

# Single Image Deraining: From Model-Based to Data-Driven and Beyond

Wenhan Yang<sup>1</sup>, Member, IEEE, Robby T. Tan, Member, IEEE, Shiqi Wang<sup>2</sup>, Member, IEEE, Yuming Fang<sup>3</sup>, Senior Member, IEEE, and Jiaying Liu<sup>4</sup>, Senior Member, IEEE

**Abstract**—The goal of single-image deraining is to restore the rain-free background scenes of an image degraded by rain streaks and rain accumulation. The early single-image deraining methods employ a cost function, where various priors are developed to represent the properties of rain and background layers. Since 2017, single-image deraining methods step into a deep-learning era, and exploit various types of networks, i.e., convolutional neural networks, recurrent neural networks, generative adversarial networks, etc., demonstrating impressive performance. Given the current rapid development, in this paper, we provide a comprehensive survey of deraining methods over the last decade. We summarize the rain appearance models, and discuss two categories of deraining approaches: model-based and data-driven approaches. For the former, we organize the literature based on their basic models and priors. For the latter, we discuss the developed ideas related to architectures, constraints, loss functions, and training datasets. We present milestones of single-image deraining methods, review a broad selection of previous works in different categories, and provide insights on the historical development route from the model-based to data-driven methods. We also summarize performance comparisons quantitatively and qualitatively. Beyond discussing the technicality of deraining methods, we also discuss the future possible directions.

**Index Terms**—Rain streak removal, single image, model-based, data-driven

## 1 INTRODUCTION

RAIN introduces visual degradations to captured images and videos. *Rain streaks* particularly in heavy rain can cause severe occlusion on the background scene. *Rain accumulation* [1], where distant rain streaks cannot be seen individually and together with water particles form a layer of veil on the background, significantly degrades the contrast of the scene and reduce the visibility. Fig. 1 shows examples of degradation due to rain streaks and rain accumulation. Human vision and many computer vision algorithms suffer from this degradation, since most of these algorithms assume clear weather, with no interference of rain streaks and rain accumulation. Hence, restoring images from rain, called deraining or rain removal, is much desired in many practical applications.

An early study of video deraining was started in 2004 by Garg and Nayar [3]. They analyze rain dynamic appearances, and develop an approach to remove rain streaks from

videos. Kang *et al.* [4] was a pioneer in the single image deraining by publishing a method in 2012. The method extracts the high-frequency layer of a rain image, and decomposes the layer further into rain and non-rain components using dictionary learning and sparse coding. Starting from 2017, by the publications of [1], [18], data-driven deep-learning methods that learn features automatically become dominant in the literature.

In this survey, we focus on single-image deraining, which the aim is to estimate the rain-free background layer of an image degraded by rain streaks and rain accumulation. Unlike video deraining methods, which leverage temporal redundancy and dynamics of rain, single image deraining methods exploit the spatial information of neighboring pixels and the visual properties of rain and background scenes.

The milestones of single-image deraining in the past years are presented in Fig. 2. Before 2017, the typical methods are model-based approach (or non-deep learning approach). The major developments in the model-based approach are driven by the following ideas: image decomposition (2012), sparse coding (2015), and priors based Gaussian mixture models (2016). Since 2017, single-image deraining methods enter into a period of data-driven approach (or deep learning approach). The major developments in the data-driven approach are indicated by the following ideas: deep convolutional network (2017), generative adversarial network (2019), and semi/unsupervised methods (2019). In 2017-2019, there are more than 30 deraining papers based on deep learning, significantly more than the number of deraining papers before 2017.

Model-based methods rely more on the statistical analysis of rain streaks and background scenes. The methods

- W. Yang and S. Wang are with the Department of Computer Science, City University of Hong Kong, Hong Kong. E-mail: {wyang34, shiqi.wang}@cityu.edu.hk.
- R.T. Tan is with the Yale-NUS College, Singapore 138527, and also with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119077. E-mail: tanrobby@gmail.com.
- Y. Fang is with the School of Information Management, Jiangxi University of Finance and Economics, Jiangxi 330013, China. E-mail: fa0001ng@e.ntu.edu.sg.
- J. Liu is with the Wangxuan Institute of Computer Technology, Peking University, Beijing 100871, China. E-mail: liujiaying@pku.edu.cn.

Manuscript received 27 Dec. 2019; accepted 5 May 2020. Date of publication 19 May 2020; date of current version 1 Oct. 2021.

(Corresponding author: Jiaying Liu.)

Recommended for acceptance by P. Favaro.

Digital Object Identifier no. 10.1109/TPAMI.2020.2995190

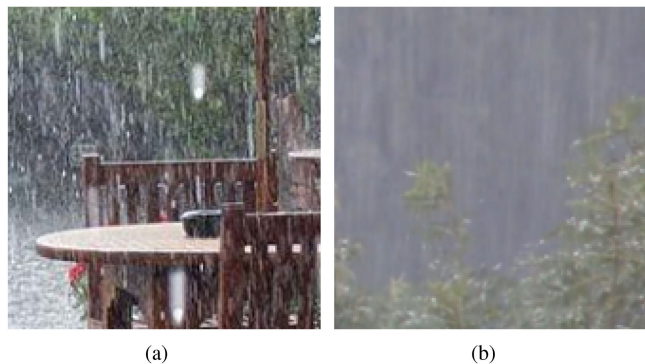


Fig. 1. Different types of visibility degradation caused by rain. (a) *Rain streaks* cause severe occlusion on the background scene. (b) *Rain accumulation* significantly degrades the contrast of the scene and reduce the visibility.

enforce handcrafted priors on both rain and background layers, then build a cost function and optimize it. The priors are extracted from various ways: Luo *et al.* [5] learn dictionaries for both rain streak and background layers, Li *et al.* [6] build Gaussian mixture models from clean images to model background scenes, and from rain patches of the input image to model rain streaks, Zhu *et al.* [7] enforce a certain rain direction based on rain-dominated regions so that the background textures can be differentiated from rain streaks.

In recent years, the popularity of data-driven methods has overtaken model-based methods. These methods exploits deep networks to automatically extract hierarchical features, enabling them to model more complicated mappings from rain images to clean ones. Some rain-related constraints are usually injected into the networks to learn more effective features, such as rain masks [1], background features [8], etc. Architecture wise, some methods utilize recurrent network [1], or recursive network [9] to remove rain

progressively. There are also a series of works focusing on the hierarchical information of deep features, e.g., [10], [11].

While deep networks lead to a rapid progress in deraining performance, many of these deep-learning deraining methods train the networks in a fully supervised way. This can cause a problem, since to obtain paired images of rain and rain-free images is intractable. The simplest solution is to utilize synthetic images. Yet, there are domain gaps between synthetic rain and real rain images, which can make the deraining performance not optimum. To overcome the problem, unsupervised/semi-supervised methods that exploit real rain images [12] and [13] are introduced.

Our paper aims to provide a comprehensive survey on single-image deraining methods. We believe it can provide a useful starting point to understand the main development of the field, the limitations of existing methods, and the future possible directions. The rest of the paper is organized as follows. Section 2 introduces the rain appearance model. Section 3 provides a detailed survey of single-image rain removal methods, including their synthetic rain models, deraining challenges, methods architectures, and the related technical development. A particular emphasis is placed on the deep-learning based methods as they offer the most significant progress in the recent years. Subsequently, Section 4 gives detailed discussion on technical developments of network architectures, basic blocks, and summaries of loss functions and databases. Section 5 summarizes the quantitative comparisons of a number of single-image rain removal methods and shows qualitative comparisons. Finally, the paper is concluded in Section 6.

## 2 RAINDROP APPEARANCE MODELS

The shape of a raindrop is usually approximated by a spherical shape [14]. As shown in Fig. 3, considering a point B on the surface of the raindrop with a surface normal  $\hat{n}$ , rays of

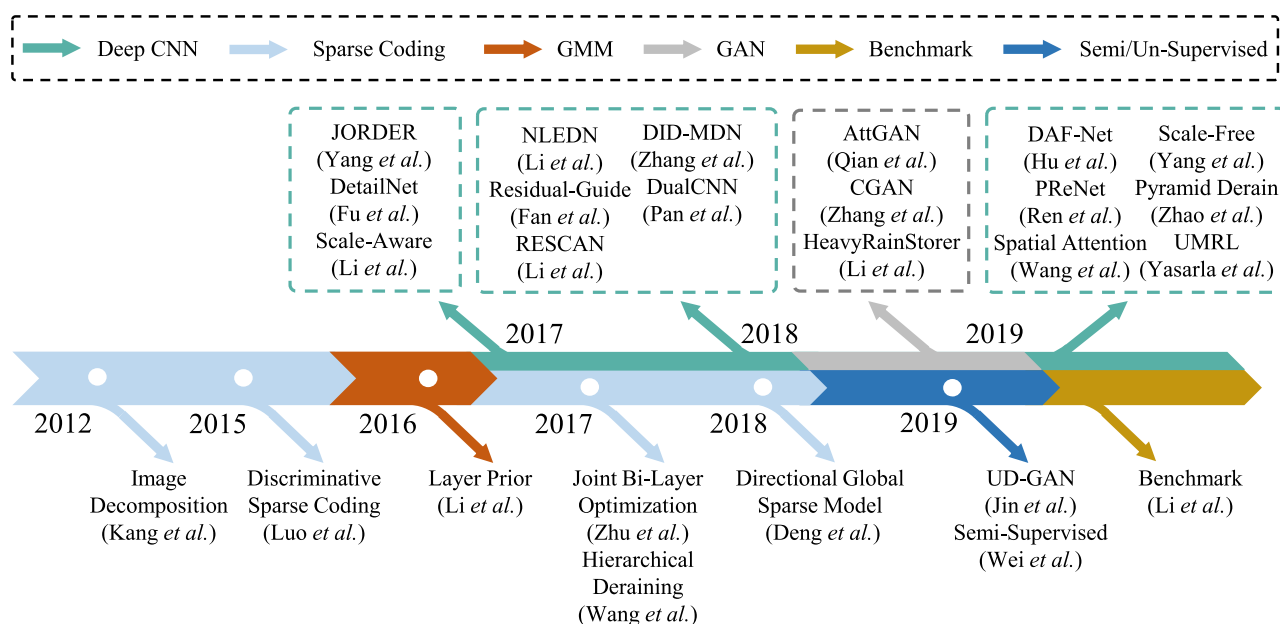


Fig. 2. Milestones of single image deraining methods: image decomposition, sparse coding, Gaussian mixture models, deep convolutional network, generative adversarial network, and semi/unsupervised learning. Before 2017, the typical methods are model-based approach (or non-deep learning approach). Since 2017, single-image deraining methods enter into a period of data-driven approach (or deep learning approach).

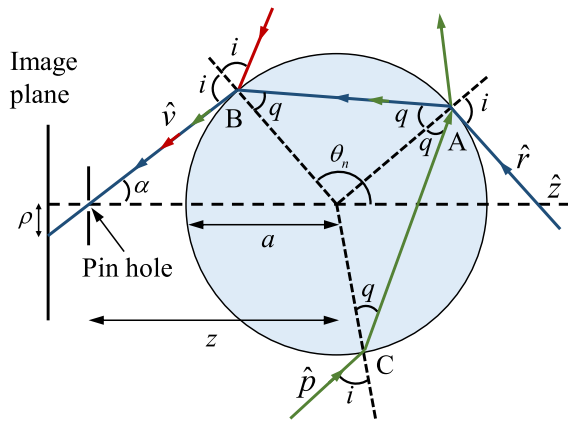


Fig. 3. A raindrop's appearance [14] is a complex mapping of the environmental radiance, which is determined by reflection, refraction, and internal reflection.

light ( $\hat{r}$ ,  $\hat{s}$  and  $\hat{p}$ ) are directed toward the observer via refraction, specular reflection, and internal reflection, respectively. Hence, the radiance  $L(\hat{n})$  at point B is approximated as the sum of the radiance  $L_r$  of refracted ray, radiance  $L_s$  of specularly reflected ray and radiance  $L_p$  of internally reflected ray

$$L(\hat{n}) = L_r(\hat{n}) + L_s(\hat{n}) + L_p(\hat{n}). \quad (1)$$

Considering that the radiances depend on the environmental radiance  $L_e$  in the direction of the reflected or refracted ray, Eq. (1) can be expressed as

$$L(\hat{n}) = RL_e(\hat{r}) + SL_e(\hat{s}) + PL_e(\hat{p}), \quad (2)$$

where  $R$ ,  $S$  and  $P$  denote the fractions of incident environmental radiance that reaches the camera after refraction, reflection and internal reflection, respectively. We refer to these fractions ( $R, S, P$ ) as radiance transfer functions.

Moreover, we can reach the composite raindrop model written as [14]

$$L(\hat{n}) = (1 - k(i, \mu))^2 L_e(\hat{r}) + SL_e(\hat{s}) + PL_e(\hat{p}), \quad (3)$$

where  $i = (\pi - \theta_n + \alpha)$  is the incident angle,  $\mu$  is the refractive index of the water and  $k$  is the Fresnel's reflectivity coefficient for unpolarized light. Based on the statistics from [14], the radiance of the raindrop is mainly decided by the refraction, and the appearance of the raindrop is mainly based on refraction through the drop.

