DSDNet: Toward single image deraining with self-paced curricular dual stimulations

Yong Du [a], Junjie Deng [b,c], Yulong Zheng [a], Junyu Dong [a], Shengfeng He [c,*]

a Ocean University of China, Qingdao, China

b South China University of Technology, Guangzhou, China

c, * Singapore Management University, Singapore. Corresponding author: shengfenghe@smu.edu.sg

Published in Computer Vision and Image Understanding, 2023 April, 230, 103657. DOI: 10.1016/j.cviu.2023.103657

Abstract

A crucial challenge regarding the single image deraining task is to completely remove rain streaks while still preserving explicit image details. Due to the inherent overlapping between rain streaks and background scenes, the texture details could be inevitably lost when clearing rain away from the degraded image, making the two purposes contradictory. Existing deep learning based approaches endeavor to resolve the two issues successively in a cascaded framework or to treat them as independent tasks in a parallel structure. However, none of the models explores a proper interaction between rain distributions and hidden feature responses, which intuitively would provide more clues to facilitate the procedures of rain streak removal as well as detail restoration. In this paper, we investigate the impact of rain streak detection for single image deraining and propose a novel deep network with dual stimulations, namely, DSDNet. The proposed DSDNet utilizes a dual-stream pipeline to separately estimate rain streaks and a loss of details, and more importantly, an additional mask that indicates both location and intensity of rains is jointly predicted. In particular, the rain mask is involved in a tailored stimulation strategy that is deployed into each stream of the proposed model, serving as guidance for allowing the network to focus on rain removal and detail recovery in rain regions rather than non-rain areas. Moreover, we incorporate a self-paced semi-curriculum learning design to alleviate the learning ambiguity brought by the prediction of the rain mask and thus accelerate the training process. Extensive experiments demonstrate the proposed method outperforms the state-of-the-art methods on several benchmarks, including in both synthetic and real-world scenarios. The effectiveness of the proposed method is also validated via joint single image deraining, detection, and segmentation tasks.

Keywords: Single image deraining, Rain distributions, Stimulation strategy, Semi-curriculum learning

Rain causes degradation unavoidably when capturing images in outdoor scenes. Such a kind of deterioration would severely hamper the practicality of many image processing and computer vision algorithms, which usually take an assumption of a clear environment without considering the condition of rain. Therefore, rain removal technology has become a desired preprocessing step for various applications, such as object detection (He et al., 2017), video surveillance (Shehata et al., 2008), and visual tracking (Comaniciu et al., 2003). Most deraining methods aim to restore a clean image from its rainy observation with a rain streak layer, under a plain degradation model. Due to the two components (i.e., the clean background and the rain streaks) both unknown, it is an ill-posed problem that makes the deraining task extremely challenging.

Compared to video-based deraining problem (Tripathi and Mukhopadhyay, 2014, Kim et al., 2015, Jiang et al., 2017) which can exploit the intrinsic temporal information embedded in neighboring frames as additional constraints, the single image based deraining problem is naturally more underdetermined. Existing single image deraining methods can be roughly divided into two categories: model-based and deep learning based methods. Model-based methods (Kang et al., 2011, Chen and Hsu, 2013, Luo et al., 2015, Li et al., 2016) utilize optimization frameworks with hand-crafted priors for rain removal, such as low-rank representation (Chen and Hsu, 2013), sparse coding (Luo et al., 2015), and Gaussian mixture model (GMM) (Li et al., 2016). However, these methods generally exhibit a slow inference due to their high computational complexity. Also, it is hard to design proper priors to characterize the degraded images, especially in medium- and heavy-rain scenes, due to the messy statistics (e.g., directions and shapes) of rain streaks.

Deep learning based methods (Li et al., 2017, Yang et al., 2017, Li et al., 2018, Ren et al., 2019), on the other hand, have achieved conspicuous improvements via learning an implicit mapping from rainy

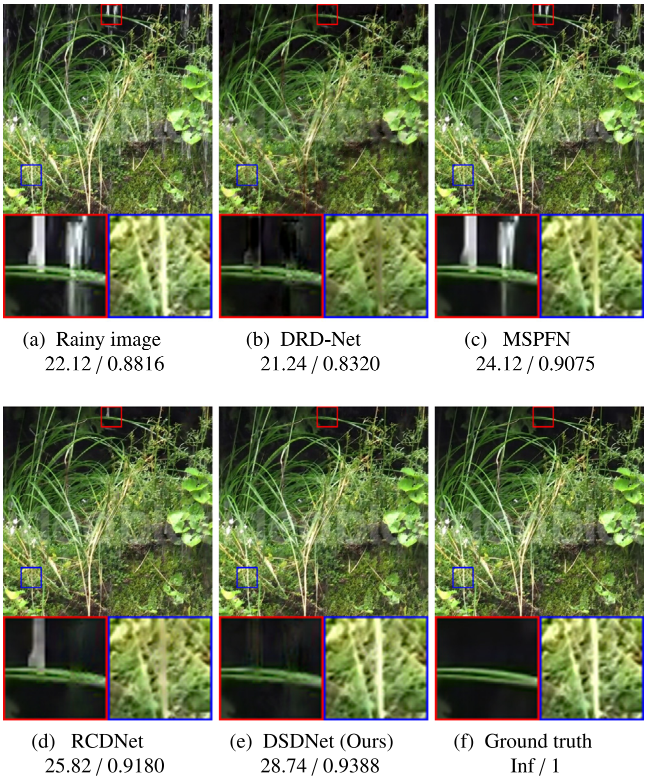

(a) Rainy image 22.12 / 0.8816
(b) DRD-Net 21.24 / 0.8320
(c) MSPFN 24.12 / 0.9075
(d) RCDNet 25.82 / 0.9180
(e) DSDNet (Ours) 28.74 / 0.9388
(f) Ground truth Inf / 1

Fig. 1. A comparison of real-world rain images in the SPA-Data dataset. PSNR/SSIM values are listed below the corresponding image. All the comparisons are limited by either (i) an over-smoothing problem, as shown by the stem in the blue box, or (ii) failing to remove rain streaks properly, as displayed by the region enclosed in the red box. While with the proposed stimulation mechanism, our framework is more suitable to cover both detail preserving and rain streak removal purposes, thus reaching better deraining performance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

images to their clean counterparts in an end-to-end manner. The pioneering work is proposed by Fu et al. (2017a) that utilizes a three-layer CNN with extracted high-pass components of rainy images as input. And it is later extended by replacing the simple convolutional layers with Resblocks (Fu et al., 2017b). Consequently, many other formulations have emerged, such as joint rain detection and removal (Yang et al., 2017), generative learning that uses an additional regularization for superior visualization (Zhang et al., 2019), and multi-scale information aggregation that exploits the correlations of rain streaks for rain removal (Jiang et al., 2020).

However, most deep learning based methods still suffer from two main issues: (1) failing to remove rain streaks accurately, particularly in the case of heavy rain, and (2) generating an over-smoothed background with a loss of details. This is due to the ill-posed property of the deraining problem, such that the deep models would be hindered by the inherent overlapping between rain streaks and background scenes from distinguishing either of them individually. Although recent methods make some efforts toward both issues, limited effectiveness has been accomplished. For example, Yang et al. (2017) propose to perform joint rain detection, estimation and removal successively in a cascaded framework. Nevertheless, the multi-task objective would create ambiguities in learning the shared parameters, leading to a trade-off between rain streak removal and detail preserving. In contrast, just as pointed out by Pan et al. (2018) that jointly estimating the structures and details in a dual-branch structure can be highly effective for low-level vision tasks, DRD-Net (Deng et al., 2020) uses a parallel pipeline to separately estimate the rain streaks and the detail complement. But the effect of rain streak detection for single image deraining is ignored. Specifically, there have been no explorations regarding a suitable relationship between rain distributions and the hidden feature responses in a deep model, which is all-important for the deraining task in two aspects: (1) additional cues would be provided by rain distributions for the model to predict rain streaks, and (2) concerning detail recovery, an ideal network should concentrate more on rain regions, where the contextual details are more likely to be damaged, rather than non-rain areas. Several typical failures by previous works can be seen in Fig. 1. The lack of information from rain distributions leads to incomplete removal of rain streaks or an over-smoothing problem. Note that even a rain mask, which indicates both location and intensity of rains, is caught in Yang et al. (2017), it is simply leveraged via being concatenated with the hidden feature maps, resulting in an underutilization of the auxiliary information for guiding the rain removal process.

