+ \mathbb{E}[\log P(S = fake|X_{fake})]. \tag{12}$$

In our multi-class discriminator network, however, the discriminator does not only learn to determine the correctness of the restored image, but also strives to classify the type of the degradation from the restored image:

$$\mathcal{L}_c = \mathbb{E}[\log P(C = c_i)|X_{real})] + \mathbb{E}[\log P(C = c_i|X_{fake})]. \tag{13}$$

Hence, the discriminator D is trained to minimize $\mathcal{L}_s + \mathcal{L}_c$, while the generator is trained to minimize $\mathcal{L}_c - \mathcal{L}_s$. This approach havs been proven to be effective in tackling mode collapse in a standard GAN [14]. To summarize, the loss function for the discriminator is:

$$\mathcal{L}_{dis} = \mathcal{L}_c + \mathcal{L}_s. \tag{14}$$

For the generator, we also apply the MSE loss to compute the difference between the predicted clean image $\mathbf{J}_{pred}$ and the ground truth clean image $\mathbf{J}_{gt}$:

$$\mathcal{L}_{gen} = \mathcal{L}_c + \mathcal{L}_{mse}(\mathbf{J}_{pred}, \mathbf{J}_{gt}) - \mathcal{L}_s. \tag{15}$$

Updating Relevant Encoders As mentioned, different bad weathers are formed based on different physical principles. The classification error of one degradation type $\mathcal{L}_{c_i}$ may not be effective to update the encoders of other degradation types j, where $j \neq i$. In our approach, the classification error of degradation type i is only used to update encoder $\mathcal{E}_i$, as shown in Fig. 2. By backpropagating the adversarial loss specifically to the appropriate encoder, we strengthen the ability of the multi-encoder generator to map images from different domains to a common feature space.

4. Implementation

4.1. Datasets

We train our network on different bad weather datasets, including "Outdoor-Rain" [25], "Snow100K" [38], and "Raindrop" [41]. "Outdoor-Rain" contains 9,000 training samples and 1,500 validation samples. "Snow100K" contains 100k synthetic snow images and the corresponding snow-free ground truth images. "Raindrop" comprises 1,119 pairs of real adherent-raindrop images and the corresponding ground truth background images. Since the number of images in each datasets are not equal, we sample 9000 images from "Snow100K" at each training epoch. For the small "Raindrop" dataset, we over-sample them by performing data augmentation, such as rotation, affine transformation, noise and random cropping. As a result, in each

Table 1: A comparison of our algorithm with the baseline methods performed on *Test 1* [25] dataset.

Method		Test 1	
Metric		PSNR	SSIM
DetailsNet [8] + Dehaze	DHF	13.36	0.583
	DRF	15.68	0.640
RESCAN [29] + Dehaze	DHF	14.72	0.587
	DRF	15.91	0.615
Pix2Pix [17]		19.09	0.710
CycleGAN [61]		17.62	0.656
HRGAN [25]		21.56	0.855
Ours		**24.71**	**0.898**

Table 2: A comparison of our algorithm with the baseline methods performed on *Raindrop* [41] dataset.

Metric	Pix2Pix [17]	AttentGAN [41]	Quan et al. [42]	Ours
PSNR	28.02	31.57	**31.44**	31.12
SSIM	0.8547	0.9023	0.9263	**0.9268**

Table 3: A comparison of our algorithm with the baseline methods performed on *Snow100K-L* [37] test dataset.

Metric	DetailsNet [8]	DesnowNet [38]	Ours
PSNR	19.18	27.17	**28.33**
SSIM	0.7495	**0.8983**	0.8820

epoch, the number of samples from all the datasets is uniform at 9,000.

4.2. Training Details

Our network is trained on all of the bad weather datasets in an end-to-end manner. To reduce the training expenses under limited resources, we adopt the first-order approximation in [35] and split the training data into two disjoint sets *Set1* and *Set2*. We simultaneously optimize the parameters of the image restoration network on *Set1* and the architecture-search parameters on *Set2*. The learning rate is set to 0.002 initially and is divided by 2 after every 5 epochs until the 40^{th} epoch. We use Adam optimizer [22] with a weight decay 10^{-4} to optimize the network parameters.