For a moving raindrop, its appearance changes significantly. The raindrop becomes a rain streak, and its appearance relies on the brightness of the raindrop, background scene radiances, and camera's exposure time. The change of pixel's intensity value caused by a rain streak can be approximated as [14]

$$\Delta I = -\beta I_b + \alpha, \quad (4)$$

with  $\beta = \frac{\tau}{T}$ , and  $\alpha = \tau \bar{E}_r$ , where  $\tau$  is the time when a drop remains within a pixel, and  $T$  is the exposure time.  $\bar{E}_r$  is the time-averaged irradiance caused by the drop. Based on Eq. (4), we can reach two conclusions: 1) A raindrop causes an intensity change and moves faster than the integration time of a typical video camera; 2) the intensity change of a rain streak correlates linearly to the background intensity  $I_b$ .

Based on the derived numerical bounds [14], empirically we obtain:  $0 < \beta < 0.039$  and  $0 < \tau < 1.18$ . In most real cases,  $\alpha$  dominates the appearance of  $\Delta I$ , thus

$$\Delta I = \alpha. \quad (5)$$

As a result, in most rain synthetic models, *rain streaks are assumed to be superimposed on the background image.*

### 3 LITERATURE SURVEY

In this section, we first review a few rain synthesis models proposed in some existing methods. Unlike in the previous section (Section 2), the models we discuss here are only loosely based on physics and thus, to our knowledge, their correctness has not been verified both theoretically or experimentally. Despite this, the methods that use these models show, to some extent, the effectiveness of the models on real-image deraining. Having discussed various rain synthetic models, we briefly explain the challenges in image deraining, and subsequently survey on existing deraining methods comprehensively.

#### 3.1 Synthetic Rain Models

*Additive Composite Model.* The most simple and popular rain model used in existing studies is the additive composite model [4], [6], which follows Eq. (5) and is expressed as

$$\mathbf{O} = \mathbf{B} + \mathbf{S}, \quad (6)$$

where  $\mathbf{B}$  denotes the background layer, and  $\mathbf{S}$  is the rain streak layer.  $\mathbf{O}$  is the image degraded by rain streaks. Here, the model assumes that the appearance of rain streaks is simply superimposed to the background, and there is no rain accumulation in the rain degraded image.

*Screen Blend Model.* Luo *et al.* [5] propose a non-linear composite model, called screen-blend model

$$\mathbf{O} = 1 - (1 - \mathbf{B}) \circ (1 - \mathbf{S}) = \mathbf{B} + \mathbf{S} - \mathbf{B} \circ \mathbf{S}, \quad (7)$$

where  $\circ$  denotes the operation of point-wise multiplication. Unlike the additive composite model in Eq. (6), the background and rain layers influence the appearance of each other. Luo *et al.* [5] claim that the screen blend model can model some visual properties of real rain images, such as the effect of internal reflections, and thus generate visually more authentic rain images. The combination of rain and background layers are signal-dependent. Implying, when the background is dim, the rain layer will dominate the appearance of the rain image; and, when the background is bright, the background layer will dominate the image.

*Heavy Rain Model.* Yang *et al.* [1] propose a rain model that includes both rain streaks and rain accumulation. This is the first model in the deraining literature that includes the two rain phenomena. Rain accumulation or rain veiling effect is a result of water particles in the atmosphere and distant rain-streaks that cannot be seen individually. The visual effect of rain accumulation is similar to mist or fog, which leads to low contrast. Considering two main aspects of rain: the Koschmieder model to approximate the visual appearance of a scene in a turbid medium, and overlapping rain streaks that have different directions and shapes, a novel rain model is introduced

TABLE 1  
Summary of Rain Synthetic Models in the Literature

Method	Degradation Factors	Main Features	Publication
Additive Composite Model (ACM)	Streak	Simple and effective	Li <i>et al.</i> 2016 [6]
Screen Blend Model (SBM)	Streak	Streaks and backgrounds are combined nonlinearly	Luo <i>et al.</i> 2015 [5]
Heavy Rain Model (HRM)	Streak, Accumulation	Overlapping streaks generating accumulation	Yang <i>et al.</i> 2017 [1]
Rain Model with Occlusion (ROM)	Streak, Occlusion	Considering rain occlusions	Liu <i>et al.</i> 2018 [15]
Comprehensive Rain Removal (CRM)	Streak, Occlusion, Accumulation, Flow	Considering comprehensive visual degradation	Yang <i>et al.</i> 2019 [2]
Depth-Aware Rain Model (DARM)	Streak, Accumulation	Streaks and accumulation modeling correlated with depth	Hu <i>et al.</i> 2019 [16]

$$\mathbf{O} = \alpha \circ \left( \mathbf{B} + \sum_{t=1}^s \mathbf{S}_t \right) + (1 - \alpha)\mathbf{A}, \quad (8)$$

where  $\mathbf{S}_t$  denotes the rain-streak layer that has the same streak direction.  $t$  indexes the rain-streak layer and  $s$  is the maximum number of the rain-streak layers.  $\mathbf{A}$  is the global atmospheric light, and  $\alpha$  is the atmospheric transmission.

*Rain Model With Occlusion.* Liu *et al.* [15] extend the heavy rain model to an occlusion-aware rain model for modeling rain in video. The model separates rain streaks into two types: transparent rain streaks that are added to the background layers, and opaque rain streaks that totally occlude the background layers. The locations of these opaque rain streaks are indicated by a map, called the reliance map. The formulation of this rain model is expressed as

$$\mathbf{O} = \beta \circ \left( \mathbf{B} + \sum_{t=1}^s \mathbf{S}_t \right) + (1 - \beta) \circ \mathbf{R}, \quad (9)$$

where  $\mathbf{R}$  is the rain reliance map and  $\beta$  is defined as

$$\beta_t(i, j) = \begin{cases} 1, & \text{if } (i, j) \in \Omega_S, \\ 0, & \text{if } (i, j) \notin \Omega_S, \end{cases} \quad (10)$$

where  $\Omega_S$  is defined as the rain occluded region.

*Comprehensive Rain Model.* Yang *et al.* [2] combine all above mentioned degradation factors into a comprehensive rain model for modeling rain appearance in video. It considers the temporal properties of rain scenes, particularly the fast-changing rain accumulations that usually cause flicker. This visible intensity changes along the temporal dimension is called rain accumulation flow. Besides, it also considers other factors including rain streaks, rain accumulation, and rain occlusion, which are formulated as

$$\mathbf{O} = \beta \circ \left[ \left( \mathbf{B} + \sum_{t=1}^s \mathbf{S}_t \right) + (1 - \alpha)\mathbf{A} + \mathbf{U} \right] + (1 - \beta) \circ \mathbf{R}, \quad (11)$$

where  $\mathbf{U}$  is the rain accumulation flow.

*Depth-Aware Rain Model.* Hu *et al.* [16] further connect  $\alpha$  to the scene depth  $d$ , to create a depth-aware rain model

$$\mathbf{O} = \beta \circ \left( 1 - \sum_{t=1}^s \mathbf{S}_t - (1 - \alpha)\mathbf{A} \right) + \sum_{t=1}^s \mathbf{S}_t + \alpha\mathbf{A}, \quad (12)$$

where  $\mathbf{S}_t$  and  $\alpha$  are connected with the scene depth written as

$$\begin{aligned} \mathbf{S}_t(i, j) &= \mathbf{S}_{\text{Pattern}}(i, j) \cdot t_r(i, j), \\ t_r(i, j) &= e^{-\alpha \max(d_M, d(i, j))}, \\ A(i, j) &= 1 - e^{-\beta d(i, j)}, \end{aligned} \quad (13)$$

where  $\mathbf{S}_{\text{Pattern}}(i, j)$  is an intensity image of uniformly-distributed rain streaks in the image space, and  $t_r(i, j)$  is the rain streak intensity map relying on the depth.  $d(i, j)$  denotes the depth and  $\alpha$  controls the rain streak intensity.  $\beta$  determines the thickness of fog, where a larger  $\beta$  denotes a thicker fog.

*Discussions.* Following these different rain models, various rain degradation can be synthetically rendered. A summary of rain synthetic models in the literature is provided in Table 1. In general, heavy rain models [1], [13] and depth-aware rain models [16] cover the most comprehensive rain degradation for single rain image synthesis. However, as we mentioned in the beginning of this section that all these models are heuristic; implying that they might not entirely correct physically. Nevertheless, as shown in the literature, they can be effective, at least to some extent, for image deraining.

### 3.2 Deraining Challenges

The goal of single image deraining is to recover the clean and rain-free background scene from a rain degraded image. However, there are a few challenges to accomplish the goal:

- *Difficulties in modeling rain images* In the real world, rain can visually appear in many different ways. Rain streaks can vary in terms of sizes, shapes, scales, densities, directions, etc. Similarly, rain accumulation depends on various water-particles and atmospheric conditions. Moreover, rain appearance significantly relies also on the textures and the depth of the background scenes. All these cause difficulties in modeling the appearance of rain, which consequently cause the rendering of physically-correct rain images to be a complex task.
- *Ill-posedness of deraining problem* Even with a simple rain model that considers only rain-streaks, to estimate the background scene from a degraded image is an ill-posed problem. The reason is that we only have the pixel intensity values produced by lights carrying fused information of rain and background scenes. To make the matter worse, in some cases the background information can be totally occluded by rain streaks or dense rain accumulation or both.
- *Difficulties in finding proper priors* As rain and background information might overlap in the feature

space, it is non-trivial to separate them. Background textures can be falsely deemed as rain, resulting in incorrect deraining. Hence, strong priors for background textures and rain are necessary. However, finding these priors is difficult, since background textures are diverse, and some have similarity to the appearance of rain-streaks or rain accumulation.

- *Real paired ground-truths* Most of deep-learning methods rely on paired rain and clean background images to train their networks. However, to obtain real rain images and their exact pairs of clean background images is intractable. Even for a static background, lighting conditions always change. This difficulty does not only impact on deep-learning methods, but also for evaluating the effectiveness of any method. Currently, for qualitative evaluation, all methods rely on human subjective judgment on whether the restored images are good; and for quantitative evaluation, all current methods rely on synthetic images. Unfortunately, up to now, there are significant gaps between synthetic and real images.

In the following section, we will discuss how existing deraining methods deal with these challenges.

### 3.3 Single-Image Deraining Methods

We categorize single-image deraining methods into two basic approaches: *model-based* (non-deep-learning) and *data-driven* (deep-learning) approaches. We will discuss the existing methods of the two approaches in detail in the subsequent sections. A summary of previous works is given in Tables 2 and 3. In addition, for the sake of completeness, we will also briefly discuss adherent raindrop removal methods, since adherent raindrops (i.e., water droplets attached to a lens or windscreen) are also part of rain degradation; although, in some situations, they can be avoided by placing the camera under a shelter.

#### 3.3.1 Model-Based Methods

Existing model-based methods employ optimization frameworks for deraining, as shown in the top panel of Fig. 4. These methods deal only with rain streaks and ignore the presence of rain accumulation. A general optimization framework can be expressed as

$$\hat{\mathbf{B}} = \arg \min_{\mathbf{B}} \|\mathbf{I} - \mathbf{B} - \mathbf{S}\|_2^2 + \Upsilon(\mathbf{B}) + \Psi(\mathbf{S}) + \Omega(\mathbf{B}, \mathbf{S}), \quad (14)$$

where  $\Upsilon(\mathbf{B})$  denotes the priors on the background layers,  $\Psi(\mathbf{S})$  represents the priors on rain streak layers, and  $\Omega(\mathbf{B}, \mathbf{S})$  is the joint prior to describe the intrinsic relationship between rain streaks and background layers. Different prior terms are designed to better describe and separate the rain streak from the background layers.

*Sparse Coding Methods.* Sparse coding [56] represents the input vectors as a sparse linear combination of basis vectors. The collection of these basis vectors is called dictionary, which is used to reconstruct the certain type of signals, e.g., rain streaks and background signals in the deraining problem. Lin *et al.* [4] make the first attempt on single-image deraining via *image decomposition* using a morphological

component analysis. The initially extracted high-frequency component of a rain image is further decomposed into rain and non-rain components by dictionary learning and sparse coding. This pioneer work successfully removes sparse light rain streaks. However, it significantly relies on the bilateral filter preprocessing, and thus generates blurred background details.

In a successive work, Luo *et al.* [5] enforce the sparsity of rain, and introduce a mutual exclusivity property into a discriminative sparse coding (DSC) to facilitate accurately the separation of the rain/background layers from their non-linear composite. Benefiting from the mutual exclusivity property, the DSC preserves clean texture details; however, it shows some residual rain streaks in the output, particularly for large and dense rain streaks. Zhanget *al.* [60] construct a model to learn a set of low-rank representation-based and sparsity-based convolutional filters to efficiently represent background rain-free images and rain streaks, respectively. To further improve the modeling capacity, Zhu *et al.* [7] construct an iterative layer separation process to remove rain streaks from the background layer, as well as to remove background's texture details from the rain streak layer using layer-specific priors. Quantitatively, the method obtains comparable performance on some synthetic datasets with that of deep learning-based methods published in the same period of time, i.e., JORDER [1] and DDN [18]. However, qualitatively on real images, the method tends to fail in handling heavy rain cases, where rain streaks may move in different directions.