To address the above problems, in this paper, we propose a convolutional neural network named DSDNet that separately estimates rain streaks and the loss of details in a dual-stream pipeline. With the aid of an additional prediction of a rain mask, extra information excavated from rain distributions can be utilized to guide the process of both tasks. In particular, to model a felicitous communication between the rain mask and the rainy image features, we design a *stimulation module* that respectively deployed into each stream of DSDNet. This strategy is to stimulate the image features such that the feature responses located in rain regions would be emphasized, while those belonging to non-rain areas can be depressed. In this way, the proposed DSDNet can focus more on the restoration of the rain region, therefore promoting a higher accuracy for rain streak removal as well as better detail reinforcement. In addition, we present a self-paced semi-curriculum learning mechanism to alleviate the learning ambiguity caused by the prediction of the rain mask. Extensive experiments conducted on synthetic rainy images from three benchmarks show that the proposed DSDNet outperforms the state-of-the-art deraining methods. Moreover, the potential generalization ability of DSDNet is validated on two benchmarks collected in real-world scenes. The superiority of the proposed method is also demonstrated via the evaluation of joint image deraining, detection, and segmentation tasks.

In summary, the contributions of this work are three-fold:

- We study an effective integration of regional cues for tackling the ill-posed single image deraining problem. To this end, we tailor a stimulation strategy that is embedded in both streams of the proposed DSDNet. It introduces the entanglement between hidden feature responses and the rain distribution, enabling an innovative detect-derain-restore paradigm for single image deraining.
- We present a self-paced semi-curriculum learning strategy intending to progressively decrease and finally get rid of the dependency on the ground truth mask during training. It can thus facilitate the mitigation of the ambiguity brought by the prediction of the rain mask and accelerate the training process.
- We perform favorably against the state-of-the-art single image deraining methods on several benchmarks in both synthetic and real-world scenarios. Also, the effectiveness of the proposed DSDNet is demonstrated via being spread to other applications, *i.e.*, detection, and segmentation.

The rest of this paper is organized as follows. Section 2 reviews the existing related single image deraining methods. Section 3 elaborates the proposed DSDNet. And Section 4 gives a detailed discussion based on the experimental evaluations of the proposed method. Finally, Section 5 concludes this paper.

2. Related work

Traditional model-based methods mainly delve into modeling optimization functions by enforcing various hand-crafted priors on the degraded images. For example, Chen and Hsu (2013) proposed a low-rank appearance model to represent and remove the spatiotemporally correlated rain streaks. With a sparsity-based regularization, Luo et al. (2015) proposed to separately estimate a rain streak layer and a background layer. Li et al. (2016) utilized a sequence of Gaussian mixture models to capture patch-based priors for accommodating multiple orientations and scales of the rain streaks. However, model-based approaches usually suffer from high computational complexity, as well as difficulty in designing efficient priors when facing a condition of heavy rain.

In recent years, deep neural networks have achieved tremendous success in dealing with the single image deraining task. Fu et al. (2017a) firstly proposed to use a simple three-layer CNN for rain removal. Zhang et al. (2019) proposed a conditional generative adversarial network (CGAN) with an additional regularization for a better quality of visualization. To iteratively reconstruct the rainy images, several works (Li et al., 2017; Yang et al., 2017) were proposed to leverage recurrent networks to remove the rain streaks stage by stage. As the dependencies of image features across different stages were neglected by previous works, RESCAN (Li et al., 2018) and PReNet (Ren et al., 2019) introduced advanced recurrent modules, *i.e.*, convolutional Long Short-Term Memory (convLSTM) and convolutional Gated Recurrent Unit (convGRU), to take the stage-wise result as input for the further restoration process. To take advantage of spatial contextual information for learning more effective features, formulas like multi-scale information aggregation (Zhang and Patel, 2018; Jiang et al., 2020), squeeze-and-excitation (SE) operation (Yang and Lu, 2019; Deng et al., 2020; Wang et al., 2020b), and spatial attention mechanism (Wang et al., 2019) were also widely utilized. Note that all the above methods were implemented using cascaded frameworks. Instead, DRDNet (Deng et al., 2020) was proposed to use a parallel structure to respectively tackle the rain streak removal and detail recovery problems. However, it ignores the influence of rain distribution which is a critical factor for rain removal. And in fact, all of the existing related methods show limited performance without investigating a proper interaction between rain distributions and feature responses, while the proposed DSDNet uses a self-contained stimulation strategy to achieve this goal.

3. Method

3.1. Overview

Following the spirit of Pan et al. (2018), we consider resolving rain streak removal as well as detail recovery problems in a dual-stream pipeline. And therefore, a Dual Stimulated Deraining Network (DSDNet) is established, which consists of a stimulated rain streak estimation stream (RStream) and a stimulated detail reinforcement stream (DStream). Fig. 2 illustrates the overall architecture of the proposed DSDNet.

In particular, as the distribution of rain is undoubtedly influential for both issues, our method additionally predicts a rain mask that contains information on the location and intensity of rain. Specifically, to reduce the system complexity, the rain mask is produced by the backbone of the RStream as a side output instead of introducing an extra subnetwork. This is because the two tasks, *i.e.*, rain detection and rain streak estimation, are quite similar. And the mask is further fed into two independent stimulation modules that are respectively embedded in both streams, providing more constraints to stimulate hidden feature responses. To alleviate the learning ambiguity, we also propose a self-paced semi-curriculum learning strategy for the estimation of the rain mask.

3.2. Network architecture

3.2.1. Stimulation module

The most possible reason for either an incorrect removal of rain streaks or a loss of texture details is the ill-posed property of the deraining problem. Nevertheless, there exists an implicit yet reasonable constraint that the severity of the degradation is closely related to the distribution (*i.e.*, location and intensity) of rain. This motivates us to treat the hidden feature responses differently according to a predicted rain mask. Specifically, we implement a stimulation module that respectively equipped in each stream of DSDNet to rescale the extracted features, with guidance from the rain mask.

Given a sequence of feature maps $\boldsymbol{F} \in \mathbb{R}^{C\times H\times W}$ generated by a backbone network and a single-channel rain mask $\boldsymbol{M} \in \mathbb{R}^{1\times H\times W}$ where the element values are ranged from 0 to 1 (0 indicates a non-rain area, while 1 denotes the highest intensity of rains), the proposed stimulation module $\mathcal{SM}(\cdot)$ produces stimulated features $\mathcal{SM}(\boldsymbol{F}, \boldsymbol{M}) \in \mathbb{R}^{C\times H\times W}$ via the following model:

$$\mathcal{SM}(\boldsymbol{F}, \boldsymbol{M}) = \alpha \odot E(\boldsymbol{M}) \odot \boldsymbol{F} + (1-\alpha) \odot \boldsymbol{F}, \tag{1}$$

where $\odot$ indicates an element-wise multiplication operation and α denotes a learnable parameter. $E(\cdot)$ is an expansion operation, for example, the stimulation maps $E(\boldsymbol{M}) \in \mathbb{R}^{C\times H\times W}$ is produced by a C-times stack of the rain mask $\boldsymbol{M}$ along the channel dimension.

Note that as a representative one of the few deraining methods associated with a subtask of rain detection, JORDER (Yang et al., 2017) adopted to directly concatenate the rain mask with the hidden features, which cannot fully utilize the information from the distribution of the rains to improve the feature representations. Also, the goal of rain detection would distract from its total objective due to the cascaded structure design. In contrast, our DSDNet prefers to perform deraining in a parallel pipeline with multiple aims. More importantly, the proposed stimulation strategy, which serves as an ancillary regularization, is capable of dynamically emphasizing the hidden features in rain regions while deemphasizing those in the non-rain areas. In this way, both rain streak estimation and detail reinforcement processes can be benefitted.