5. Experiments

We evaluate our method using both synthetic and real bad weather images, including rain, snow and adherent raindrop. The test dataset for rain (with fog) is the test set from HRGAN [25]. The baseline methods for deraining comprise of the state-of-the-art heavy rain removal HRGAN [25], RESCAN [29], and DetailsNet [8]. The test dataset for snow is from the Snow100K-L test set adopted from [37]. Since there are only a few works related to snow recently, we compare DeSnowNet [37] and the baseline methods from that paper. Lastly, the test dataset for adherent raindrop is from Qian et al. [41]. We compare our method with the most recent raindrop removal methods [42] and Qian et al. [41].

5.1. Qualitative Results

Rain and Fog We show the results produced by the various methods on synthetic rain images in Fig. 5. One can observe that our network, while is trained for multiple bad weather types, achieve competitive performance compared with dedicated state-of-the-art deraining methods. We provide more results in the supplementary material.

Snow We show the results produced by the various method on synthetic snow images from the Snow100K dataset [38] in Fig. 9. [1]

Raindrop We list the results of our method compared with recent raindrop removal methods in Fig. 7. Although our method does not produce the best result in terms of the PSNR, we still achieve a competitive performance without incorporating extra information such as that obtained from the edge attention mechanism in [42].

5.2. Quantitative Results

Table 1 and 2 demonstrate the quantitative results of our proposed method compared with dedicated state-of-the-art methods and generic image restoration methods. The quantitative results are evaluated based on two metrics: PSNR [15] and SSIM [49]. For the raindrop removal task, our method does not yield the best result in terms of the PSNR, but achieve a competitive performance.

6. Ablation Study

To study the effectiveness of each of the components in our proposed network, we conduct an ablation study, which the results are shown in Table. 4. As can be seen, our network with the feature search performs better than simple concatenation. We also conduct an ablation study on the categorical adversarial training component. The quantitative results tested on the deraining task are shown in Table. 4.

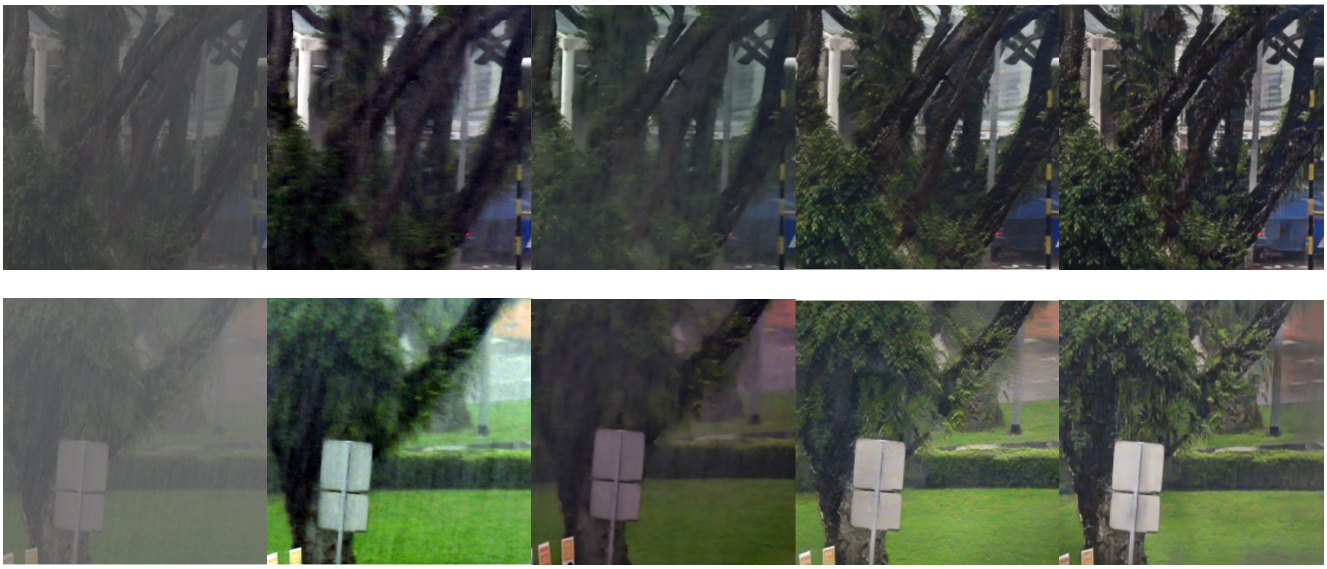
7. Conclusion

In this paper, we propose a novel all-in-one bad weather image enhancement solution that can handle multiple types of bad weather degradations using only one single network. The competitive performance of our network stems from

[1] Since the original code of [38] is not available, the qualitative results are based on our implementation. The quantitative results in Table 3 are directly obtained from the paper.

