SDRNET: SALIENCY-GUIDED DYNAMIC RESTORATION NETWORK FOR RAIN AND HAZE REMOVAL IN NIGHTTIME IMAGES

Wanning Zhu, Lin Tan, and Libao Zhang*

School of Artificial Intelligence, Beijing Normal University, Beijing 100875, China

1. Introduction

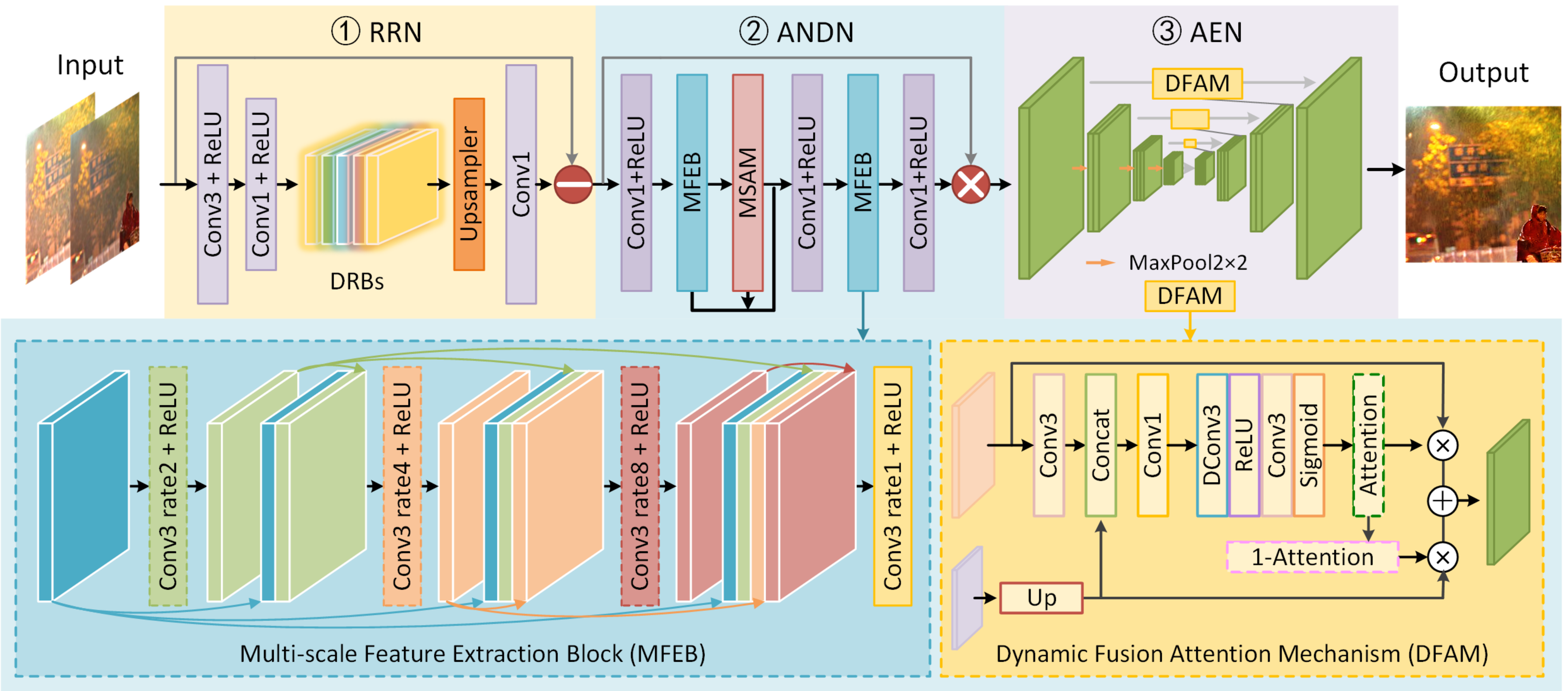

Fig. 1. The structure of the proposed method.