To model rain streak directions and sparsity, Deng *et al.* [17] formulate a directional group sparse model (DGSM), which includes three sparse terms representing the intrinsic directional and structural knowledge of rain streaks. It can effectively remove blurred rain streaks but fail to remove sharp rain streaks.

*Gaussian Mixture Model.* Li *et al.* [6] apply Gaussian mixture models (GMMs) to model both rain and background layers. The GMMs of the background layer is obtained offline from real images with diverse background scenes. A selected rain patch from the input image that has no background textures is proposed to train the GMMs of the rain layer. The total variation is utilized to remove small sparkle rains. The method is capable to effectively remove rain streaks of small and moderate scales, but fail to handle large and sharp rain streaks.

#### 3.3.2 Deep Learning Based Methods

*Deep CNNs.* The era of deep-learning deraining starts in year 2017. Yang *et al.* [1] construct a joint rain detection and removal network. It can handle heavy rain, overlapping rain streaks, and rain accumulation. The network can detect rain locations by predicting the binary rain mask, and take a recurrent framework to remove rain streaks and clear up rain accumulation progressively. The method achieves good results in heavy rain cases. However, it might falsely remove vertical textures and generate underexposed illumination.

In the same year, Fu *et al.* [18], [19] made an attempt to remove rain streaks via a deep detail network (DetailNet). The network takes only the high frequency details as input, and predicts the residue of the rain and clean images. The

TABLE 2  
An Overview of Single-Image Rain Removal Methods (Part1)

Method	Category	Rain Model	Variables or Priors	Key Idea	Publication
Image Decomposition	Sparse Representation	ACM	Low/high-frequency; Rain/No-rain dictionaries.	Rain images are decomposed into the low/high-frequency (HF) parts. Then, the "nonrain component" is removed from HF part by dictionary learning and sparse coding. details.	Kang <i>et al.</i> 2012 [4]
DSC		SBM	Rain/No-rain codes; Rain/No-rain dictionaries.	Patches of two layers are sparsely approximated by very high discriminative codes over a learned dictionary with strong mutual exclusivity property.	Luo <i>et al.</i> 2015 [5]
Bi-Layer Optimization		ACM	Centralized sparse prior; Rain direction prior; Rain layer prior.	A joint optimization process is used that alternates between removing rain-streak details from the estimated background and removing non-streak details from the estimated rain streak layer.	Zhu <i>et al.</i> 2017 [7]
CCRR		ACM	Convolutional sparse coding; Convolutional low rank coding.	The method learns a set of low-rank representation-based and sparsity-based convolutional filters to efficiently represent background rain-free images and rain streaks, respectively.	Zhang <i>et al.</i> 2017 [60]
Hierarchical Deraining		ACM	High/low-frequency; Sensitivity of variance across color channels; Principal direction of an image patch.	A three-layer hierarchical scheme is designed. The first layer uses sparse coding to classify the high-frequency part into rain/snow and non-rain/snow components. The second layer relies on guided filtering. The third layer enhances the visual quality with the sensitivity of variance across color channels.	Wang <i>et al.</i> 2017 [32]
DGSM		ACM	Unidirectional TV; Rotation Angle.	A global sparse model is formulated to consider the intrinsic directional, structural knowledge of rain streaks, and the property of image back-ground information.	Deng <i>et al.</i> 2018 [17]
CNN		Raindrop	Clean image.	A three-layer CNN is constructed to learn the mapping between the corrupted image patches to clean ones.	Eigen <i>et al.</i> 2013 [27]
JORDER JORDER-E		HRM	Binary rain mask; Rain intensity; Residual.	A multi-task architecture is constructed to learn the binary rain streak map, the appearance of rain streaks, and the clean background. A recurrent network is also built to remove rain streaks and clears up the rain accumulation progressively.	Yang <i>et al.</i> 2017 [11] Yang <i>et al.</i> 2019 [33]
DetailNet		ACM	Residual.	A deep detail network taking as input the high frequency detail and predict the residual between the input image and the ground truth image.	Fu <i>et al.</i> 2017 [18] Fu <i>et al.</i> 2017 [19]
Scale-Aware		HRM	Rain streak; Transmission map.	Parallel sub-networks are built to predict different scales of rain streaks and the transmission maps.	Li <i>et al.</i> 2017 [20]
NLEDN	ACM	Residual.	An auto-encoder network is built with non-locally enhanced dense blocks, where a non-local feature map weighting follows four densely connected convolution layers.	Li <i>et al.</i> 2018 [10]	
Residual-Guide	ACM	Residual.	The residuals generated from shallower blocks are used to guide deeper blocks. The negative residual is predicted coarse to fine and the outputs of different blocks are fused finally.	Fan <i>et al.</i> 2018 [8]	
RESCAN	ACM	Residual; Intermediate results.	A recurrent network is constructed and the rain removal result of the previous stage is fed into the next stage. The information is flowed across stages at the feature level.	Li <i>et al.</i> 2018 [21]	
DID-MDN	ACM	Rain density; Residual.	A multi-path densely connected network is constructed to automatically detect the rain-density to guide the rain removal.	Zhang <i>et al.</i> 2018 [22]	
DualCNN	ACM	Structure layer; Detail layer.	A dual CNN estimates the two parts of the target signal: structures and details for a series of low-level vision tasks.	Pan <i>et al.</i> 2018 [23]	
DAF-Net	DARM	Depth; Attention; Residual.	The work creates a <i>RainCityscapes</i> dataset related to the scene depth. A deep network is developed based on a depth-guided attention mechanism to predict the residual map.	Hu <i>et al.</i> 2019 [16]	
Spatial Attention + Dataset	ACM	Spatial attention; Residual.	A high-quality dataset including 29.5K paired real rain and pseudo real non-rain images is constructed. Then, a novel spatial attentive network is built to effectively learn discriminative features for rain removal from local to global.	Wang <i>et al.</i> 2019 [34]	
PReNet	ACM	Residual; Intermediate results.	The network utilizes recursive computation at two levels: 1) Progressive ResNet is built by repeatedly unfolding a shallow ResNet; 2) The ResNet performs stage-wise operations processing the input and intermediate results progressively.	Ren <i>et al.</i> 2019 [9]	
Scale-Free	HRM	Residual; Wavelet decomposition results.	A recurrent network performs two-stage deraining: 1) rain removal on the low-frequency component; 2) recurrent detail recovery on high-frequency components guided by the recovered low-frequency component.	Yang <i>et al.</i> 2019 [30]	
PyramidDerain	ACM	Gaussian-Laplacian pyramid.	The proposed network combines Gaussian Laplacian image pyramid decomposition and the deep neural network. Recursive and residual network structures are employed to aggregate the features at different layers.	Fu <i>et al.</i> 2019 [11]	
UMRL	ACM	Confident map.	The network is guided based on the confidence measure about the estimate. The cycle spinning is introduced to remove artifacts and improve the deraining performance.	Yasarla <i>et al.</i> 2019 [29]	
AttGAN	Raindrop	Attention map; Clean image.	Visual attention is injected into both the generative and discriminative networks for learning to attend raindrop regions and percept their surroundings.	Qian <i>et al.</i> 2018 [28]	
CGAN	ACM	Clean image.	The work directly applies a multi-scale conditional generative adversarial network to address single image de-raining task.	Zhang <i>et al.</i> 2019 [24]	
HeavyRainRestorer	HRM	Transmission map; Atmospheric light; Rain streak; Clean image.	A two-stage network is built: a physics-driven model followed by a depth-guided generative adversarial refinement.	Li <i>et al.</i> 2019 [13]	

TABLE 3  
An Overview of Single-Image Rain Removal Methods (Part2)

Method	Category	Rain Model	Variables or Priors	Key Idea	Publication
LP	GMM	ACM	Gaussian mixture model; Total variation.	Two patch-based priors for the background and rain layers are built based on Gaussian mixture models and can accommodate multiple orientations and scales of rain streaks.	Li <i>et al.</i> 2016 [6]
Semi-supervised CNN	Semi/Un-Supervised	ACM	Rain streak; Residual.	A semi-supervised learning method formulates the residual as a specific parametrized rain streak distribution between an input rainy image and its expected network output.	Wei <i>et al.</i> 2019 [25]
UD-GAN		ACM	Rain streak; Residual.	An Unsupervised Deraining Generative Adversarial Network is built to introduce self-supervised constraints, the intrinsic priors extracted from unpaired rainy and clean images.	Jin <i>et al.</i> 2019 [12]
Benchmark	Benchmark + Application	ACM + Raindrop	-	It provides extensive study and evaluation of existing single image deraining algorithms with a new proposed large-scale dataset including both synthetic and real-world rainy images of various rain types.	Li <i>et al.</i> 2019 [26]
RobustFlow		HRM	Residual channel; Rain accumulation; Total variation.	The accuracy of the optical flow estimation under rainy conditions is improved via residual channel and optimization with layer-wise priors.	Li <i>et al.</i> 2018 [63]
RainFlow		HRM	Veiling-invariant features; Streak-invariant feature.	A deep learning pipeline is built to extract veiling-invariant features and establish a new mapping operation to remove the effect of rain-streak patterns at the feature level.	Li <i>et al.</i> 2019 [64]
PhysicsRender		HRM	Photometry of rain streak; Rain streak composition; Rainfall composition.	A physical-based rain rendering pipeline is built to synthesize rain scenes to evaluate the performance of deep object detection and semantic segmentation algorithms in degraded rainy and haze weather.	Halder <i>et al.</i> 2019 [65]

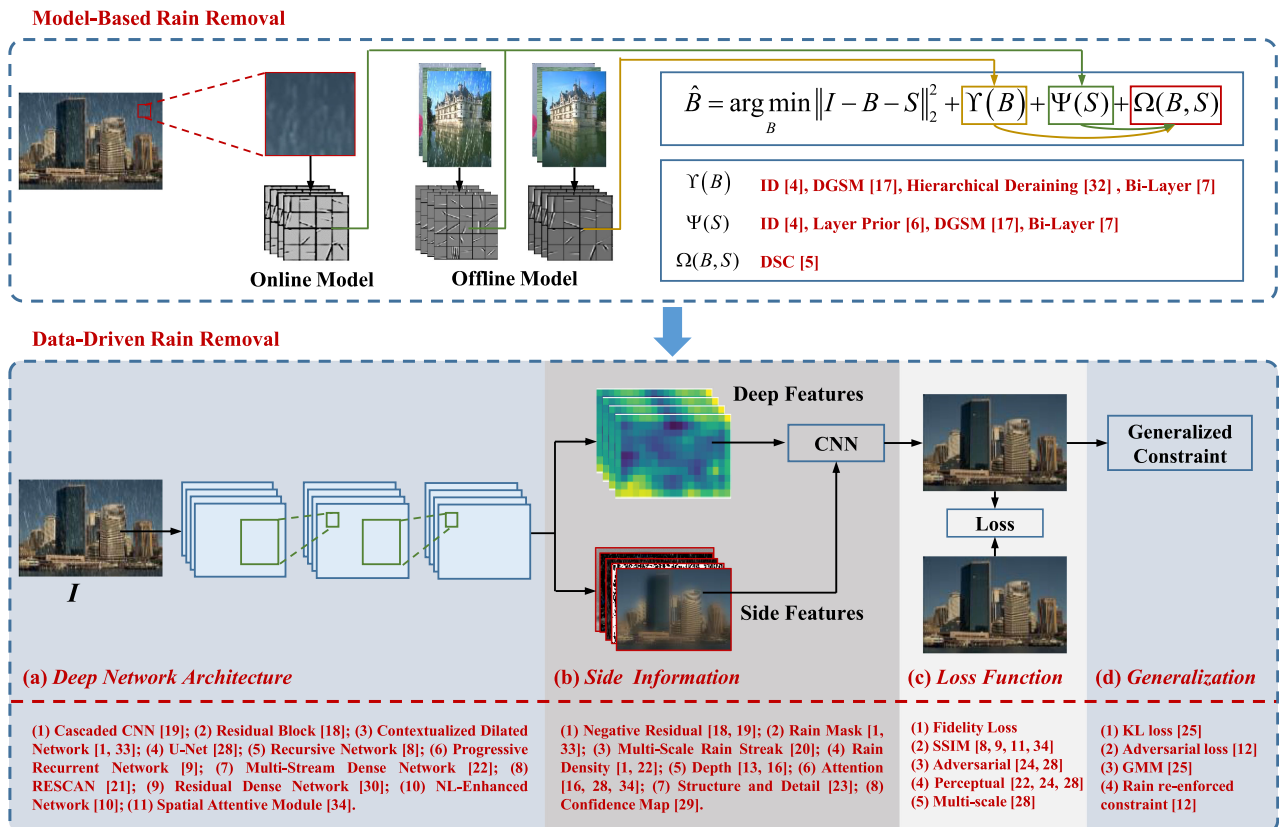


Fig. 4. The improvement of single-image rain removal, from model-based to data-driven approaches. The model-based methods employ optimization frameworks for deraining. They rely on the statistical analysis of rain streaks and background scenes, and enforce handcrafted priors on both rain and background layers. Data-driven approaches exploit deep networks to automatically extract hierarchical features, enabling them to model more complicated mappings from rain images to clean ones. Some rain-related constraints are injected into the networks to learn more effective features.