(a) input (b) RESCAN [29] (c) HRGAN [25] (d) Ours (e) GT

Figure 5: Synthetic rain and fog removal results of our method compared with state-of-the-art dedicated rain and fog removal methods.



(a) input (b) DetailsNet [8] (c) RESCAN [29] (d) HRGAN [25] (e) Ours

Figure 6: Realistic rain and fog removal results of our method compared with state-of-the-art dedicated rain and fog removal methods.

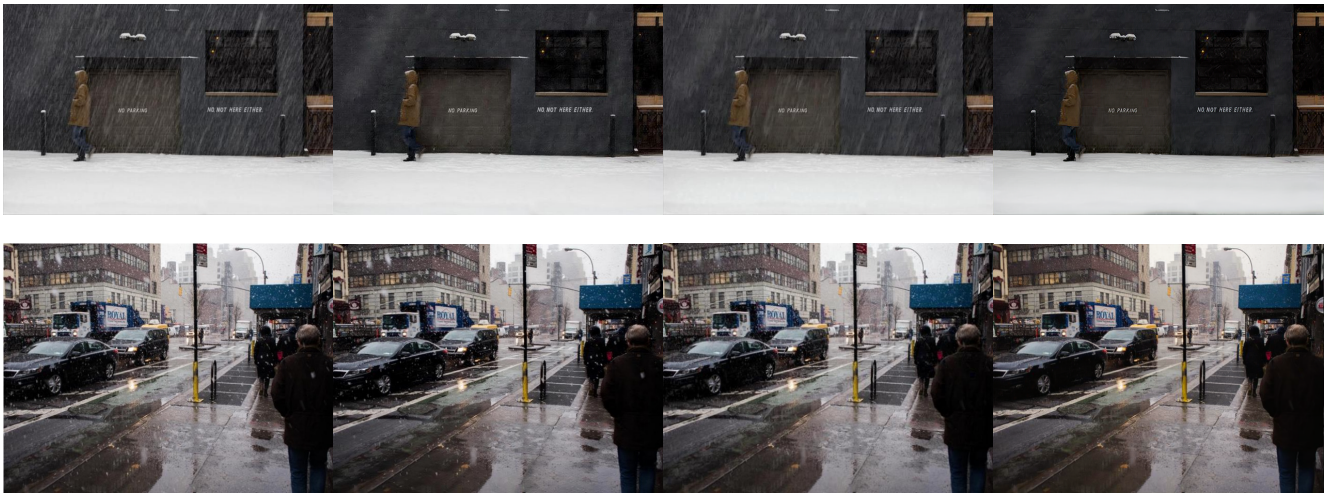
Table 4: Ablation study on our search component and categorical adversarial learning component in the proposed network. The evaluation is conducted on the test dataset *Test 1* [25] for the rain and fog removal task.

Method	Rainfog dataset [25]	
Metric	PSNR	SSIM
No Feature Search	20.82	0.827
No Categorical Discriminator	21.58	0.86
Full Architecture	**24.71**	**0.898**

our two main contributions. First, to find the most effective way to process features from different bad weather domains, we propose an architectural search equipped with, among others, four fundamental operations designed for bad weather, namely *deveiling, residue, self-attention and decomposition*. Second, we design a multi-class discriminator that classifies image degradation types and assesses image correctness simultaneously. The proposed new training scheme updates the encoders in the generator based on the classification results of the discriminator. Finally, comprehensive experiments demonstrate the effectiveness of our method compared with the dedicated state-of-the-art algorithms on rain, snow and raindrop removal tasks.

(a) input (b) AttentGAN[41] (c) Quan et al. [42] (d) Ours (e) Ground Truth

Figure 7: Raindrop removal results of our method compared with state-of-the-art dedicated raindrop removal methods.



(a) input (b) DetailsNet [8] (c) DeSnowNet [38] (d) Ours

Figure 8: Snow removal results of our method compared with state-of-the-art dedicated snow removal methods.

(a) Input (b) No Categorical Discriminator (c) No Feature Search (d) Ours

Figure 9: Comparison between different networks in ablation study (zoom to view details).

References

[1] D. Berman, T. Treibitz, and S. Avidan. Non-local image dehazing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[2] Jérémie Bossu, Nicolas Hautière, and Jean-Philippe Tarel. Rain or snow detection in image sequences through use of a histogram of orientation of streaks. *International Jour-*

nal of Computer Vision, 93(3):348–367, Jul 2011.