- We design a novel saliency-guided detail enhancement preprocessing (SDEP) method to enhance the illumination features of light source regions and texture information of non-light ones in nighttime rain-haze images.
- The proposed all-in-one nighttime dehazing network (ANDN) integrates spatially variable ambient light and transmission into one parameter to effectively remove haze in nighttime scenes.
- We propose a dynamic fusion attention module (DFAM) with deformable convolution layer to enhance the nighttime lowlight background image with unevenly distributed light according to the input features.
- We synthesize a Nighttime Rain-Haze Image Dataset (NRHID), which includes a large number of nighttime images with different fog effects and rain distributions. To our best knowledge, we are the first to generate such nighttime rain-haze image dataset for training.

2. Methodology

2.1. Nighttime Rain-Haze Image Model

We comprehensively consider the rain image model and nighttime haze image model and define the nighttime rain-haze image model as follow:

$$I(x) = R(x) + J(x)t(x) + A(x)[1 - t(x)] \tag{1}$$

where $I(x)$ is to the nighttime rain-haze image, $R(x)$ is the rain layer and $J(x)$ is the clear image. $A(x)$ is the globally-changed ambient light at night and $t(x)$ is the transmission. In our model, we first propose rain removal network (RRN) to eliminate the influence of rain layer $R(x)$. Then we design ANDN to jointly estimate $A(x)$ and $t(x)$ and further get $J(x)$. Finally, attention-based enhancement network (AEN) is proposed to enhance the nighttime lowlight image $J(x)$ and obtain the final restoration result. Before all that, saliency detection is introduced to generate the detail map as additional supervision.

2.2. Saliency-guided Detail Enhancement Preprocessing

We first use homomorphic filtering to correct the illumination distribution of the input image. Then we introduce saliency detection to obtain the saliency map reflecting the light distribution and further enhance the image by weighting. Finally, we concatenate it with the original image as the network's input to learn the auxiliary detail information.

2.3. Rain Removal Network (RRN)

In order to remove rain stripes and preserve high-frequency background information, we use five dilated residual blocks (DRBs) with different dilated rates (Fig. 2) to greatly enlarge the reception. The DRBs is preceded and followed by the convolution operation to get the distribution of rain layer $R(x)$.