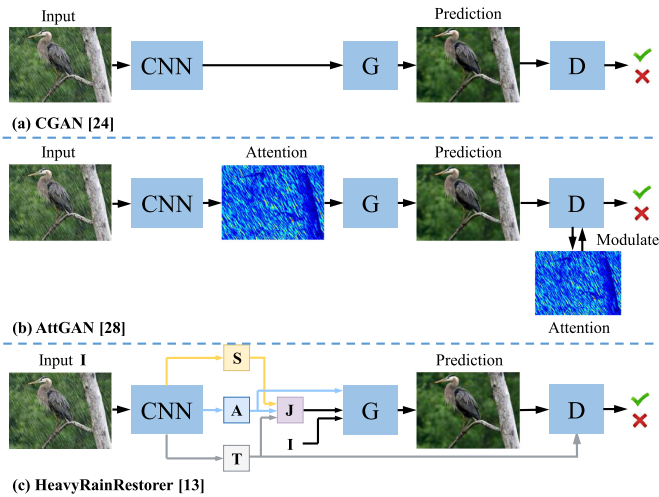


Fig. 5. Summary of GAN-based rain removal methods. To capture some visual properties of rain that cannot be modeled and synthesized, the adversarial learning is introduced to reduce the domain gaps between the generated results and real clean images.

paper shows that removing the background information in the network input is beneficial, as doing so makes the training easier and more stable. However, the method still cannot handle large and sharp rain streaks.

Following Yang *et al.* [1] and Fu *et al.* [18], [19], many CNN based methods [8], [10], [20], [21], [22], [23] are proposed. These methods employ more advanced network architectures and injecting new rain related priors. They achieve better results both quantitatively and qualitatively. However, due to the limitation of their fully supervised learning paradigm, namely using synthetic rain images, they tend to fail when dealing with some conditions of real rain that has never been seen during training.

*Generative Adversarial Networks.* To capture some visual properties of rain that cannot be modeled and synthesized, the adversarial learning is introduced to reduce the domain gaps between the generated results and real clean images. The typical network architecture consists of two parts: generator and discriminator, where the discriminator attempts to assess whether a generated result is real or fake, which provides an additional feedback to regularize the generator to produce more visually pleasing results. Zhang *et al.* [24] directly apply the conditional generative adversarial network (CGAN) for the single image rain removal task, as shown in Fig. 5a. CGAN is capable of capturing the visual properties beyond the signal fidelity, and presents results with better illumination, color and contrast distribution. However, CGAN sometimes might generate visual artifacts when the background of the testing rain image is different from those in the training set.

Li *et al.* [13] propose a single-image deraining method that combines the physics-driven network and adversarial learning refinement network, as shown in Fig. 5c. The first stage learns from the synthesized data and estimates physics-related components, i.e., rain streaks, the transmission, and the atmospheric light. At the second refinement stage, a depth-guided GAN is proposed to compensate for the lost details and to suppress the introduced artifacts at the first stage. Learning from real rain data, some visual properties of the results by these methods are significantly improved,

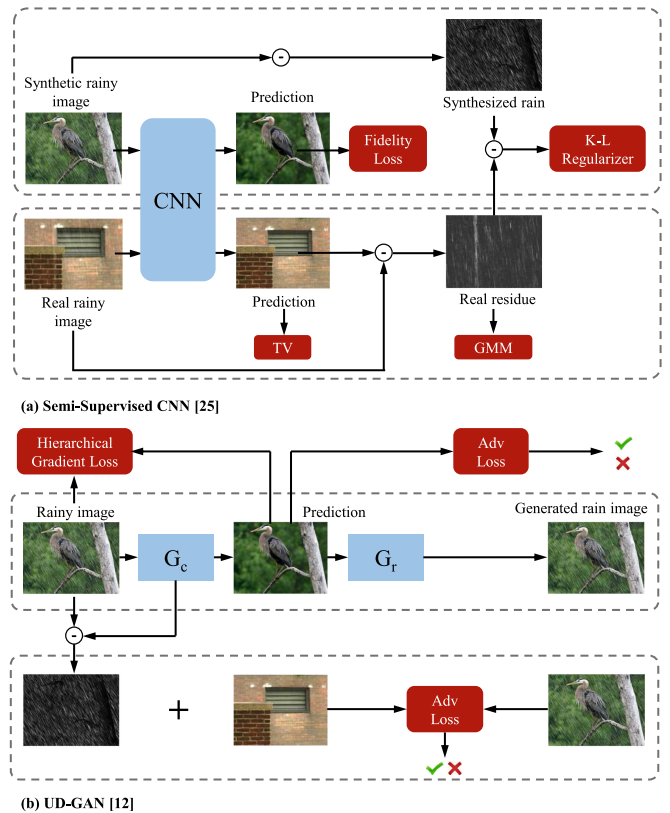


Fig. 6. Summary of semi/unsupervised learning-based rain removal methods. Semi/un-supervised learning methods make an attempt to improve the generality and scalability by learning directly from real rain data.

namely removing rain accumulation more thoroughly and achieving a more balanced luminance distribution. However, as GAN-based methods are not good at capturing fine-grained detail signals, the diversified appearances of real rain streaks are also not properly modeled in these methods.

*Semi/Unsupervised Learning Methods.* Recently, semi-supervised and unsupervised learning methods make an attempt to improve the generality and scalability by learning directly from real rain data. Wei *et al.* [25] propose a semi-supervised learning method to make use of the priors in both synthesized paired data and unpaired real data, as shown in Fig. 6a. In the proposed method, the residual is formulated as a specific parametrized rain streak distribution between an input rain image and its expected network output. The model trained on synthesized paired rain images is adapted to handle diversified rain in real scenarios with the guidance of the rain-streak distribution model. The method, however, does not show effective deraining results particularly for real rain images. This might be caused by the loose loss functions they impose to the network during the training process.

In [12], an unsupervised deraining generative adversarial network (UD-GAN) is proposed by introducing self-supervised constraints, and the intrinsic priors extracted from unpaired rain and clean images, as shown in Fig. 6b. Two collaborative modules are designed: One module is utilized to detect the difference between real rain images (real background images) and generated rain images (generated background images); while the other is introduced to adjust the luminance of the generated results, making the results more visually pleasing. The method is capable of removing real



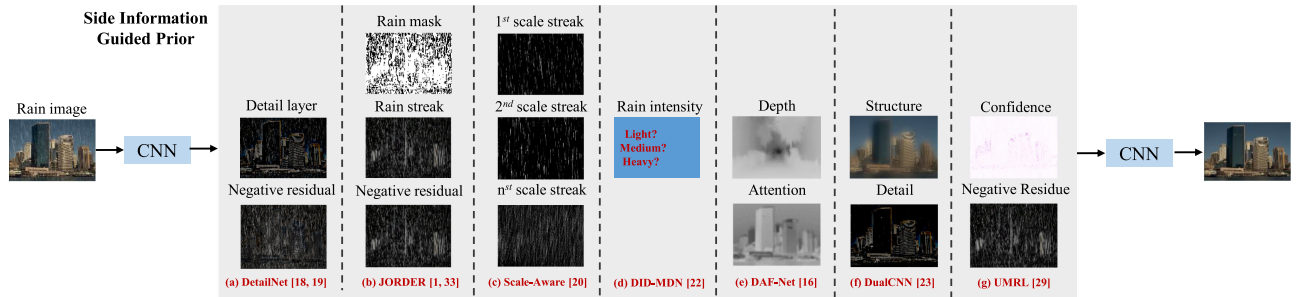


Fig. 7. Summary of the side information and priors for single-image rain removal. They injected into the networks to learn more effective features for deraining.

rain from rain images, yet inevitably losing some details, particularly when the rain streaks are dense.

**Benchmark and Applications.** Li *et al.* [26] provided extensive study and evaluation of existing single image deraining algorithms with a newly proposed large-scale dataset including both synthetic and real-world rain images with various rain types, i.e., rain streak, raindrops, and mist. The benchmark also includes a wide range of evaluation criteria including the results of different methods quantitatively and qualitatively. Comparatively, beyond extensive evaluations, our survey focuses more on discussions on technical developments of existing rain models, deraining methods from various perspectives.

There are also recent works focusing on rain-related applications, e.g., optical flow estimation in rainy conditions or object detection in rain videos. In [63], the accuracy of the optical flow estimation under rainy conditions is improved via residual channel and optimization with layer-wise priors. In [64], a deep learning pipeline is built to extract veiling-invariant features and establish a new mapping operation to remove the effect of rain-streak patterns at the feature level. In [65], a physics-based rain rendering pipeline is built to synthesize rain scenes to evaluate the performance of deep object detection and semantic segmentation algorithms in degraded rainy and hazy scenes.

### 3.3.3 Adherent Raindrop and Snow Removal

Raindrops adhered to the camera lens can severely degrade visibility of a background scene in an image. The goal of adherent raindrop removal is to detect and remove raindrops from an input image. Deraining is different from adherent raindrop removal, since rain images do not always suffer from adherent raindrop degradation, and vice versa: Adherent raindrop images do not always suffer from the degradation of rain streaks or rain accumulation. Nevertheless, we discuss it briefly here for the sake of the completeness of the survey.

In [57], Yamashita *et al.* develop a stereo system to detect and remove raindrops. Subsequently, a method [58] is built based on the image sequence instead of stereo video. You *et al.* [59] propose a motion based method to detect raindrops, and apply video completion to remove the detected regions. Eigen *et al.* [27] make the first attempt to tackle the problem of single-image raindrop removal. A three-layer CNN is trained with pairs of raindrop degraded images and the corresponding clean ones. It can handle relatively sparse and small raindrops as well as dirt, however, it fails to produce clean results for large and dense raindrops.

Recently, Qian *et al.* [28] develop an attentive GAN (AttGAN) [28] by injecting visual attention into both the generative and discriminative networks, as shown in Fig. 5b. The visual attention does not only guide the discriminative network to focus more on local consistency of the restored raindrop regions, but also make the generative network pay more attention to the contextual information surrounding the raindrop areas. Quanet *et al.* [62] introduce a double attention mechanism into deep CNN models. The double attention concurrently guides the CNN with shape-driven attention and channel recalibration. For the shape-driven attention, the physical shape priors of raindrops, i.e., convexness and contour closedness, help locate raindrops; while the channel recalibration improves the robustness of processing raindrops with varying appearances.

Similarly, snow can be regarded as small particles floating in the air, whose appearance depends on their particle sizes and shapes, as well as their distance to the camera. In [32], snow removal is also conducted via the same framework as rain removal. DesnowNet [61] makes the first efforts in the data-driven way. In this work, a multi-stage network is proposed to deal with the removal of translucent and opaque snow particles. The snow attributes of translucency and chromatic aberration are estimated. Furthermore, the residual complements of the snow-free images is estimated to restore details obscured by opaque snow.

## 4 TECHNICAL DEVELOPMENT REVIEW

In this section, we summarize the developments of existing deep-learning methods from the perspective of the network architectures, basic blocks, loss functions, and datasets. These aspects significantly influence the network’s learning capacity and thus determine the networks’ deraining performance.

### 4.1 Network Architectures

Since the publications of deep learning based deraining [1], [19], the successive methods aim to design more effective network architectures by relying on certain assumptions/constraints and general knowledge in image processing. In this section, we take a look at these developments.

**Deraining Assumptions/Constraints.** Specific networks dedicated to certain problems usually perform better than generic networks. To create these specific networks, some constraints or assumptions about the problems need to be injected. For deraining, these assumptions can relate to rain, background scenes, or other information. By incorporating some of these assumptions, a network is expected to learn

TABLE 4  
Summary of Side Information Used in Previous Works

Side information	Methods
Rain Mask	JORDER [1]
Rain Density	JORDER [1], DID-MDN [22]
Depth	DAF-Net [16], HeavyRainRestorer [13]
Attention	DAF-Net [16], SPA-Net [34], AttGAN [28]
Intermediate Results	JORDER-E [33], PReNet [9],
Bands Results	Scale-Free Rain Removal [30], PyramidNet [11]

the characteristics of rain better, and thus to separate rain layer from the background layer more robustly.

Fu *et al.* [19] assume a rain image can be decomposed into detail and base layers, where the detail layer contains the image textures and rain streaks (Fig. 7a), and the base layer mostly contains the background and the rain accumulation. The proposed network thus attempts to derain the the detail layer, before fusing the layer with the base layer to obtain the rain-free output.

Yang *et al.* [1] construct a joint rain detection and removal network as shown in Fig. 7b to detect rain locations, estimate rain densities and predict rain sequentially, which boosts the capacity of the network to process rain and non-rain regions differently. Li *et al.* [20] focus on the scale diversity of rain streaks. A scale-aware network as shown in Fig. 7c consisting of parallel subnetworks is built to make it aware of different scales of rain streaks, producing better deraining performance for real images. Zhang *et al.* [22] propose a density-aware rain removal method (DID-MDN) as shown in Fig. 7d to automatically detect the rain-density as the guidance information for the successive deraining.