3.2.2. Stimulated rain streak estimation stream

The goal of RStream is divided into two parts: The first one is to detect rain streaks and render the result as a rain mask. With the aid of the additional cues furnished by the mask, the other aim is to predict the rain streak layer. Specifically, we first input a rainy image $\boldsymbol{O}$ to a flat convolutional layer with a filter size of 3 × 3. To exploit both spatial information as well as the channel dependencies, the extracted shallow features are then fed into a series of SE-Resblocks (Li et al., 2018). The last layer of the backbone of RStream is used for fusing the long-range information from the first conv-layer, and the filter size is set to 3 × 3. Consequently, we hook up an outer conv-layer conjunct with a Sigmoid layer to squeeze the preceded multi-channel feature maps into a preliminary single-channel rain mask $\boldsymbol{M}_p$, which is formulated as follows:

$$\boldsymbol{M}_p = \mathrm{Sigmoid}(\mathrm{Conv}(\mathcal{R}(\boldsymbol{O}))), \tag{2}$$

where $\mathcal{R}(\cdot)$ indicates the backbone of RStream. Note that before being fed into the proposed stimulation module to generate enhanced features, the preliminary mask $\boldsymbol{M}_p$ would be further updated to get the rain mask $\boldsymbol{M}$ via a self-contained curriculum learning strategy. We will expound this in Section 3.3.

Finally, once the features are stimulated, they would be sent into a 3 × 3 conv-layer for reconstructing a three-channel rain streak layer $\boldsymbol{S}$. This process can be calculated as follows:

$$\boldsymbol{S} = \mathrm{Conv}(\mathcal{SM}_{\mathcal{R}}(\mathcal{R}(\boldsymbol{O}), \boldsymbol{M})), \tag{3}$$

where $\mathcal{SM}_{\mathcal{R}}(\cdot)$ denotes the stimulation module embedded in RStream. As thus, we can calculate a coarse deraining result $\boldsymbol{B}_{coarse}$ via the following formulation:

$$\boldsymbol{B}_{coarse} = \boldsymbol{O} - \boldsymbol{S}. \tag{4}$$

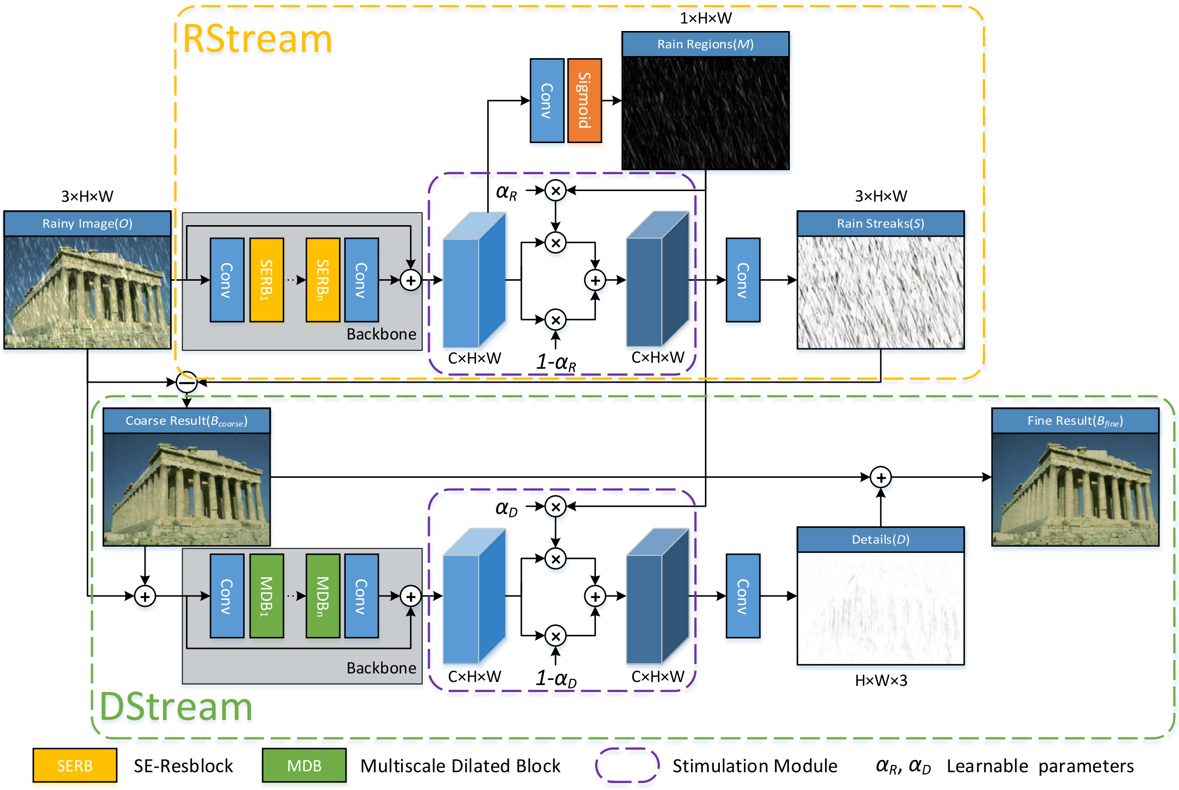


Fig. 2. The overall architecture of our DSD-Net. It consists of two parallel streams: stimulated Rain Streak Estimation Stream (RStream), and stimulated Detail Reinforcement Stream (DStream). RStream in the upper part is used for both rain region detection and rain streak prediction. The lower part is DStream which aggregates the contextual information to restore the lost details. Particularly, the tailored stimulation module is exhibited inside the red dotted box. The generated rain mask would be respectively fed into dual stimulation modules to guide rain streak estimation as well as detail reinforcement processes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.3. Stimulated detail reinforcement stream

The coarse deraining result may incur a loss of details. Apart from the ill-posed property of the problem, the limited receptive field may also be one concern. To boost the quality of the coarse result, we typically construct a DStream to predict the lost details. Since dilated convolution has shown a preferable performance in recent works (Yang et al., 2017; Li et al., 2018), here we simply employ several multiscale dilated blocks (Yang et al., 2017) as the base of the backbone. Such a kind of block is conducive to amplifying the receptive field as well as leveraging multiscale information, so as to promote the aggregation of spatial contextual information and facilitate the detail reinforcement process. Note that except for the basic blocks, all the remaining structure of DStream is the same as RStream.

During the detail reinforcement stage, we first add the coarse deraining result $\boldsymbol{B}_{coarse}$ to its corresponding rainy image $\boldsymbol{O}$, mainly for a complement of details provided by the former one. Then we inject this sum into DStream $\mathcal{D}(\cdot)$ to obtain the details $\boldsymbol{D}$, which can be formulated as follows:

$$\boldsymbol{D} = \mathrm{Conv}(\mathcal{SM}_{\mathcal{D}}(\mathcal{D}(\boldsymbol{O} + \boldsymbol{B}_{coarse}), \boldsymbol{M})), \tag{5}$$

where $\mathcal{SM}_{\mathcal{D}}(\cdot)$ indicates the stimulation module deployed in DStream. In the end, the fine deraining result $\boldsymbol{B}_{fine}$ is given by

$$\boldsymbol{B}_{fine} = \boldsymbol{B}_{coarse} + \boldsymbol{D}. \tag{6}$$

3.3. Self-paced semi-curriculum learning

Note that the rain mask that fed into the stimulation module can be obtained in advance by other upstream detection methods, as long as it can precisely represent the distribution of rains. Nonetheless, to increase the flexibility of the proposed DSDNet, we choose to jointly estimate the rain mask and rain streaks in RStream. A subsequent issue is created that the learning ambiguity would be aggravated due to a multi-objective prediction. To solve this, inspired by the curriculum learning technology (Bengio et al., 2009) that training a framework with an easy start and gradually increasing the difficulty to expedite a better convergence, we particularly propose a self-paced curriculum learning strategy to update the rain mask.

Specifically, the rain mask $\boldsymbol{M}$ is calculated from a trade-off between the ground truth mask $\boldsymbol{M}_g$ and the preliminary mask $\boldsymbol{M}_p$, which is formulated as follows:

$$\boldsymbol{M} = \beta \odot \boldsymbol{M}_g + (1 - \beta) \odot \boldsymbol{M}_p, \tag{7}$$

where β is a curriculum rate that varies with the training difficulty changes. Note that the ground truth mask $\boldsymbol{M}_g$ we used is the difference between the rainy image and its corresponding rain-free image in the luminance channel of a YCbCr color space, since it is not provided by most existing datasets. We further give the exact definition of the curriculum rate, that is

$$\beta = \begin{cases} 1, & \gamma_{max} < \mathcal{L}_{mask}, \\ \frac{\mathcal{L}_{mask} - \gamma_{min}}{\gamma_{max} - \gamma_{min}}, & \gamma_{min} < \mathcal{L}_{mask} \le \gamma_{max}, \\ 0, & \mathcal{L}_{mask} \le \gamma_{min}, \end{cases} \tag{8}$$

where $\mathcal{L}_{mask}$ is a loss term regarding the estimation of the rain mask that would be discussed in the next section, γ_{min} and γ_{max} are two constants that selected empirically.