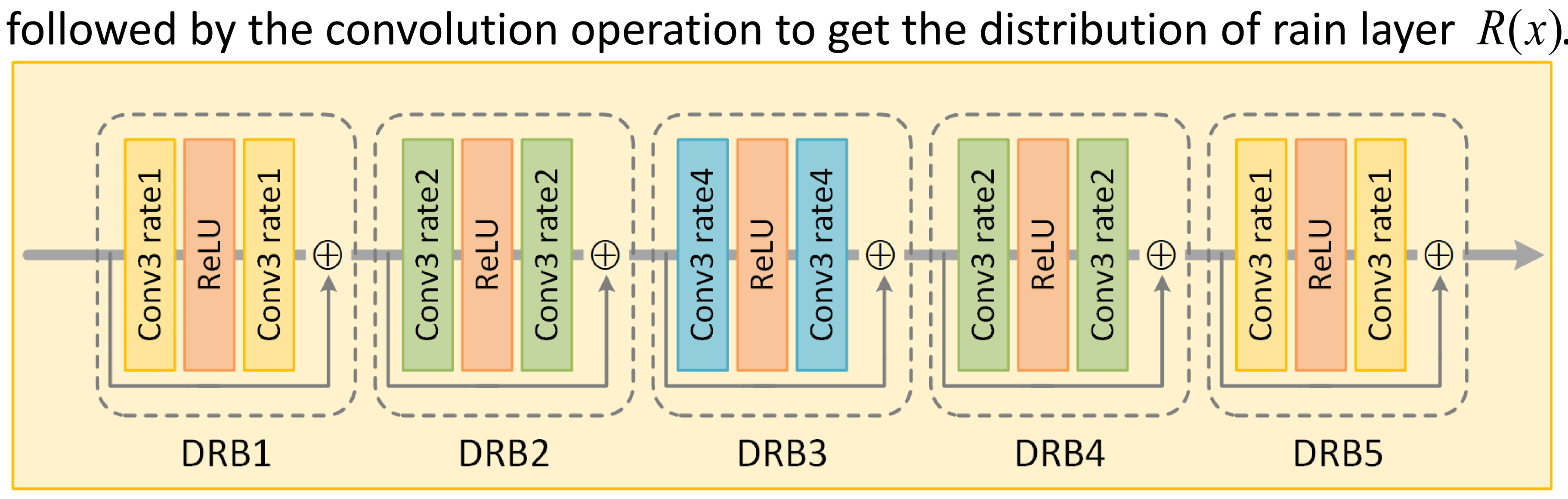

Fig. 2. The structure of dilated residual blocks (DRBs).

2.4. All-in-one Nighttime Dehazing Network (ANDN)

Due to the interference error between $t(x)$ and A, they can be unified as:

$$K(x) = \frac{(I(x) - A)/t(x) + (A - b)}{I(x) - 1} \tag{2}$$

where b is the constant bias with the default value 1.

The ambient light distribution at night is globally uneven, so A is no longer a constant and the atmospheric scattering model is not applicable to the hazy scene at night. From this aspect, we re-write the model shown in Eq. (2) and define the intermediate parameter $M(x)$ as:

$$J(x) = M(x)I(x)$$
$$M(x) = \frac{1}{t(x)} - \frac{A(x)}{t(x)I(x)} + \frac{A(x)}{I(x)} \tag{3}$$

We design ANDN to obtain $M(x)$. The network contains two modules: multi-scale feature extraction block (MFEB) and multi-scale spatial attention mechanism (MSAM). MFEB uses dilated convolution and cross layer connection to enrich receptive fields. MSAM is constructed to improve the feature representation abilities in different areas by spatial attention mapping.

2.5. Attention-based Enhancement Network (AEN)

We design a lightweight encoder-decoder network based on DFAM to enhance the low illumination image after rain and haze removal. The skip connection between the encoder and decoder is achieved by DFAM, in which the output of encoder and the input of decoder are first concatenated, and then pass through several convolutions and the sigmoid function to get the attention map. We weight the two input by it to get the final output of DFAM.

3. Experimental Results

Due to the absence of nighttime rain-haze image dataset for training, we synthesize NRHID dataset to train the proposed SDRNet and evaluate its performance on real-world nighttime rain-haze images. First, we generate hazy images according to the atmospheric scattering model. Then we synthesize a fuzzy noise image with motion blurs to imitate rain streaks. Finally, we expand the two-dimensional rain noise into a three-channel map, which is weighted proportionally with the hazy image to synthesize the rain-haze image.

Two rain removal methods and three nighttime dehazing methods are selected to form six combined restoration methods as comparison algorithms. The visual results on real-world scenes are shown in Fig. 3. The objective evaluations are shown in Table 1. The ablation study is designed to verify the effects of SDEP, AEN and PL. The experimental results are shown in Table 2.