[3] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *Trans. Img. Proc.*, 25(11):5187–5198, Nov. 2016.

[4] Jie Chen, Cheen-Hau Tan, Junhui Hou, Lap-Pui Chau, and He Li. Robust video content alignment and compensation for rain removal in a cnn framework. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[5] D. Eigen, D. Krishnan, and R. Fergus. Restoring an image taken through a window covered with dirt or rain. In *2013 IEEE International Conference on Computer Vision*, pages 633–640, Dec 2013.

[6] Qingnan Fan, Dongdong Chen, Lu Yuan, Gang Hua, Nenghai Yu, and Baoquan Chen. Decouple learning for parameterized image operators. In *The European Conference on Computer Vision (ECCV)*, September 2018.

[7] Raanan Fattal. Dehazing using color-lines. *ACM Trans. Graph.*, 34(1), Dec. 2015.

[8] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[9] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. A neural algorithm of artistic style. *CoRR*, abs/1508.06576, 2015.

[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc., 2014.

[11] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber. Lstm: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2222–2232, Oct 2017.

[12] Zhixiang Hao, Shaodi You, Yu Li, Kunming Li, and Feng Lu. Learning from synthetic photorealistic raindrop for single image raindrop removal. In *The IEEE International Conference on Computer Vision (ICCV) Workshops*, Oct 2019.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 770–778, 2016.

[14] Quan Hoang, Tu Dinh Nguyen, Trung Le, and Dinh Phung. MGAN: Training generative adversarial nets with multiple generators. In *International Conference on Learning Representations*, 2018.

[15] Q. Huynh-Thu and M. Ghanbari. Scope of validity of psnr in image/video quality assessment. *Electronics Letters*, 44(13):800–801, June 2008.

[16] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. Globally and Locally Consistent Image Completion. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2017)*, 36(4):107:1–107:14, 2017.

[17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *arxiv*, 2016.

[18] Tai-Xiang Jiang, Ting-Zhu Huang, Xi-Le Zhao, Liang-Jian Deng, and Yao Wang. A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[19] Rafal Jozefowicz, Wojciech Zaremba, and Ilya Sutskever. An empirical exploration of recurrent network architectures. In *Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, pages 2342–2350. JMLR.org, 2015.

[20] L. W. Kang, C. W. Lin, and Y. H. Fu. Automatic single-image-based rain streaks removal via image decomposition. *IEEE Transactions on Image Processing*, 21(4):1742–1755, April 2012.

[21] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. *CoRR*, abs/1511.04587, 2015.

[22] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

[23] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[24] Minghan Li, Xiangyong Cao, Qian Zhao, Lei Zhang, Chenqiang Gao, and Deyu Meng. Video rain/snow removal by transformed online multiscale convolutional sparse coding. *CoRR*, abs/1909.06148, 2019.

[25] Ruoteng Li, Loong-Fah Cheong, and Robby T. Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[26] R. Li, J. Pan, Z. Li, and J. Tang. Single image dehazing via conditional generative adversarial network. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8202–8211, June 2018.

[27] Ruoteng Li, Robby T. Tan, and Loong-Fah Cheong. Robust optical flow in rainy scenes. In *The European Conference on Computer Vision (ECCV)*, September 2018.

[28] Ruoteng Li, Robby T. Tan, Loong-Fah Cheong, Angelica I. Aviles-Rivero, Qingnan Fan, and Carola-Bibiane Schonlieb. Rainflow: Optical flow under rain streaks and rain veiling effect. In *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[29] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *The European Conference on Computer Vision (ECCV)*, September 2018.

[30] Yu Li, Robby T. Tan, Xiaojie Guo, Jiangbo Lu, and Michael S. Brown. Rain streak removal using layer priors. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[31] Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan L. Yuille, and Li Fei-Fei. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation. *CoRR*, abs/1901.02985, 2019.

[32] Chenxi Liu, Barret Zoph, Jonathon Shlens, Wei Hua, Li-Jia Li, Li Fei-Fei, Alan L. Yuille, Jonathan Huang, and Kevin Murphy. Progressive neural architecture search. *CoRR*, abs/1712.00559, 2017.

[33] Ding Liu, Bihan Wen, Yuchen Fan, Chen Change Loy, and Thomas S. Huang. Non-local recurrent network for image restoration. *CoRR*, abs/1806.02919, 2018.