Eq. (8) forces that the curriculum rate progressively decays from 1 to 0 during training. It implies that, the update of the rain mask $\boldsymbol{M}$ relies more on the ground truth mask $\boldsymbol{M}_g$ during the beginning of training for alleviating a learning ambiguity, and then increasingly shifts toward the preliminary mask $\boldsymbol{M}_p$. When the loss term reaches a low plateau, the required mask $\boldsymbol{M}$ would be equal to $\boldsymbol{M}_p$, such that getting rid of the dependency of the ground truth mask.

To some extent, our model can be seen as a paradigm with multiple knowledge representations (Yang et al., 2021; Luo et al., 2017) for single image deraining. First, in contrast with the extracted image features which are flowed in the backbone network, the rain mask involved in the proposed stimulation module can be seen as a lower-level abstraction that provides privileged information to promote both procedures of rain streak estimation and detail reinforcement. Second, the dual-stream structure establishes an entanglement between the rain streaks and the latent clean image and explores reciprocal enhancements for each other.

3.4. Objective function

We mainly consider the consistencies in three aspects for optimizing the proposed model, and design an objective function with three components: a pixel consistency loss $\mathcal{L}_{pixel}$ between the final result $\boldsymbol{B}_{fine}$ and its corresponding ground truth clean image $\boldsymbol{B}_g$, a rain streak consistency loss $\mathcal{L}_{rain}$ between the predicted rain streak image $\boldsymbol{S}$ and the corresponding ground truth rain streaks $\boldsymbol{S}_g$, and a mask consistency loss $\mathcal{L}_{mask}$ between the rain mask $\boldsymbol{M}$ and its corresponding ground truth rain mask $\boldsymbol{M}_g$. The total objective $\mathcal{L}$ is defined as follows:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{pixel} + \lambda_2 \mathcal{L}_{rain} + \lambda_3 \mathcal{L}_{mask}. \tag{9}$$

Regarding image processing tasks, mean square error (MSE) is one of the most commonly used loss functions. We use such a plain measurement to formulate each term of our objective, which is as follows:

$$\mathcal{L}_{pixel} = \sum_{i=1}^{N} \left\| \boldsymbol{B}^i_{fine} - \boldsymbol{B}^i_g \right\|^2, \tag{10}$$

$$\mathcal{L}_{rain} = \sum_{i=1}^{N} \left\| \boldsymbol{S}^i - \boldsymbol{S}^i_g \right\|^2, \tag{11}$$

$$\mathcal{L}_{mask} = \sum_{i=1}^{N} \left\| \boldsymbol{M}^i - \boldsymbol{M}^i_g \right\|^2, \tag{12}$$

where i denotes the index of training samples, and N is the total number of samples. λ_1, λ_2, and λ_3 indicate the balance factors with respect to each component.

4. Experiments

In this section, we first clarify the experimental settings used for evaluating the proposed DSDNet. And then, we demonstrate an ablation study to examine the effectiveness of different components in our framework. Subsequently, we present the assessments of DSDNet regarding both synthetic and real-world scenarios. Finally, we spread our model to other applications, *i.e.*, detection and segmentation, to manifest a more comprehensive evaluation.

4.1. Experimental settings

4.1.1. Experimental data

For the case of synthetic image deraining, we utilize three benchmark datasets to evaluate the proposed DSDNet: *Rain200L* (Yang et al., 2017), *Rain200H* (Yang et al., 2017), and *Rain1400* (Fu et al., 2017b). *Rain200L* is a light rain dataset that consists of 1800 image pairs for training and 200 image pairs for testing. *Rain200H* is a heavy rain dataset that also includes 1800 training samples and 200 testing samples. And *Rain1400* includes 14 000 rainy images which are synthesized from 1000 clean images with 14 kinds of rain streaks, among which 100 clean images (1400 rainy images) are chosen for testing and others are used for training.

Regarding real image deraining, two real-world datasets are further leveraged to validate the robustness of our approach, *i.e.*, *SPA-Data* (Wang et al., 2019) and *Real-Internet* (Wang et al., 2019). SPA-Data includes 638 492 rainy/clean training pairs and 1000 testing pairs, while Real-Internet is with 146 rainy images from the Internet collected by Wang et al. (2019). Note that the testing images from Real-Internet have no clean ground truths.

Regarding the practicality of DSDNet in the applications of detection and segmentation, we use two synthetic datasets *COCO350* (Jiang et al., 2020) and *BDD150* (Jiang et al., 2020) for investigation. They are randomly selected from *COCO* (Caesar et al., 2018) and *BDD* (Yu et al., 2018), respectively, with diverse rain streak directions and intensities synthesized by Photoshop.

4.1.2. Implementation details

Our model is implemented with the PyTorch framework and trained on an NVIDIA GeForce GTX 1080Ti GPU. During training, we utilize Adam (Kingma and Ba, 2015) as the optimizer and a batch size of 8. The learning rate is initialized to 2×10^{-4} and remained the same in the first 20 epochs. And it then linearly decayed to zero from the 21st epoch to the 200th epoch.

In the experiments, the numbers of the SE-Resblocks and the multi-scale dilated blocks in the backbones are both set to 16. To obtain the best performance, we empirically set the parameters as $\gamma_{min} = 0.005$, $\gamma_{max} = 0.1$, $\lambda_1 = 1$, $\lambda_2 = 0.1$ and $\lambda_3 = 0.1$. The input images are randomly cropped with a size of 128 × 128 before being fed into the network.

4.1.3. Comparisons and measurements

We compare our method with several state-of-the-art methods, including RESCAN (Li et al., 2018), PReNet (Ren et al., 2019), SPANet (Wang et al., 2019), JORDER-E (Yang et al., 2019), DRD-Net (Deng et al., 2020), MSPFN (Jiang et al., 2020), RCDNet (Wang et al., 2020a), and MPRNet (Zamir et al., 2021). Unless stated otherwise, all the reported results are evaluated from the models trained by the corresponding dataset in their default settings for fairness.

Regarding the results from the datasets in which the ground truth images are available, we utilize PSNR and SSIM as the measurements. And we use SSEQ, ENIQA, and BRISQUE to measure the visual quality of the results from the other datasets. Note that all the metrics are calculated in the RGB color space.

4.2. Ablation study

In this subsection, we explore the effectiveness of each component in the proposed network. We design six comparisons and test their performances on *Rain200L*, *Rain200H*, and *Rain1400*. Specifically, the settings of different models are defined as follows:

- *RS:* Only RStream is used to derain.
- *DS:* Only DStream is used to derain.
- *RS+SM:* Incorporating the proposed stimulation module into RStream to remove rain streaks with additional cues provided by rain mask.
- *DS+SM:* Incorporating the proposed stimulation module into DStream to remove rain streaks with additional cues provided by rain mask.
- *RS+DS:* Combining RStream and DStream to jointly remove rain streaks and recover image details.
- *RS+DS+SM:* Incorporating the stimulation module into RStream and DStream respectively.
- *RN+DN+SM+SSL:* Our final DSDNet which additionally exploits a self-paced semi-curriculum learning strategy.

Table 1 lists all the quantitative results. First, baseline performances towards rain streaks removal and details reinforcement are given by *RS* and *DS*. After introducing the stimulation module into *RS* and *DS* respectively, average PSNR improvements of 0.18 dB and 0.24 dB are

Table 1
Ablation study of different network designs.

Method	Rain200H			Rain200L			Rain1400		
	PSNR (dB)	SSIM	Time (h)	PSNR (dB)	SSIM	Time (h)	PSNR (dB)	SSIM	Time (h)
RS	27.92	0.8624	2.5	37.87	0.9823	2.38	31.43	0.9178	6.65
DS	28.31	0.8746	6.0	37.72	0.9822	7.54	31.50	0.9199	18.38
RS+SM	28.27	0.8725	3.6	37.94	0.9824	3.85	31.54	0.9182	11.49
DS+SM	28.69	0.8809	7.0	37.91	0.9824	9.44	31.65	0.9222	22.40
RS+DS	28.55	0.8779	12.0	37.92	0.9819	13.89	31.60	0.9206	28.11
RS+DS+SM	28.85	0.8885	14.6	38.05	**0.9828**	16.25	31.76	0.9234	34.24
RS+DS+SM+SSL	**28.91**	**0.8895**	6.0	**38.06**	0.9827	10.56	**31.94**	**0.9263**	22.82

Table 2
Quantitative evaluations for synthetic image deraining on three benchmark datasets. The top-3 performances are marked in red, blue, and green.