Aside from assumptions related to rain, some assumptions are related to the background scenes. Hu *et al.* [16] analyze the complex visual effects in real rain and formulate a rain imaging model related to the scene depth. An end-to-end deep neural network as shown in Fig. 7e is developed to extract depth-attentional features and to regress a residual map for predicting the clean image. In [23], a dual CNN as shown in Fig. 7f is presented, where two branches learn the estimation of two parts of the target signal: structures and details.

Another type of constraints is the confidence information of the residual between the estimated background layer and the ground-truth. In Fig. 7g, the model [29] learns about rain streaks by being guided by a per-pixel confidence map. This map is used to weigh wrongly estimated pixels in the back-propagation process, so that the network can pay more attention to these pixels during the training process. A summary of deraining side information used in previous works is provided in Table 4.

*Image Processing Knowledge.* Some ideas present in the image processing literature can also be useful in designing deraining network architectures, for instance: multi-scale structure, Laplacian pyramid, wavelet transform, etc. A scale-free network [30] as shown in Fig. 8a pays attention to the scale variety of rain streaks in real scenes, and constructs a scale-free deraining architecture by unrolling a wavelet transform into a recurrent neural network, which can handle various kinds of rain at different scales. Guided by the hierarchical representation of the wavelet transform, a recurrent network consisting of two stages is built: 1) rain removal on the low-frequency component; 2) recurrent detail recovery on high-frequency components guided by the recovered low-frequency component.

PyramidDerain [11] pursues a light-weighted pyramid of network as shown in Fig. 8b to remove rain from a single image. The decomposed Gaussian Laplacian image pyramid is combined with a deep network. The learning paradigm at each pyramid layer can be simplified, and the obtained network becomes shallow and has less parameters. The model is quite light-weighted and achieves comparable state-of-the-art performance.

Li *et al.* [31] propose a recurrent network to remove rain streaks progressively as shown in Fig. 8c. The intermediate result from the last recurrence is taken as the input of the next recurrence, and the features are also forwarded and fused by RNN units, e.g., GRU and LSTM, across recurrences. Ren *et al.* [9] utilize recursive computation to obtain more effective processing as shown in Fig. 8d. The PReNet performs stage-wise operations that process the input and intermediate results to generate the clean output images progressively.

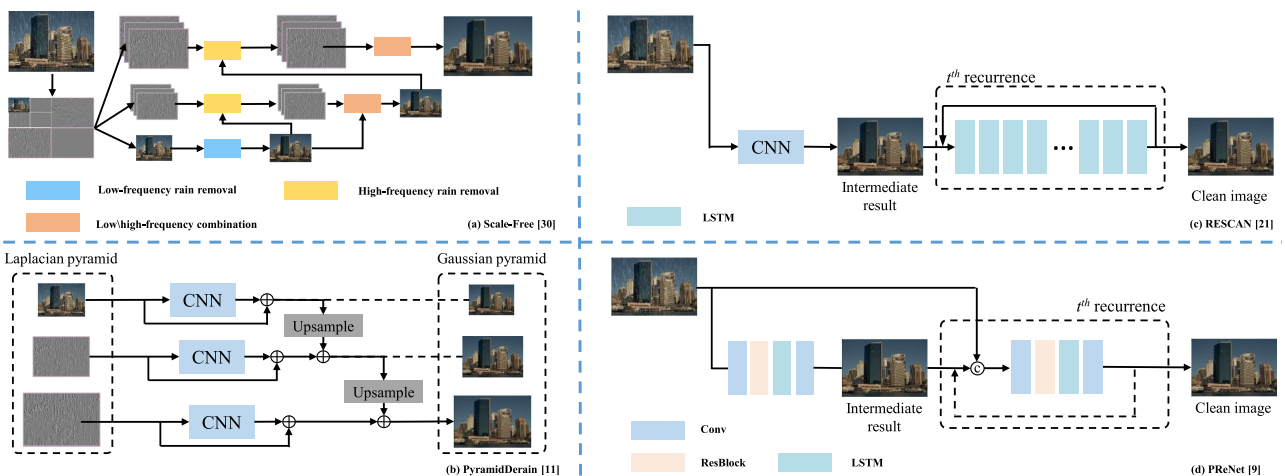


Fig. 8. Summary of the network improvement for single-image deraining. More effective network architectures are designed by relying on certain assumptions/constraints and general knowledge in image processing.

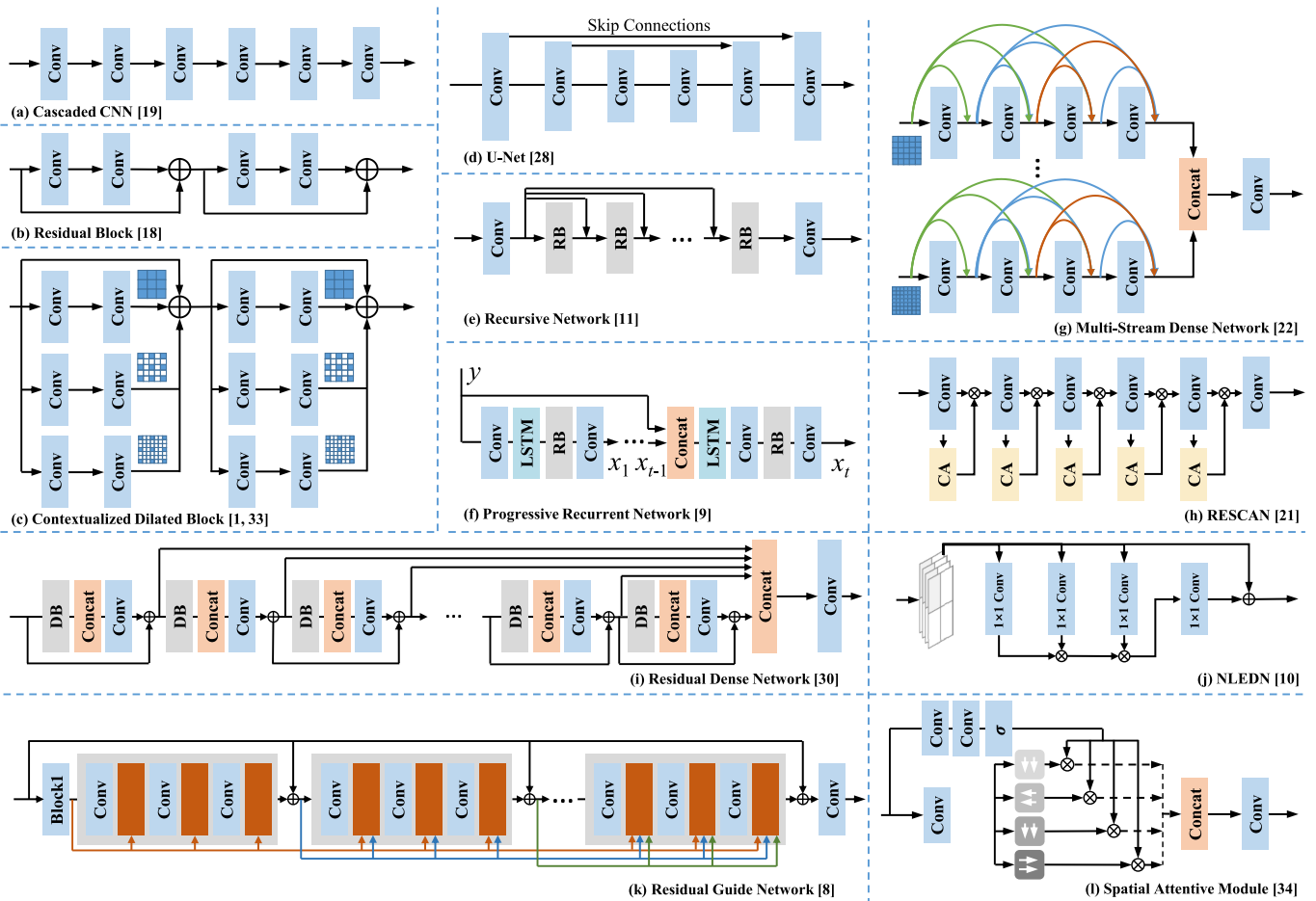


Fig. 9. The basic block improvement for single-image rain removal. The trend of newly proposed methods is to have more complex basic blocks with more powerful modeling capacities, which are further stacked into a more complex deraining network.

### 4.2 Basic Blocks

With the development of deep-learning based methods, the trend of newly proposed methods is to have more complex basic blocks with more powerful modeling capacities, which are further stacked into a more complex deraining network.

*Using Existing Architectures.* DetailNet [18], [19] (Figs. 9a and 9b) introduces a residual network and a cascaded CNN for rain removal. AttGAN [28] (Fig. 9d) utilizes U-Net as the baseline of the generator, which is effective to fuse the information from different scales to obtain global information while maintaining local details. RESCAN [31] (Fig. 9h) introduces channel-wise attention to adjust the relative weighting among channels for better separating rain streaks from background layers.

Multi-stream dense network [22] (Fig. 9g) combines dense block and convolutional networks. Residual dense network [30] (Fig. 9i) integrates dense blocks into residual networks. In [8] (Figs. 9k and 9e), basic blocks are connected in the recursive way, where the input feature is also forwarded to the intermediate features of the network. In [9] (Fig. 9f), residual blocks are also aggregated in the recursive way and the LSTMs are selected to connect different recurrences.

*Multi-Path Architectures.* One of the common architectures is the multi-path network. As shown in Figs. 9c, 9g and 9l, the networks have different paths possessing different properties, i.e., kernel sizes, dilation factors, and filter directions, to gather different kinds of information. In Figs. 9c and 9e,

different paths have different receptive fields, and thus can obtain both global information and maintain local structural details. In Fig. 9l, the spatial redundancies are aggregated from different directions to form visual attention.

*Recursive Architectures.* In [8], [9], [11], [31], recursive blocks are nested and aggregated in the recursive way as shown in Figs. 9f, 9k, 9e and 9i. The networks perform stage-wise operations that process the input and intermediate results to generate the output clean images progressively. Inter-stage recursive computation of different blocks is sometimes adopted to propagate information across the blocks.

Non-locally enhanced encoder-decoder network [10] as shown in Fig. 9j incorporates nonlocal operations to the design of an end-to-end network for deraining. The non-local operation calculates the feature at a spatial position as a weighted sum of the features at a specific range of positions. In [34], a spatial attentive module as shown in Fig. 9l employs recurrent neural networks with ReLU and identity matrix initialization, to accumulate global contextual information in four directions. It utilizes another branch to capture the spatial contexts to selectively highlight the transformed rain features.

### 4.3 Loss Functions

In existing deraining methods, several loss functions have been proposed to regularize the training of the deraining network.

*Fidelity-Driven Metrics.* Most studies need to use signal fidelity-driven matrices as the loss functions, such as Mean Squared Error (MSE) (L2), mean squared error (MAE) (L1), and SSIM [41]. They are defined as follows:

$$\mathcal{L}_{\text{MSE}}(x, \hat{x}) = \|x - \hat{x}\|_2^2, \quad (15)$$

$$\mathcal{L}_{\text{MAE}}(x, \hat{x}) = \|x - \hat{x}\|, \quad (16)$$

$$\mathcal{L}_{\text{SSIM}}(x, \hat{x}) = \frac{(2\mu_x\mu_{\hat{x}} + c_1)(2\sigma_{x\hat{x}} + c_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + c_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + c_2)}, \quad (17)$$

where  $x$  and  $\hat{x}$  are the ground truth and predicted clean images.  $\mu_x$  and  $\mu_{\hat{x}}$  are the average of  $x$  and  $\hat{x}$ , respectively.  $\sigma_x$  and  $\sigma_{\hat{x}}$  are the variance of  $x$  and  $\hat{x}$ , respectively.  $c_1$  and  $c_2$  are two numbers to stabilize the division with weak denominator.

*Rain-Related Loss.* The rain-related variable prediction loss makes some outputs of the network predict the rain-related variable. For example, in [1], the streak and the binary streak maps are connected to the corresponding losses as follows:

$$\begin{aligned} \mathcal{L}_S(s, \hat{s}) &= \|s - \hat{s}\|_2^2, \\ \mathcal{L}_R(r, \hat{r}) &= -\sum_i (\log \hat{r}_{i,1} r_{i,1} + \log(1 - \hat{r}_{i,2})(1 - r_{i,2})), \end{aligned} \quad (18)$$

where  $s$ ,  $\hat{s}$ ,  $r$  and  $\hat{r}$  are the ground truth rain streak, predicted rain streak, ground truth rain mask, and predicted rain mask, respectively.  $i$  indexes the spatial pixel location.

*Multi-Scale Loss.* The multi-scale loss [28] constrains the deraining network at different scales, which is expressed as

$$\mathcal{L}_{\text{MS}}(x, \hat{x}) = \sum_i \|x_{s_i} - \hat{x}_{s_i}\|_2^2, \quad (19)$$

where  $s_i$  indexes the scale, and  $x_{s_i}$  and  $\hat{x}_{s_i}$  are the down-sampled versions of  $x$  and  $\hat{x}$  with the scaling factor  $s_i$ .