[34] Hanxiao Liu, Karen Simonyan, and Yiming Yang. DARTS: differentiable architecture search. *CoRR*, abs/1806.09055, 2018.

[35] Hanxiao Liu, Karen Simonyan, and Yiming Yang. DARTS: Differentiable architecture search. In *International Conference on Learning Representations*, 2019.

[36] Jiaying Liu, Wenhan Yang, Shuai Yang, and Zongming Guo. Erase or fill? deep joint recurrent rain removal and reconstruction in videos. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[37] Yun-Fu Liu, Da-Wei Jaw, Shih-Chia Huang, and Jenq-Neng Hwang. Desnownet: Context-aware deep network for snow removal. *CoRR*, abs/1708.04512, 2017.

[38] Y. Liu, D. Jaw, S. Huang, and J. Hwang. Desnownet: Context-aware deep network for snow removal. *IEEE Transactions on Image Processing*, 27(6):3064–3073, June 2018.

[39] Jinshan Pan, Sifei Liu, Deqing Sun, Jiawei Zhang, Yang Liu, Jimmy Ren, Zechao Li, Jinhui Tang, Huchuan Lu, Yu-Wing Tai, and Ming-Hsuan Yang. Learning dual convolutional neural networks for low-level vision. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[40] Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc V. Le, and Jeff Dean. Efficient neural architecture search via parameter sharing. *CoRR*, abs/1802.03268, 2018.

[41] Rui Qian, Robby T. Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[42] Yuhui Quan, Shijie Deng, Yixin Chen, and Hui Ji. Deep learning for seeing through window with raindrops. In *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[43] Wenqi Ren, Si Liu, Hua Zhang, Jin shan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, 2016.

[44] Weihong Ren, Jiandong Tian, Zhi Han, Antoni Chan, and Yandong Tang. Video desnowing and deraining based on matrix decomposition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[45] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf. Learning to deblur. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(7):1439–1451, July 2016.

[46] Soo-Chang Pei, Yu-Tai Tsai, and Chen-Yu Lee. Removing rain and snow in a single image using saturation and visibility features. In *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, pages 1–6, July 2014.

[47] R. T. Tan. Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2008.

[48] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 5998–6008. Curran Associates, Inc., 2017.

[49] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.

[50] Wei Wei, Lixuan Yi, Qi Xie, Qian Zhao, Deyu Meng, and Zongben Xu. Should we encode rain streaks in video as deterministic or stochastic? In *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[51] Jing Xu, Wei Zhao, Peng Liu, and Xianglong Tang. An improved guidance image based method to remove rain and snow in a single image. *Computer and Information Science*, 5, 04 2012.

[52] A. Yamashita, Y. Tanaka, and T. Kaneko. Removal of adherent waterdrops from images acquired with stereo camera. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 400–405, 2005.

[53] Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Joint rain detection and removal via iterative region dependent multi-task learning. *CoRR*, abs/1609.07769, 2016.

[54] W. Yang, R. T. Tan, J. Feng, J. Liu, S. Yan, and Z. Guo. Joint rain detection and removal from a single image with contextualized deep networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2019.

[55] S. You, R. T. Tan, R. Kawakami, Y. Mukaigawa, and K. Ikeuchi. Adherent raindrop modeling, detectionand removal in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(9):1721–1733, 2016.

[56] Han Zhang, Ian J. Goodfellow, Dimitris N. Metaxas, and Augustus Odena. Self-attention generative adversarial networks. *ArXiv*, abs/1805.08318, 2018.

[57] He Zhang and Vishal M. Patel. Density-aware single image de-raining using a multi-stream dense network. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[58] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, July 2017.

[59] X. Zhang, H. Li, Y. Qi, W. K. Leow, and T. K. Ng. Rain removal in video by combining temporal and chromatic properties. In *2006 IEEE International Conference on Multimedia and Expo*, pages 461–464, July 2006.

[60] Yulun Zhang, Kunpeng Li, Kai Li, Bineng Zhong, and Yun Fu. Residual non-local attention networks for image restoration. In *International Conference on Learning Representations*, 2019.

[61] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

[62] Lei Zhu, Chi-Wing Fu, Dani Lischinski, and Pheng-Ann Heng. Joint bi-layer optimization for single-image rain streak removal. In *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[63] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V. Le. Learning transferable architectures for scalable image recognition. *CoRR*, abs/1707.07012, 2017.