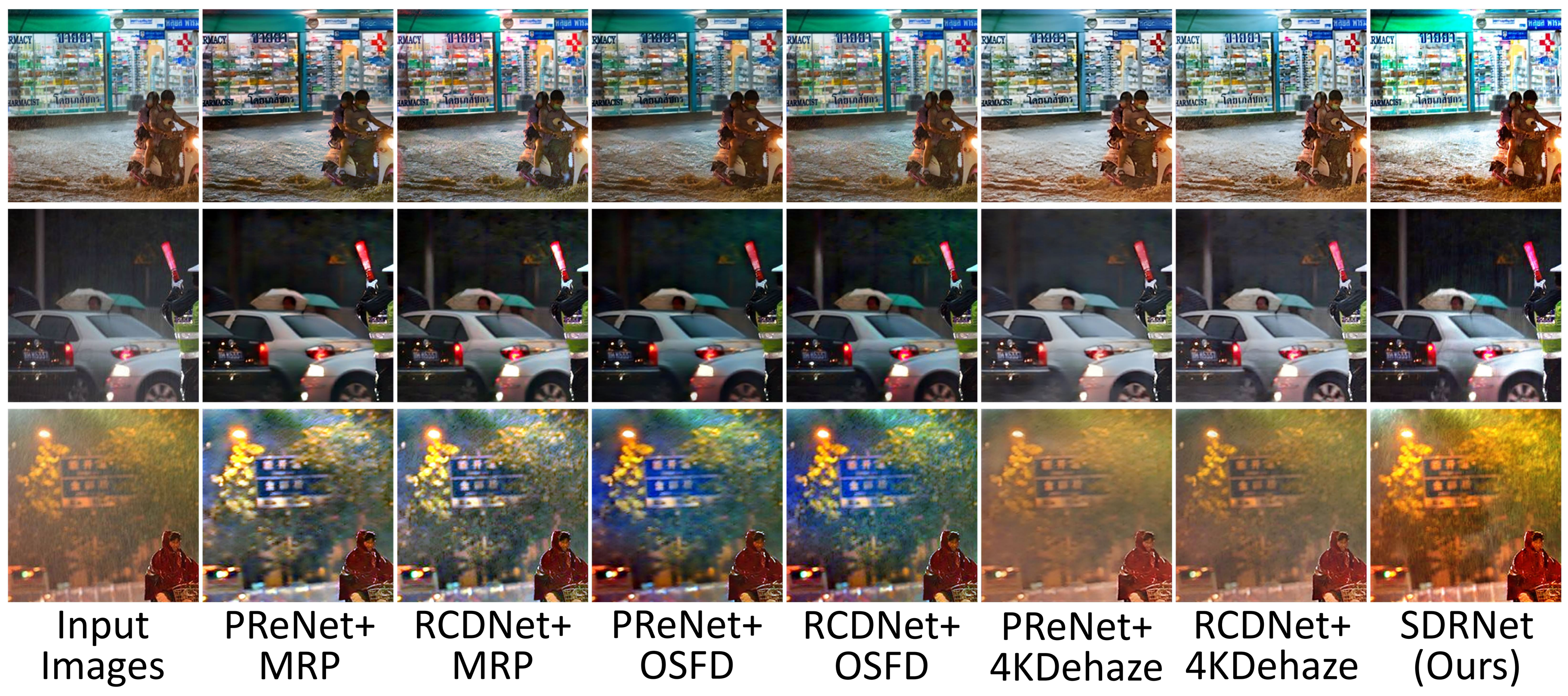

Fig. 3. Visual results of different restoration methods on real-world scenes.

Table 1. Quantitative assessment on real-world scenes.

Modules	σ	r	ENIQA	FADE
PReNet+MRP	0.007093	1.177832	0.134911	0.265723
RCDNet+MRP	0.008329	1.303544	0.117843	0.246797
PReNet+OSFD	0.003227	1.007588	0.139119	0.269112
RCDNet+OSFD	0.003281	1.127328	0.105297	0.245694
PReNet+4KDehaze	0.000604	1.318255	0.154282	0.394633
RCDNet+4KDehaze	0.000645	1.392977	0.099464	0.337738
SDRNet (Ours)	0.009304	1.393866	0.090527	0.210097

Table 2. PSNR and SSIM values of ablation study.

Modules	PSNR	SSIM
w/o LP	33.52	0.9100
w/o AEN	34.47	0.9381
w/o SDEP	34.73	0.9425
SDRNet (ours)	35.42	0.9428