*Perception-Driven Loss.* Applying perceptual and adversarial losses [28] improves the perceptual quality of generated results. The perceptual loss is formulated as

$$\mathcal{L}_P(x, \hat{x}) = \|F(x) - F(\hat{x})\|_2^2, \quad (20)$$

TABLE 5  
Summary of Loss Functions Used in Existing Works

Loss Function	Methods
MSE (L2)	JORDER [1], [33] DetailNet [18], [19], DID-MDN [22], DAF-Net [16], CGAN [24], DualCNN [23], PRNet [9], RESCAN [31], Scale-Free [35], Residual guided net [8], Semi-Supervised [12], Scale-Aware [35]
MAE (L1)	PyramidDerain [11], NLEDN [10], SPA-Net [34], UD-GAN [12]
SSIM	PyramidDerain [11], PRNet [9], SPA-Net [34], Residual guided net [8]
Adversarial	AttGAN [28], CGAN [24]
Perceptual	AttGAN [28], DID-MDN [22], CGAN [24]
Multi-Scale	AttGAN [28]
Variable	JORDER [1], [33], Semi-Supervised [12], DID-MDN [22], DAF-Net [16], AttGAN [28], Scale-Aware [35], SPA-Net [34], UD-GAN [12]

where  $F(\cdot)$  is a pretrained CNN transformation.

*Adversarial Loss.* The adversarial loss used for deraining network is represented as

$$\mathcal{L}_{\text{AdvGen}} = -\log D(\hat{x}), \quad (21)$$

where  $D(\cdot)$  is a discriminator network that differentiates the generated  $\hat{x}$  and the ground truth  $x$ . A summary of loss functions used in previous works is given in Table 5.

#### 4.4 Datasets

There are a few benchmarking datasets for image deraining, as introduced in Table 6. These datasets are useful for network training as well as for evaluation:

- *Rain12* [6] includes 12 synthesized rain images with only one type of rain streaks.
- *Rain100L* and *Rain100H* [1] include the synthesized rain images with only one type and five types of rain streaks, respectively.

TABLE 6  
Summary of Datasets Used in Previous Works

Dataset	Number (#train/#test)	Highlight	Rain Model	Publication
Rain12	12	Only for testing.	ACM	Li <i>et al.</i> [6]
Rain100L	1,800/100	Synthesized with only one type of rain streaks (light rain case).	ACM	Yang <i>et al.</i> [1]
Rain100H	1,800/100	Synthesized with five types of rain streaks (heavy rain case).	ACM	Yang <i>et al.</i> [1]
Rain800	700/100	Clean images are selected from BSD500 and UCID [36].	ACM	Zhang <i>et al.</i> [24]
Rain14000	9,100/4,900	1,000 clean image used to synthesize 14,000 rain images.	ACM	Fu <i>et al.</i> [18]
Rain12000	12,000/4000	The data has three kinds of densities.	ACM	Zhang <i>et al.</i> [22]
RealDataset	28,500/1,000	The ground truth data is synthesized based on temporal redundancy and visual properties.	ACM	Wang <i>et al.</i> [34]
NYU-Rain	13,500/2,700	Background images and the depth information are selected from NYU-Depth V2 [37].	HRM	Li <i>et al.</i> [13]
Outdoor-Rain	9,000/1,500	The background images are collected from [28], and the depth information is estimated by [38].	HRM	Li <i>et al.</i> [13]
RainCityscapes	9,432/1,188	The rain-free images are selected from the training and validation sets of Cityscape [39]. Rain patches are selected from [6].	DARM	Hu <i>et al.</i> [16]
MPID	1,561/419	The MPID dataset covers a much larger diversity of rain models, including both synthetic and real-world images, serving both human and machine visions.	ACM + HRM	Li <i>et al.</i> [26]

- *Rain800* [24]'s training set consists of a total of 700 images, where 500 images are randomly chosen from the first 800 images in the UCID dataset [36] and 200 images are randomly chosen from the BSD500's training set [40]. The testing set consists of a total of 100 images, where 50 images are randomly chosen from the last 500 images in the UCID dataset and 50 images are randomly chosen from the testing set of the BSD-500 dataset.
  - *Rain14000* [18] includes 1,000 clean images from UCID dataset [36], BSD dataset [40] and Google image search being used to synthesize rainy images.
  - *Rain12000* [22] consists of 12,000 images in the training set, where each image is assigned a label based on its corresponding rain-density level (i.e., light, medium and heavy). There are 4,000 images per rain-density level in the dataset. The synthesized testing set includes 1,200 images.
  - *RealDataset* [34] includes 29,500 rain/rain-free image pairs that cover a wide range of natural rain scenes, where the rain-free images are synthesized based on temporal redundancy and visual properties.
  - *NYU-Rain* [13] is a new synthetic rain dataset taking images from NYU-Depth V2 [37] as background and the provided depth information to generate rain streak and accumulation layers. The dataset also considers the effect of image blurring presented in the rain image. It contains 16,200 image samples, where 13,500 images are used as the training set.
  - *Outdoor-Rain* [13]'s background images are collected from [28], and the depth information used to synthesize accumulation is produced by [38]. The dataset includes 9,000 training images and 1,500 validation images.
  - *MPID* [26]'s training set includes 2,400 synthetic rain streak image pairs, 861 synthetic raindrop image pairs, and 700 synthetic rain and mist image pairs. The testing set includes 200 synthetic rain streak image pairs, 149 synthetic raindrop image pairs, and 70 synthetic rain and mist image pairs, as well as 50 real rain streak images, 58 real raindrop images, and 30 real rain and mist images. The testing set also includes 2,496 and 2,048 real captured images in the driving and surveillance video conditions with human annotated object bounding boxes.
  - *RainCityscapes* [16] consists of 262 training images and 33 testing images from the training and validation sets of Cityscape [39], which are selected as the clean background images. Rain patches are selected from [6]. There are in total 9,432 training images and 1,188 testing images.
- 5) Deep Detail Network, DetailNet [19], DDN [18],
  - 6) Directional Global Sparse Model, DGSM [17],
  - 7) Recurrent Squeeze-and-Excitation Context Aggregation Net, RESCAN [31],
  - 8) Progressive Recurrent Network, PReNet [9],
  - 9) Enhanced JOint Rain DETection and Removal, JORDER-E [33],
  - 10) Heavy Rain Image Restoration, HeavyRain Restorer [13],
  - 11) Spatial Attentive Network, SPANet [16],
  - 12) Semi-supervised Image rain Removal, SSIR [25], and
  - 13) Density-aware Image De-raining using Multi-stream Dense Network, DID-MDN [22].

LP is built based on Gaussian mixture models. ID, DSC, and JCAS are designed based on sparse coding. JORDER-E, DetailNet, DDN, DID-MDN, PReNet, SPANet, and RESCAN are deep-learning based methods. HeavyRainRestorer integrates deep CNN and generative adversarial learning for rain removal. In our experiments, JORDER-E, DetailNet, PReNet, and RESCAN are trained on Rain100H. SPANet is trained on RealDataset. DID-MDN is trained on Rain800. HeavyRainRestorer is trained on NYU-Rain and Outdoor-Rain.

Peak signal-to-noise ratio (PSNR) and SSIM [43] are used for performance evaluation. A few different metrics are used, particularly when we do not have ground-truths, which we call non-reference metrics: Naturalness Image Quality Evaluator (NIQE) [44], Perception-based Image Quality Evaluator (PIQE) [45], Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [46], Integrated Local NIQE (IL-NIQE) [47], Spatial-Spectral Entropy based Quality (SSEQ) [48], SR Metrics [49], Entropy-based No-reference Image Quality Assessment (ENIQA) [50], Blind Image Quality Assessment through Anisotropy (BIQAA) [51], and Blind Image Quality Assessment (BIQA) [52], BLind Image Integrity Notator using DCT Statistics (BLIINDS-II) [53]. These metrics measure the visual quality in different ways including human perception in lightness distortion, texture preservation, spatial domain statistics, and natural preservation, etc.

## 5.1 Quantitative Evaluation

We compare the quantitative results of different rain removal methods in Fig. 10. The numbers are obtained from [54]. To observe the trend of the performance changes over the years, we order different methods by year in Fig. 10. This figure provides some interesting information. First, most of the deep learning-based methods achieve significantly superior performance to model-based methods. For example, DDN obtains more than 3, 7, and 0.7 dB on Rain100L, Rain100H, and Rain1400, respectively. Second, the best performance of different methods gradually converges. The performance gaps between RESCAN, PReNet and JORDER-E are considerably close.

## 5.2 Qualitative Evaluation

We also show the visual results of different methods in Fig. 12. The input images shown in the figure are diversified and difficult to be handled, including large rain streaks and dense rain accumulation. The top two panels clearly show that, JORDER-E (Fig. 10j) and PReNet (Fig. 10k) are better at handling large rain streaks. JORDER-E (Fig. 10j) and

## 5 PERFORMANCE SUMMARY

We select a number of recent deraining algorithms from different categories to be evaluated:

- 1) Image Decomposition, ID [4],
- 2) Discriminative Sparse Coding, DSC [5],
- 3) Gaussian mixture model Layer Prior, LP [6],
- 4) Joint Convolutional Analysis and Synthesis Sparse Representation, JCAS [42],

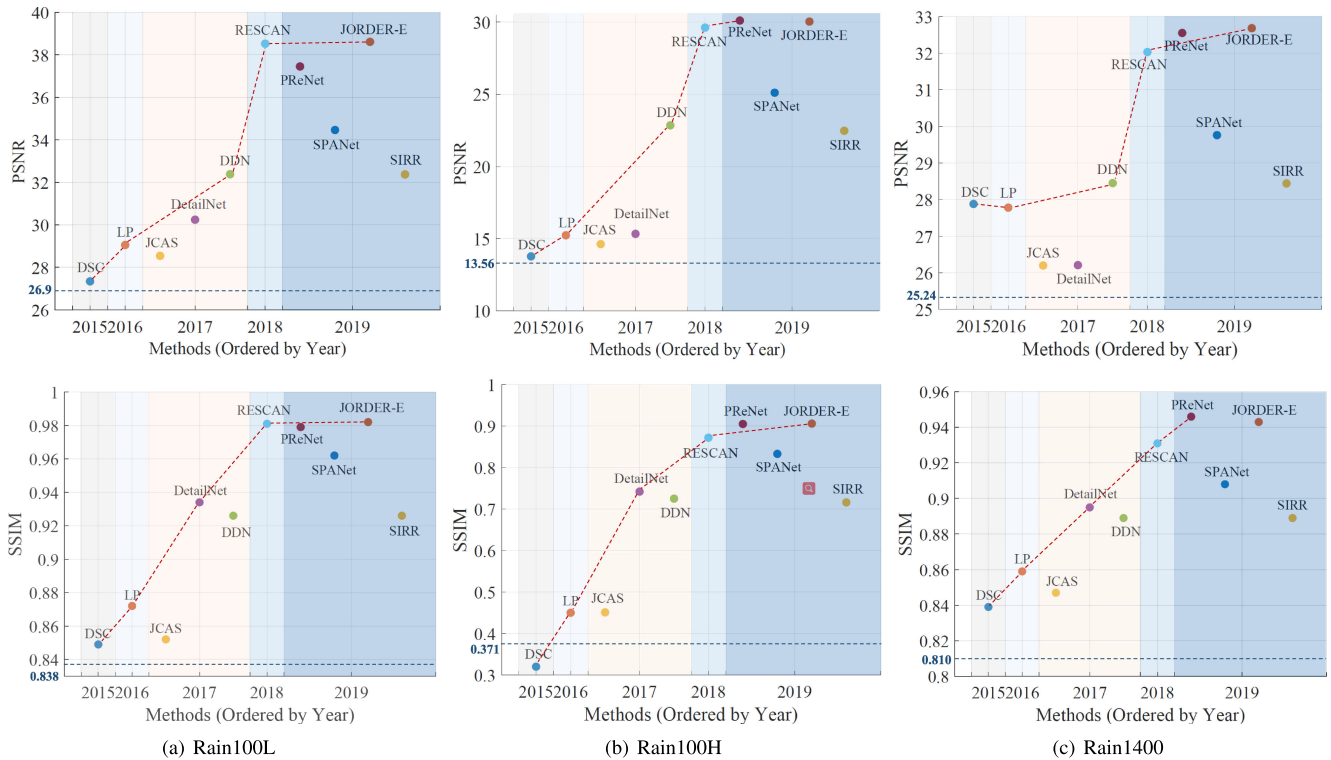


Fig. 10. The objective results of different methods. Top panel: PSNR. Bottom panel: SSIM. All methods are sorted by year. A red curve connects the top performance from 2015 to 2019. It is shown that, the objective performance gains converge gradually.

HeavyRainRemoval (Fig. 10f) achieve better results in removing rain accumulation and enhancing the visibility from the bottom three panels.

We also use qualitative assessment metrics for performance comparison of different methods by evaluating the consistency between the results of deraining methods and the subjective results by Mean Opinion Score (MOS).<sup>1</sup> There are 20 rain images for the evaluation. These images are processed by the methods and their results are evaluated by human annotators. 40 participants are invited in the subjective experiment. Each of them is required to provide subjective results for 550 image pairs.