Method	Venue & Year	Rain200L		Rain200H		Rain1400		Average	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
RESCAN	ECCV18	37.09	0.9787	26.64	0.8363	30.91	0.9090	31.55	0.9080
PReNet	CVPR19	36.69	0.9799	27.96	0.8894	30.56	0.9136	31.74	0.9276
SPANet	CVPR19	35.87	0.9762	25.86	0.8530	28.90	0.8875	30.21	0.9056
JORDER-E	PAMI19	37.75	0.9825	27.88	0.8730	31.11	0.9158	32.25	0.9238
DRD-Net	CVPR20	37.15	0.9811	28.17	0.8701	29.11	0.8912	31.48	0.9141
MSPFN	CVPR20	30.12	0.9166	25.53	0.7969	31.14	0.9137	28.93	0.8757
RCDNet	CVPR20	38.41	0.9839	28.71	0.8875	31.37	0.9196	32.83	0.9303
MPRNet	CVPR21	38.21	0.9842	28.27	0.8806	31.88	0.9277	32.79	0.9308
DSDNet		38.06	0.9827	28.91	0.8895	31.94	0.9263	32.97	0.9328

achieved. This implies that the guidance of rain distributions provided by the stimulation module can significantly boost the performances of baselines. Moreover, we observe that combining *RS* and *DS* surpass that using only either by average PSNR improvements of 0.28 dB and 0.18 dB, revealing that the two subnetworks with different structures can learn the complementary information to promote the deraining performance. On this basis, it is not surprising that an average increase of 0.20 dB in PSNR is further accomplished by embedding the stimulation modules into each subnetwork. Finally, it can be seen that our curriculum learning strategy mainly helps to accelerate the training, despite only slight improvement gains in the measurements. Fig. 3 shows the visual results of the ablation study on *Rain200H*. We can see that the deraining result of our DSDNet (*RS+DS+SM+SSL*) is the most visibly plausible one in that the background (sky) is purer compared to the other results.

4.3. Experiments on synthetic data

Table 2 reports the quantitative evaluation for synthetic image deraining on three benchmark datasets. In particular, DSDNet exhibits powerful ability when dealing with heavy-rain scenes (*Rain200H* and *Rain1400*), with increases of 0.20 dB and 0.06 dB in PSNR compared to that of the second best method MPRNet. This implies that a stimulation module design that yields a prediction of dense rain distributions can facilitate rain streak estimation and detail restoration. Another interesting observation is that, when coping with light-rain situations (*Rain200L*), our method seems not to achieve the desired performance (in third place). We speculate that this is due to the information yielded by a sparse rain distribution being deficient, which sometimes may even mislead the learning process. Nevertheless, it still can be seen that the proposed DSDNet reaches superior performance on average against all the competitors.

We visualize the restored results from the test datasets by different comparisons in Fig. 4. It can be observed that our DSDNet can sufficiently remove the rain streaks and subtly recover the details on different rain conditions compared with other methods. Moreover, thanks to our stimulation mechanism, the DSD-Net can be enforced to focus on the rain regions to obtain better deraining performance as well as alleviate an over-smoothing problem that existed in most deraining methods.

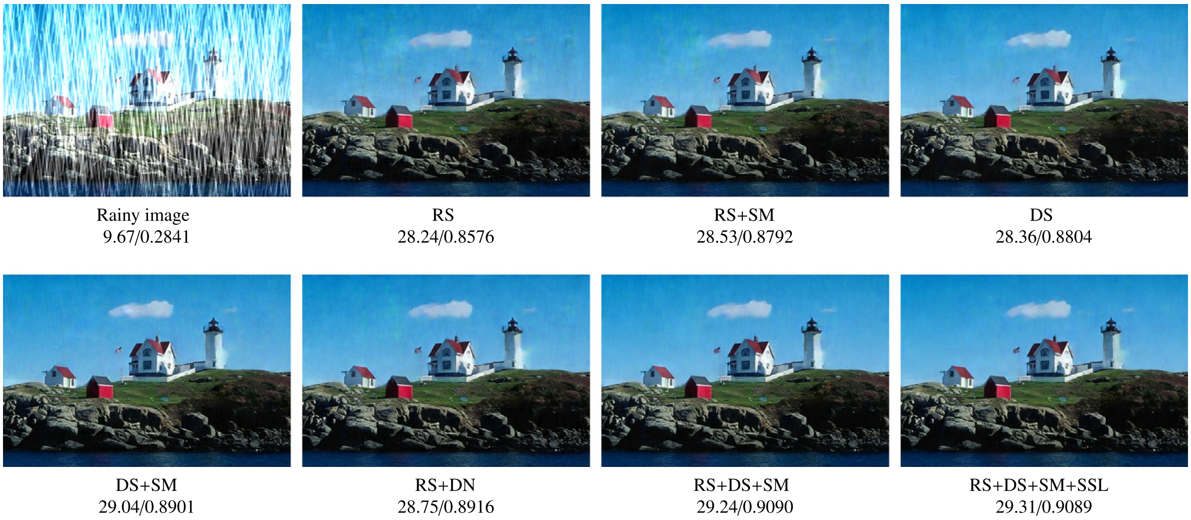
Table 3
Quantitative evaluation for real-world image deraining on the SPA-Data dataset. The top-2 performances are marked in red and blue.

Metric	MPRNet	PReNet	SPANet	JORDER-E
PSNR (dB)	32.73	39.44	38.53	40.48
SSIM	0.9398	0.9813	0.9875	0.9818
	DRD-Net	MSPFN	RCDNet	**DSDNet**
PSNR (dB)	35.91	39.59	38.55	40.62
SSIM	0.9595	0.9789	0.9763	0.9818

4.4. Experiments on real-world data

To validate the robustness of the proposed method, we conduct experiments on real-world image deraining. Fig. 5 illustrates the visual comparisons on the SPA-Data dataset in different scenarios, including nature, fence, urban, and building. We can see that for the real rain patterns, our DSD-Net also shows superiority compared to other methods. Particularly, for the regions with rain streaks, our method restores fine details while for the regions without rains, our method would not blur the background details. For example, in the first row of Fig. 5, DSDNet successfully preserves the stem of the plant in the blue bounding box while other methods inaccurately remove it as rain streaks. This is also proved by the quantitative evaluation summarized in Table 3. The metrics of the proposed method beat all the other methods, demonstrating its effectiveness in real-world scenarios.

We also perform another evaluation on the Real-Internet dataset. Note that all the comparisons are trained on *Rain200H* dataset. We select four hard samples with various rain streaks and complex backgrounds to verify the generalization ability of all the methods and display the results in Fig. 6. We can see that most of the competitors can remove the rain streaks well but simultaneously blur some background details, such as the veins of the leaves in the fourth row of Fig. 6. In contrast, our DSDNet can preserve most of the details meanwhile removing rain streaks more accurately. Note that despite DSDNet, which also adopts a parallel structure, can achieve a similar qualitative result enclosed in the red box with ours (the fourth row), more rain prints can be found in its recovered background. This indicates the effectiveness of our stimulation module which considers rain distribution as auxiliary information for image deraining. The quantitative results are reported in Table 4.