The comparison results are visualized in Fig. 11. Based on the compared pairs, we further fit a Bradley-Terry [55] model to estimate the MOS score for each method so that they can be ranked. We can infer the MOS score for each input sample, and then combine the results of different samples via geometric mean, which is denoted as the *average MOS* in Table 7. We can also directly infer the MOS score with the accumulated ranking results of all samples, which is denoted as *MOS* value in Table 7.

In general, the paper published in 2019 are on average superior to previous methods on the dataset. However, the superiority of the qualitative comparison is not the same as that of the quantitative one, which reflects the disagreements between optimizing the quantitative metrics on the synthesized data and achieving better visual quality on real images. This is due to the domain gap between the real rain images and synthesized data.

1. More information about the subjective evaluation dataset, results and the evaluation website can be found at: [https://flyywh.github.io/Single\\_rain\\_removal\\_survey/](https://flyywh.github.io/Single_rain_removal_survey/).

We also observe that, all non-reference metrics are not in agreement with MOS and the qualitative values. We calculate Spearman rankorder correlation coefficient (SROCC), Kendallrank-order correlation coefficient (KROCC), and Pearson linear correlation coefficient (PLCC) in Table 8, where large absolute values denote that the metric can obtain more consistent results with respect to human perception. One can see that the values for the best result are only 0.2216, 0.1473, 0.1864 for SROCC, KROCC, and PLCC, respectively.

We conclude that all the existing metrics are not suitable to measure the performance of deraining, and thus there is

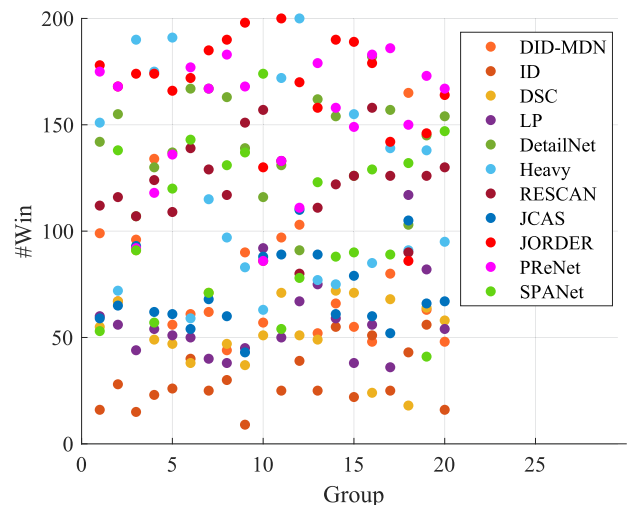


Fig. 11. Visualization of all paired comparisons. The horizontal axis denotes the comparison group ID while the vertical axis denotes the winning time in the comparison.

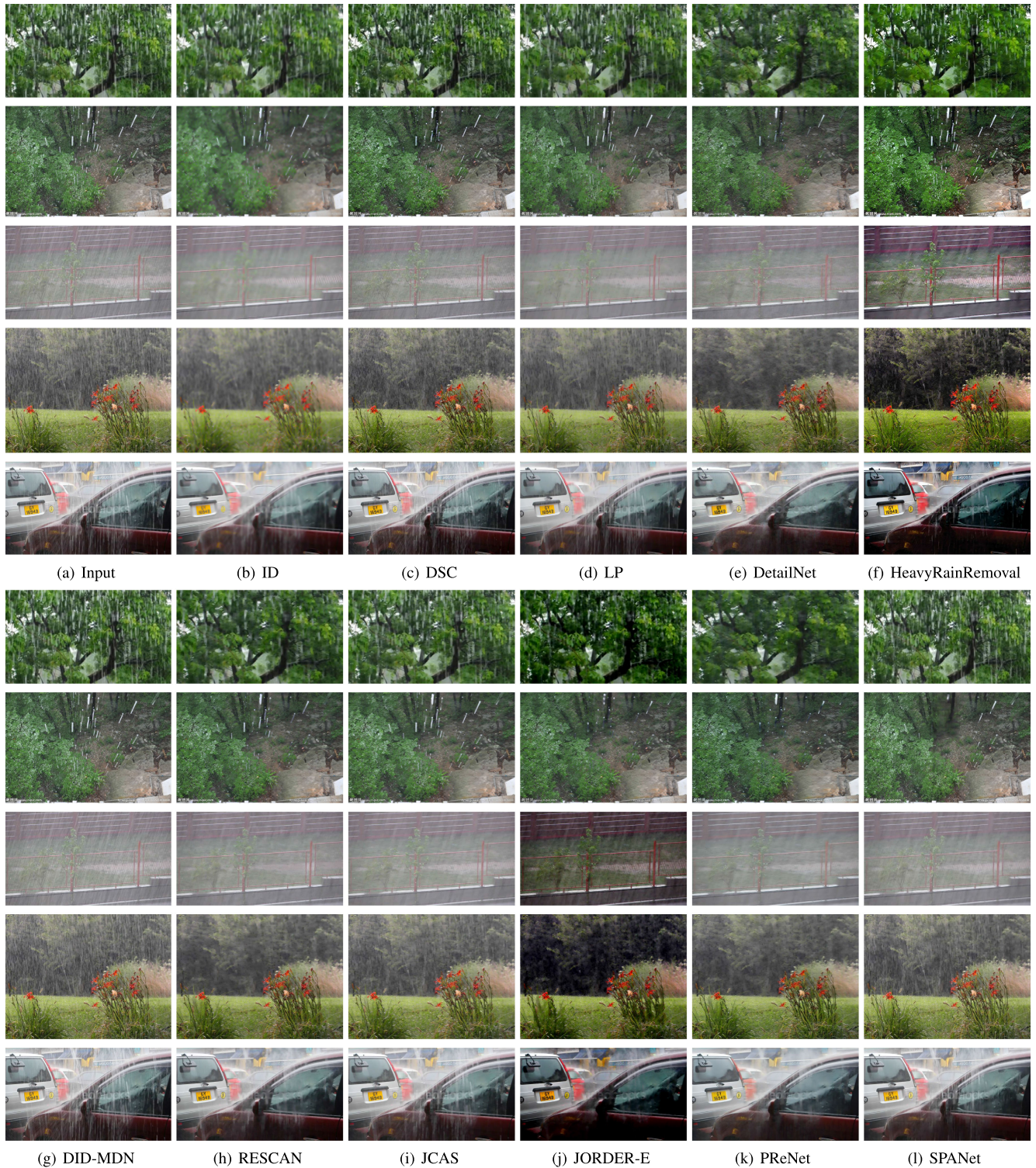


Fig. 12. Visual results of different methods. 1st and 2nd panels: rain images with large rain streaks. 3th and 4th panels: rain images including rain accumulation. 5th panel: rain images including both large rain streaks and accumulation.

great potential for future works on the deraining performance evaluation.

### 5.3 Computational Complexity

Table 9 compares the runtime of different state-of-the-art methods. All sparse coding based methods are implemented in MATLAB and tested on a CPU, following the original setting of all the released codes, while other

methods are accelerated by a GPU. ID-CGAN is implemented in Torch7. The rest is implemented in Pytorch. One can observe that JORDER-E, HeavyRainRemoval, and URML employ many more parameters than other methods. The comparison results on both performance and parameter number show that PReNet is an impressive method quantitatively and qualitatively, while keeping a light-weighted framework.

TABLE 7  
The Non-Reference Metric Results of Different Methods

Methods	Input	ID	DSC	LP	DetailNet	Heavy	DID-MDN	RESCAN	JCAS	JORDER-E	PreNet	SPANet	↑ or ↓
NIQE	5.38	4.46	4.55	5.46	5.34	4.97	5.07	3.78	4.97	3.93	4.78	3.92	↓
PIQE	38.36	39.41	40.04	64.86	36.50	52.81	38.97	24.28	35.64	29.77	45.72	30.40	↓
BRISQUE	34.00	31.44	32.51	41.01	32.99	37.32	30.12	25.70	32.91	26.51	30.28	28.38	↓
ILNIQE	31.09	31.20	29.16	41.78	30.72	31.84	26.32	30.63	26.53	28.46	29.70	30.58	↓
SSEQ	24.91	28.66	29.02	44.84	22.83	32.76	26.41	21.94	17.94	23.27	26.48	25.15	↓
SR-Metric	7.90	7.73	7.46	4.76	8.12	7.26	8.06	7.90	8.29	7.82	7.66	7.85	↑
ENIQA	0.1508	0.2012	0.1886	0.2631	0.1323	0.2091	0.1331	0.1347	0.1166	0.1394	0.1499	0.1445	↓
BIQAA	0.0107	0.0036	0.0051	0.0031	0.0123	0.0069	0.0085	0.0041	0.0165	0.0040	0.0043	0.0078	↓
BIQI	42.84	11.75	34.12	18.58	40.05	-5.15	37.05	35.40	29.81	34.46	31.02	24.34	↑
BLIINDS-II	22.55	19.80	22.00	22.85	21.78	25.33	20.60	12.58	20.13	12.53	24.03	16.08	↓
FRISQUE	54.99	54.32	52.95	38.40	58.97	52.75	25.29	57.56	66.70	58.49	58.95	58.54	↑
Average MOS	-	0.1042	0.2232	0.2440	1.8377	1.0974	0.4004	1.2413	0.3527	3.1366	2.3252	1.0	↑
MOS	-	0.1466	0.2804	0.3050	1.8897	1.2510	0.4825	1.3037	0.4233	3.1988	2.3802	1.0	↑

Heavy denotes HeavyRainRemoval. Red, blue, and green denote the best, second best, and third best results.

## 6 FUTURE DIRECTIONS

### 6.1 Integration of Physics Model and Real Images

Many existing learning-based methods rely on synthetic rain images to train the networks, since to obtain paired rain images and their exact clean ground-truths is intractable. While such a training scheme shows some degree of success, to improve the performance, we need to incorporate both real rain images the training process; otherwise, the network will never been exposed to the real rain images, impeding the network's effectiveness in the testing stage. Incorporating real rain images, however, can pose problems, because to obtain the paired clean background images is intractable. Consequently, there is no loss for a network to learn. To address this problem, we may rely on physics-based constraints. An attempt in [13] has shown the feasibility of this direction. Specifically, it combines the power of a physics model and a generative adversarial network, which can accept unpaired ground-truths. In the future, more works are expected in this direction to make efforts to combine of physical models and real rain images.

### 6.2 Rain Modeling

The current synthetic rain models can only cover limited types of rain streaks, e.g., a range of scales, shapes, directions, etc. However, in practice, the appearance of rain streaks is diverse, due to many different factors that can

influence rain conditions, e.g., 3D environments, distances, wind directions/speed, etc. Currently, when the distributions of captured rain streaks are different from the synthetic images in the training, the methods tend to fail to remove rain properly. The studies of [12], [25] attempt to model the rain appearance via the generation model and unpaired learning. However, observing their generated rain images, one can visibly see that they are not as diverse as real rain can be, and they are visibly not real enough. The latter can also cause problems, since it means there are significant gaps between synthetic and real rain images.

### 6.3 Evaluation Methodology

With a rapid growth of works on rain removal, it is still challenging to measure whether a method is sufficiently effective. As shown in Section 5.2, existing quality assessment methods are still far from capturing real visual perception of human. Thus, there is a potential direction which the community can pay more attention to. The quality assessment of rain removal methods can be considered from two aspects. First, for human vision, the metric should be designed to model the typical distortions caused by rain and deraining methods, and to describe the human preferences to different deraining results. Wuet *al.* [66] make the first attempt in that direction. However, the attempt is preliminary and there is plenty of room for further exploration. Second, for machine vision, we could consider the performance of high-level vision tasks in rain conditions. The MPID dataset makes the preliminary attempt by constructing task-driven evaluation sets for traffic detection. In the future, we hope that more large-scale task-driven evaluation sets with more applications in more diverse rain conditions.

### 6.4 More Related Tasks and Real Applications

When existing deraining methods are applied to real applications, there are a few factors that should be considered. First, the runtime of the method. Current methods are far from the requirement of real-time processing (30 fps). How to accelerate existing methods is a future challenge. Second, real rain images usually contain more complicated visual degradation. For example, the surveillance videos are compressed and also include compression distortion, e.g.,

TABLE 8  
The Evaluation Results of All Quality Assessment Models

Methods	SROCC	KROCC	PLCC
NIQE	0.0780	0.0461	0.0700
PIQE	0.2118	0.1437	0.1215
BRISQUE	0.1896	0.1297	0.1508
ILNIQE	0.0778	0.0508	0.1458
SSEQ	0.2216	0.1473	0.1257
SR-Metric	0.1132	0.0760	0.1129
ENIQA	0.1333	0.0932	0.1487
BIQAA	0.1365	0.0927	0.1383
BIQI	0.2001	0.1299	0.1558
BLIINDS2	0.1705	0.1208	0.1752
FRISQUE	0.2083	0.1407	0.1864



TABLE 9  
The Model Complexity and Running Time of Different Methods

Baseline	ID	DSC	LP	UGSM	JCAS	DetailNet	DID-MDN
Time (Seconds)	283.69	398.66	1177.3	2.51	188.28	0.61	0.53
GPU/GPU	(C)	(C)	(C)	(C)	(C)	(G)	(G)
Para.	-	-	-	-	-	57,369	372,839
Baseline	JORDER-E	RESCAN	ID-CGAN	SPANet	URML	HeavyRainRemoval	PRNet
Time (Seconds)	0.13	0.61	0.50	1.72	2.02	0.73	0.11
GPU/GPU	(G)	(G)	(G)	(G)	(G)	(G)	(G)
Para.	4,169,024	149,823	263,686	283,716	984,356	40,627,038	168,963

The size of the input image is  $512 \times 512$ . C and G denote CPU and GPU, respectively.

blocking artifacts. Effective deraining methods also need to take care of these issues. Third, there are scenarios where composite degradation might be involved, e.g. night-time rain conditions, mixture of raindrop and rain streak, etc. It will be interesting to detect the degradation types and handle them in a unified framework adaptively.