Rainy image 9.67/0.2841

RS 28.24/0.8576

RS+SM 28.53/0.8792

DS 28.36/0.8804

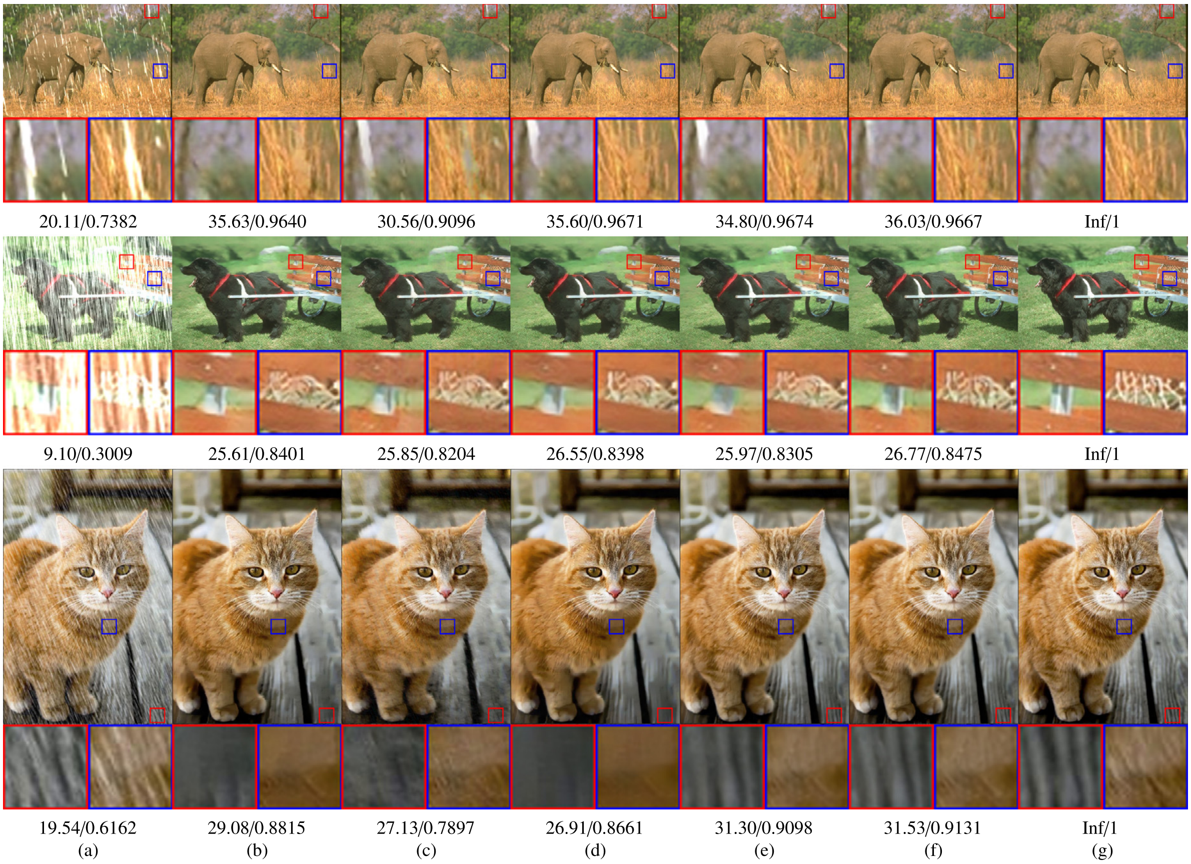
DS+SM 29.04/0.8901

RS+DN 28.75/0.8916

RS+DS+SM 29.24/0.9090

RS+DS+SM+SSL 29.31/0.9089

Fig. 3. Visual comparisons on Rain200H of different models in ablation study. PSNR/SSIM are listed below the images.



20.11/0.7382 | 35.63/0.9640 | 30.56/0.9096 | 35.60/0.9671 | 34.80/0.9674 | 36.03/0.9667 | Inf/1

9.10/0.3009 | 25.61/0.8401 | 25.85/0.8204 | 26.55/0.8398 | 25.97/0.8305 | 26.77/0.8475 | Inf/1

19.54/0.6162 (a) | 29.08/0.8815 (b) | 27.13/0.7897 (c) | 26.91/0.8661 (d) | 31.30/0.9098 (e) | 31.53/0.9131 (f) | Inf/1 (g)

Fig. 4. Visual comparison of synthetic rainy images in the Rain200L, Rain200H, and Rain1400 datasets. From (a) to (g): (a) input rainy image, (b) PReNet (Ren et al., 2019), (c) DRD-Net (Deng et al., 2020), (d) RCDNet (Wang et al., 2020a), (e) MPRNet (Zamir et al., 2021), (f) Ours, and (g) ground truth image. PSNR/SSIM are listed below the images.

4.5. Evaluation on other applications

In this subsection, we investigate the practicality of the DSDNet by spreading it to other computer vision tasks. The overall process is that, first, we restore the rainy images from *COCO350* and *BDD150* datasets by using all the state-of-the-art comparisons. Then we feed the results into two popular algorithms for two typical purposes: YOLOv3 (Redmon and Farhadi, 2018) for object detection and RefineNet (Lin et al., 2017) for semantic segmentation. Tables 5 and 6 shows the quantitative evaluation of synthetic image deraining as well as the subsequent detection and segmentation. Visual comparisons are shown in Fig. 7.

We can find that rain streaks can greatly degrade the quality of images, leading to poor detection accuracy and segmentation precision. The detector usually misses the targets or outputs some uncertain results when suffering from rain. However, when we feed the rain-free images produced by our DSDNet into the detector, the detection precision is improved by 21.55% which performs best against other competing deraining methods. When it comes to semantic segmentation, the DSDNet also achieves superior performance with 54.03% mPA

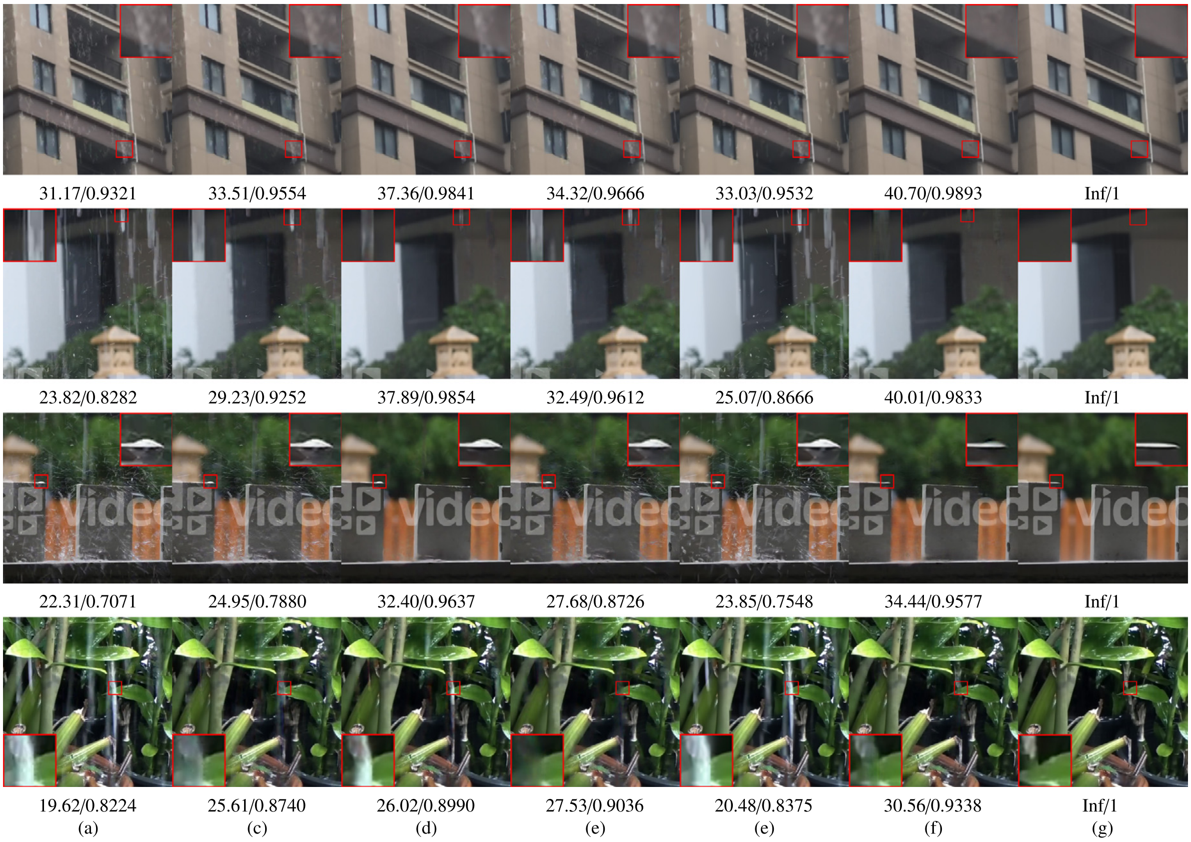


Fig. 5. Visual comparison of real rainy images in the SPA-Data dataset. From (a) to (g): (a) input rainy image, (b) DRD-Net (Deng et al., 2020), (c) MSPFN (Jiang et al., 2020), (d) RCDNet (Wang et al., 2020a), (e) MPRNet (Zamir et al., 2021), (f) Ours, and (g) ground truth image. PSNR/SSIM are listed below the images. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

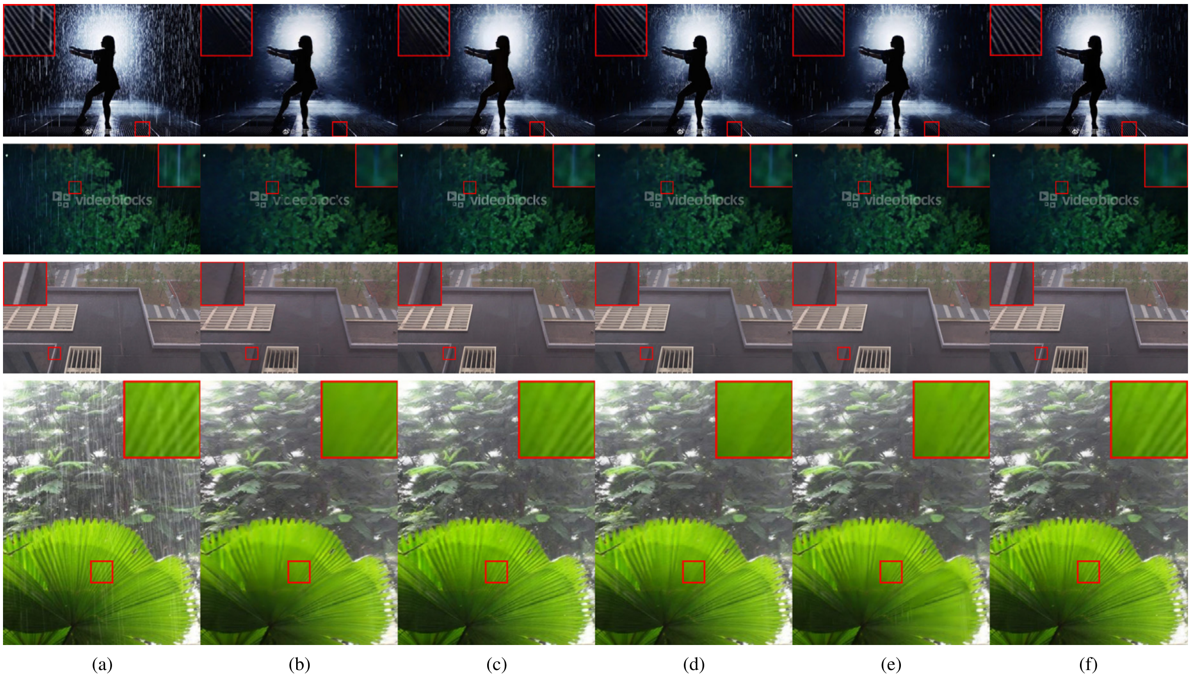


Fig. 6. Visual comparison of real rainy images in the Real-Internet dataset. From (a) to (g): (a) input rainy image, (b) SPANet (Wang et al., 2019), (c) DRD-Net (Deng et al., 2020), (d) RCDNet (Wang et al., 2020a), (e) MPRNet (Zamir et al., 2021), and (f) Ours.

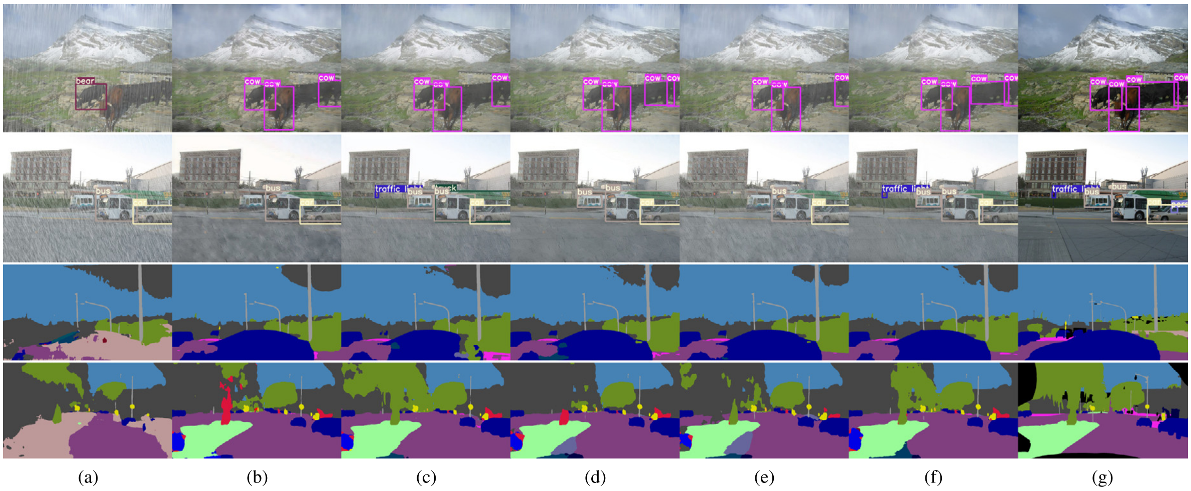


Fig. 7. Visual comparison of recovered images for detection and semantic segmentation. From (a) to (g): (a) rainy image, (b) SPANet (Wang et al., 2019), (c) DRD-Net (Deng et al., 2020), (d) RCDNet (Wang et al., 2020a), (e) MPRNet (Zamir et al., 2021), (f) Ours, and (g) ground truth image. The first and second rows are the results for detection by YOLOv3 (Redmon and Farhadi, 2018) on the COCO350 dataset. The third and fourth rows are the semantic segmentation results obtained from RefineNet (Lin et al., 2017) on the BDD150 dataset.

Table 4
Quantitative evaluation for real-world image deraining on the Real-Internet dataset. All the comparisons are trained on the Rain200H dataset. And all the metrics are the lower the better. The top-2 performances are marked in red and blue.

Method	SSEQ↓	ENIQA↓	BRISQUE↓
RESCAN	27.9040	0.1895	30.6595
PReNet	28.0166	0.1816	30.9984
SPANet	29.7238	0.1980	31.8892
JORDER-E	28.3976	0.1886	31.7157
DRD-Net	27.3162	0.1733	28.8796
MSPFN	32.7218	0.2000	35.4624
RCDNet	28.8457	0.1907	32.0494
MPRNet	28.5558	0.1871	32.4257
DSDNet	26.7733	0.1731	29.3061

Table 5
Quantitative evaluation on COCO350. The IoU threshold is set to 0.5. All the comparisons are trained on the Rain200H dataset. The top-3 performances are marked in red, blue, and green.

Method	PSNR (dB)	SSIM	Precision (%)	Recall (%)
Rain Input	13.24	0.5429	37.82	39.80
RESCAN	15.12	0.6065	46.08	48.48
PReNet	15.60	0.6080	46.42	48.64
SPANet	16.12	0.6195	47.21	49.63
JORDER-E	15.30	0.6029	49.28	51.27
DRD-Net	15.85	0.6265	51.97	53.88
MSPFN[a]	15.95	0.6398	56.35	58.23
RCDNet	15.27	0.6204	50.78	52.85
MPRNet	14.94	0.6048	47.77	49.65
DSDNet	16.16	0.6279	53.51	55.30
DSDNet[a]	16.61	0.6704	59.37	60.76

[a] Denotes that the method is trained on the dataset used in Jiang et al. (2020).

and 41.51% mIoU. The results show that the rain-free images produced by DSDNet retrieve more details (higher PSNR/SSIM) and credible information to effectively promote the performance of detection and semantic segmentation.

4.6. Running time

To evaluate the practicality of the proposed method, we summarize the average running times of RESCAN, SPANet, JORDER-E, DRD-Net, RCDNet, and our DSDNet in Table 7. All these experiments are performed on the COCO350 dataset in which the size of the test images is 640 × 480, with an NVIDIA 3090 GPU.

Table 6
Quantitative evaluation on BDD150. All the comparisons are trained on the Rain200H dataset. The top-3 performances are marked in red, blue, and green.

Method	PSNR (dB)	SSIM	mPA (%)	mIoU (%)
Rain Input	16.81	0.6764	38.06	28.38
RESCAN	19.18	0.7762	46.09	36.39
PReNet	19.71	0.7926	46.98	34.31
SPANet	20.30	0.7936	47.66	35.49
JORDER-E	19.34	0.7564	47.30	34.85
DRD-Net	20.00	0.8016	49.57	38.46
MSPFN[a]	20.75	0.8564	53.47	39.74
RCDNet	19.88	0.8095	48.69	36.60
MPRNet	18.89	0.7740	47.02	36.80
DSDNet	20.38	0.8011	49.18	37.05
DSDNet[a]	21.41	0.8687	54.03	41.51

[a] Denotes that the method is trained on the dataset used in Jiang et al. (2020).

Table 7
Running time of the comparisons for deraining of the images from COCO350 with a size of 640 × 480.

Method	RESCAN	SPANet	JORDER-E
Runtime (s)	0.153	0.362	0.321
	DRD-Net	RCDNet	DSDNet
Runtime (s)	0.442	0.416	0.694

It can be observed that DSDNet requires a bit more time for inference than the other methods. This may be caused by the cost of the proposed stimulation modules. On the other hand, the running time of our model is quite acceptable considering a trade-off between deraining performance and complexity.

4.7. Limitation

Although we demonstrate the cruciality of the interaction between the rain distribution and the hidden feature responses of the model for single image deraining, it would be limited by the accuracy of the prediction of the rain mask. Also, our model still needs the ground truth mask at the beginning of training. Therefore, it is interesting to explore the possibility of introducing self-supervised learning based strategies in the future.

5. Conclusion

In this paper, we explore the utility of rain region masks and propose a novel stimulation module to incorporate rain region masks into rain streak removal and detail restoration. We construct a dual-stream network to jointly estimate rain streaks and restore details with the guidance of the rain region masks. The stimulation module enables the network to focus on the rain regions which helps to get better deraining results as well as alleviate the over-smoothing problem in rain-free regions. To handle the ambiguity of rain region detection, we design a self-paced semi-curriculum learning strategy to gradually increase the difficulty of the learning process by replacing the ground truth rain region mask with our predicted mask under the control of the well-designed curriculum rate. Extensive quantitative and qualitative results on synthetic and real-world datasets indicate that our method outperforms the state-of-the-art single image deraining approaches and demonstrates effectiveness in different rain conditions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This project is supported by the National Natural Science Foundation of China (No. 62102381, U1706218, 41927805, 61972162); Shandong Natural Science Foundation, China (ZR2021QF035); Fundamental Research Funds for the Central Universities, China (202113035); the National Key R&D Program of China (2018AAA0100600); the China Postdoctoral Science Foundation (2020M682240, 2021T140631); Guangdong Natural Science Funds for Distinguished Young Scholar, China (No. 2023B1515020097); Guangdong International Science and Technology Cooperation Project (No. 2021A0505030009); Guangdong Natural Science Foundation (No. 2021A1515012625); Guangzhou Basic and Applied Research Project (No. 202102021074). All authors approved the version of the manuscript to be published.

References

Bengio, Y., Louradour, J., Collobert, R., Weston, J., 2009. Curriculum learning. In: ICML. pp. 41–48.

Caesar, H., Uijlings, J., Ferrari, V., 2018. Coco-stuff: Thing and stuff classes in context. In: CVPR. pp. 1209–1218.

Chen, Y.-L., Hsu, C.-T., 2013. A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In: ICCV. pp. 1968–1975.

Comaniciu, D., Ramesh, V., Meer, P., 2003. Kernel-based object tracking. IEEE TPAMI 25 (5), 564–577.

Deng, S., Wei, M., Wang, J., Feng, Y., Liang, L., Xie, H., Wang, F.L., Wang, M., 2020. Detail-recovery image deraining via context aggregation networks. In: CVPR. pp. 14560–14569.

Fu, X., Huang, J., Ding, X., Liao, Y., Paisley, J., 2017a. Clearing the skies: A deep network architecture for single-image rain removal. IEEE TIP 26 (6), 2944–2956.

Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., Paisley, J., 2017b. Removing rain from single images via a deep detail network. In: CVPR. pp. 3855–3863.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. In: ICCV. pp. 2961–2969.

Jiang, T.-X., Huang, T.-Z., Zhao, X.-L., Deng, L.-J., Wang, Y., 2017. A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors. In: CVPR. pp. 4057–4066.

Jiang, K., Wang, Z., Yi, P., Chen, C., Huang, B., Luo, Y., Ma, J., Jiang, J., 2020. Multi-scale progressive fusion network for single image deraining. In: CVPR. pp. 8346–8355.

Kang, L.-W., Lin, C.-W., Fu, Y.-H., 2011. Automatic single-image-based rain streaks removal via image decomposition. IEEE TIP 21 (4), 1742–1755.

Kim, J.-H., Sim, J.-Y., Kim, C.-S., 2015. Video deraining and desnowing using temporal correlation and low-rank matrix completion. IEEE TIP 24 (9), 2658–2670.

Kingma, D.P., Ba, J., 2015. Adam: A methodfor stochastic optimization. In: ICLR.

Li, R., Cheong, L.-F., Tan, R.T., 2017. Single image deraining using scale-aware multi-stage recurrent network. arXiv preprint arXiv:1712.06830.

Li, Y., Tan, R.T., Guo, X., Lu, J., Brown, M.S., 2016. Rain streak removal using layer priors. In: CVPR. pp. 2736–2744.

Li, X., Wu, J., Lin, Z., Liu, H., Zha, H., 2018. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: ECCV. pp. 254–269.

Lin, G., Milan, A., Shen, C., Reid, I., 2017. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In: CVPR. pp. 1925–1934.

Luo, M., Chang, X., Li, Z., Nie, L., Hauptmann, A.G., Zheng, Q., 2017. Simple to complex cross-modal learning to rank. CVIU 163, 67–77.

Luo, Y., Xu, Y., Ji, H., 2015. Removing rain from a single image via discriminative sparse coding. In: ICCV. pp. 3397–3405.

Pan, J., Liu, S., Sun, D., Zhang, J., Liu, Y., Ren, J., Li, Z., Tang, J., Lu, H., Tai, Y.-W., et al., 2018. Learning dual convolutional neural networks for low-level vision. In: CVPR. pp. 3070–3079.

Redmon, J., Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Ren, D., Zuo, W., Hu, Q., Zhu, P., Meng, D., 2019. Progressive image deraining networks: A better and simpler baseline. In: CVPR. pp. 3937–3946.

Shehata, M.S., Cai, J., Badawy, W.M., Burr, T.W., Pervez, M.S., Johannesson, R.J., Radmanesh, A., 2008. Video-based automatic incident detection for smart roads: The outdoor environmental challenges regarding false alarms. IEEE TITS 9 (2), 349–360.

Tripathi, A.K., Mukhopadhyay, S., 2014. Removal of rain from videos: a review. Signal Image Video Process. 8 (8), 1421–1430.

Wang, Y., Song, Y., Ma, C., Zeng, B., 2020b. Rethinking image deraining via rain streaks and vapors. In: ECCV. pp. 367–382.

Wang, H., Xie, Q., Zhao, Q., Meng, D., 2020a. A model-driven deep neural network for single image rain removal. In: CVPR. pp. 3103–3112.

Wang, T., Yang, X., Xu, K., Chen, S., Zhang, Q., Lau, R.W., 2019. Spatial attentive single-image deraining with a high quality real rain dataset. In: CVPR. pp. 12270–12279.

Yang, Y., Lu, H., 2019. Single image deraining using a recurrent multi-scale aggregation and enhancement network. In: ICME. IEEE, pp. 1378–1383.

Yang, W., Tan, R.T., Feng, J., Guo, Z., Yan, S., Liu, J., 2019. Joint rain detection and removal from a single image with contextualized deep networks. IEEE TPAMI 42 (6), 1377–1393.

Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., Yan, S., 2017. Deep joint rain detection and removal from a single image. In: CVPR. pp. 1357–1366.

Yang, Y., Zhuang, Y., Pan, Y., 2021. Multiple knowledge representation for big data artificial intelligence: framework, applications, and case studies. Front. Inf. Technol. Electron. Eng. 22 (12), 1551–1558.

Yu, F., Xian, W., Chen, Y., Liu, F., Liao, M., Madhavan, V., Darrell, T., 2018. Bdd100k: A diverse driving video database with scalable annotation tooling. arXiv preprint arXiv:1805.04687, 2, 6.

Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.-H., Shao, L., 2021. Multi-stage progressive image restoration. In: CVPR. pp. 14821–14831.

Zhang, H., Patel, V.M., 2018. Density-aware single image de-raining using a multi-stream dense network. In: CVPR. pp. 695–704.

Zhang, H., Sindagi, V., Patel, V.M., 2019. Image de-raining using a conditional generative adversarial network. IEEE TCSVT.