## 7 CONCLUSION

We have surveyed single-image deraining methods based on model-based and data-driven approaches. We discussed the rain models, the challenges of single-image deraining, and the basic ideas of the model-based and data-driven methods. In our discussion, the model-based methods are categorized further into: layer decomposition, sparse coding and GMMs; and, data driven based methods are grouped into: deep CNN, generative adversarial network, and semi/unsupervised learning methods. We also learned that data-driven methods generally perform better than the model-based methods. However, there are still a few open problems, particularly in the data-driven approach. Problems such as fusing physics models and real-rain images, more accurate rain models, evaluation methodology, and real applications of deraining still need further developments.

## ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China under Grant 2018AA A0102702, the Hong Kong RGC Early Career Scheme under Grant 9048122 (CityU 21211018), National Natural Science Foundation of China under contract No.61772043, and Beijing Natural Science Foundation under contract No.L182002. The work of Robby T. Tan was supported by MOE2019-T2-1-130.

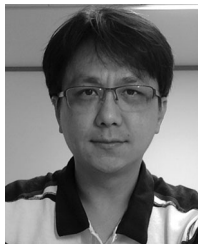
## REFERENCES

- [1] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1685–1694.
- [2] W. Yang, J. Liu, and J. Feng, "Frame-consistent recurrent video deraining with dual-level flow," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 1661–1670.
- [3] K. Garg and S. K. Nayar, "Detection and removal of rain from videos," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2004, pp. I–I.
- [4] L. W. Kang, C. W. Lin, and Y. H. Fu, "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1742–1755, Apr. 2012.
- [5] Y. Luo, Y. Xu, and H. Ji, "Removing rain from a single image via discriminative sparse coding," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 3397–3405.
- [6] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Rain streak removal using layer priors," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2736–2744.
- [7] L. Zhu, C. Fu, D. Lischinski, and P. Heng, "Joint bi-layer optimization for single-image rain streak removal," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2545–2553.
- [8] Z. Fan, H. Wu, X. Fu, Y. Huang, and X. Ding, "Residual-guide network for single image deraining," in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 1751–1759.
- [9] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3932–3941.
- [10] G. Li, X. He, W. Zhang, H. Chang, L. Dong, and L. Lin, "Non-locally enhanced encoder-decoder network for single image de-raining," in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 1056–1064.
- [11] X. Fu, B. Liang, Y. Huang, X. Ding, and J. Paisley, "Lightweight pyramid networks for image deraining," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 22, 2019, doi: 10.1109/TNNLS.2019.2926481.
- [12] X. Jin, Z. Chen, J. Lin, Z. Chen, and W. Zhou, "Unsupervised single image deraining with self-supervised constraints," in *Proc. IEEE Int. Conf. Image Process.*, 2019, pp. 2761–2765.
- [13] R. Li, L.-F. Cheong, and R. T. Tan, "Heavy rain image restoration: Integrating physics model and conditional adversarial learning," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 1633–1642.
- [14] K. Garg and S. K. Nayar, "Vision and rain," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 3–27, Oct. 2007.
- [15] J. Liu, W. Yang, S. Yang, and Z. Guo, "Erase or fill? Deep joint recurrent rain removal and reconstruction in videos," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3233–3242.
- [16] X. Hu, C.-W. Fu, L. Zhu, and P.-A. Heng, "Depth-attentional features for single-image rain removal," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8014–8023.
- [17] L.-J. Deng, T.-Z. Huang, X.-L. Zhao, and T.-X. Jiang, "A directional global sparse model for single image rain removal," *Appl. Math. Modelling*, vol. 59, pp. 662–679, 2018.
- [18] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1715–1723.
- [19] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2944–2956, Jun. 2017.
- [20] R. Li, L.-F. Cheong, and R. T. Tan, "Single image deraining using scale-aware multi-stage recurrent network," Dec. 2017, *arXiv: 1712.06830*.
- [21] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *Proc. IEEE Eur. Conf. Comput. Vis.*, 2018, pp. 262–277.
- [22] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 695–704.
- [23] J. Pan et al., "Learning dual convolutional neural networks for low-level vision," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3070–3079.
- [24] H. Zhang, V. Sindagi, and V. M. Patel, "Image de-raining using a conditional generative adversarial network," in *IEEE Trans. Circuits Syst. Video Technol.*, early access, 2020.

- [25] W. Wei, D. Meng, Q. Zhao, Z. Xu, and Y. Wu, "Semi-supervised transfer learning for image rain removal," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3872–3881.
- [26] S. Li *et al.*, "Single image deraining: A comprehensive benchmark analysis," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3833–3842.
- [27] D. Eigen, D. Krishnan, and R. Fergus, "Restoring an image taken through a window covered with dirt or rain," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 633–640.
- [28] R. Qian, R. T. Tan, W. Yang, J. Su, and J. Liu, "Attentive generative adversarial network for raindrop removal from a single image," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 2482–2491.
- [29] R. Yasarla and V. M. Patel, "Uncertainty guided multi-scale residual learning-using a cycle spinning CNN for single image de-raining," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8397–8406.
- [30] W. Yang, J. Liu, S. Yang, and Z. Guo, "Scale-free single image deraining via visibility-enhanced recurrent wavelet learning," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2948–2961, Jun. 2019.
- [31] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Rescan: Recurrent squeeze-and-excitation context aggregation net," in *Proc. IEEE Eur. Conf. Comput. Vis.*, 2018, pp. 262–277.
- [32] Y. Wang, S. Liu, C. Chen, and B. Zeng, "A hierarchical approach for rain or snow removing in a single color image," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3936–3950, Aug. 2017.
- [33] 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 Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 6, pp. 1377–1393, Jun. 2020.
- [34] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, and R. W. Lau, "Spatial attentive single-image deraining with a high quality real rain dataset," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 12262–12271.
- [35] R. Li, L.-F. Cheong, and R. T. Tan, "Single image deraining using scale-aware multi-stage recurrent network," *ArXiv e-prints*, Dec. 2017.
- [36] M. S. Gerald Schaefer, "UCID: An uncompressed color image database," *Storage and Retrieval Methods Appl. Multimedia*, vol. 5307, pp. 472–480, 2003.
- [37] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 746–760.
- [38] C. Godard, O. Mac Aodha, and G. J. Brostow, "Unsupervised monocular depth estimation with left-right consistency," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, 2017, pp. 6602–6611.
- [39] M. Cordts *et al.*, "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3213–3223.
- [40] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
- [41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [42] S. Gu, D. Meng, W. Zuo, and L. Zhang, "Joint convolutional analysis and synthesis sparse representation for single image layer separation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 1717–1725.
- [43] A. C. Brooks, X. Zhao, and S. Member, "Structural similarity quality metrics in a coding context: Exploring the space of realistic distortions," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1261–1273, Aug. 2008.
- [44] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
- [45] N. Venkatanath, D. Praneeth, B. M. Chandrasekhar, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *Proc. IEEE Nat. Conf. Commun.*, 2008, pp. 1–6.
- [46] A. Mittal, A. K. Moorthy, and A. C. Bovik, "Blind/referenceless image spatial quality evaluator," in *Proc. Conf. Rec. Asilomar Conf. Signals Syst. Comput.*, 2011, pp. 723–727.
- [47] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
- [48] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Process.: Image Commun.*, vol. 29, no. 8, pp. 856–863, 2014.
- [49] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang, "Learning a no-reference quality metric for single-image super-resolution," *Comput. Vis. Image Understanding*, vol. 158, pp. 1–16, May 2017.
- [50] X. Chen, Q. Zhang, M. Lin, G. Yang, and C. He, "No-reference color image quality assessment: From entropy to perceptual quality," *EURASIP J. Image Video Process.*, vol. 2019, no. 1, Sep. 2019, Art. no. 77. [Online]. Available: <https://doi.org/10.1186/s13640-019-0479-7>
- [51] S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," *J. Opt. Soc. America A*, vol. 24, no. 12, pp. B42–B51, Dec. 2007.
- [52] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
- [53] M. A. Saad, A. C. Bovik, and C. Charrier, "A DCT statistics-based blind image quality index," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 583–586, Jun. 2010.
- [54] H. Wang, Y. Wu, M. Li, Q. Zhao, and D. Meng, "A survey on rain removal from video and single image," Sep. 2019, *arXiv: 1909.08326*.
- [55] R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs: The method of paired comparisons," *Biometrika*, vol. 39, no. 3/4, pp. 324–345, Dec. 1952.
- [56] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 895–900.
- [57] A. Yamashita, Y. Tanaka, and T. Kaneko, "Removal of adherent waterdrops from images acquired with stereo camera," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 400–405.
- [58] A. Yamashita, I. Fukuchi, and T. Kaneko, "Noises removal from image sequences acquired with moving camera by estimating camera motion from spatio-temporal information," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2009, pp. 3794–3801.
- [59] S. You, R. T. Tan, R. Kawakami, Y. Mukaigawa, and K. Ikeuchi, "Adherent raindrop modeling, detection and removal in video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 9, pp. 1721–1733, Sep. 2016.
- [60] H. Zhang and V. M. Patel, "Convolutional sparse and low-rank coding-based rain streak removal," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2017, pp. 1259–1267.
- [61] Y. Liu, D. Jaw, S. Huang, and J. Hwang, "DesnowNet: Context-aware deep network for snow removal," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 3064–3073, Jun. 2018.
- [62] Y. Quan, S. Deng, Y. Chen, and H. Ji, "Deep learning for seeing through window with raindrops," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 2463–2471.
- [63] R. Li, R. T. Tan, and L.-F. Cheong, "Robust optical flow in rainy scenes," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 299–317.
- [64] R. Li, R. T. Tan, L.-F. Cheong, A. I. Aviles-Rivero, Q. Fan, and C.-B. Schonlieb, "RainFlow: Optical flow under rain streaks and rain veiling effect," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 7303–7312.
- [65] S. Halder, J. Lalonde, and R. D. Charette, "Physics-Based Rendering for Improving Robustness to Rain," in *Proc. Int. Conf. Comput. Vis.*, Seoul, Korea, 2019, pp. 10202–10211.
- [66] Q. Wu, L. Wang, K. N. Ngan, H. Li, and F. Meng, "Beyond synthetic data: A blind deraining quality assessment metric towards authentic rain image," in *Proc. IEEE Int. Conf. Image Process.*, 2019, pp. 2364–2368.



**Wenhan Yang** (Member, IEEE) received the BS and PhD (Hons.) degrees in computer science from Peking University, Beijing, China, in 2012 and 2018, respectively. He is currently a postdoctoral research fellow with the Department of Electrical and Computer Engineering, National University of Singapore. He was a visiting scholar with the National University of Singapore, from 2015 to 2016. His current research interests include deep-learning based image processing, bad weather restoration, and related applications and theories.



**Robby T. Tan** (Member, IEEE) received the PhD degree in computer science from the University of Tokyo, Tokyo, Japan. He is currently an associate professor with Yale-NUS College and Electrical and Computing Engineering (ECE) Department, National University of Singapore. Previously, he was an assistant professor with Utrecht University. His research interests include machine learning and computer vision, particularly in dealing with bad weather, physics-based, and motion analysis.



**Shiqi Wang** (Member, IEEE) received the BS degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2008, and the PhD degree in computer application technology from Peking University, Beijing, China, in 2014. From 2014 to 2016, he was a postdoctoral fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. From 2016 to 2017, he was with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore, as a research fellow. He is currently an assistant professor with the Department of Computer Science, City University of Hong Kong. He has proposed more than 40 technical proposals to ISO/MPEG, ITU-T, and AVS standards, and authored/coauthored more than 150 refereed journal/conference papers. His research interests include video compression, image/video quality assessment, and image/video search and analysis. He received the Best Paper Award from IEEE ICME 2019, IEEE Multimedia 2018, PCM 2017, and is the co-author of a paper that received the Best Student Paper Award in the IEEE ICIP 2018.



**Yuming Fang** (Senior Member, IEEE) received the BE degree from Sichuan University, Chengdu, China, the MS degree from the Beijing University of Technology, Beijing, China, and the PhD degree from the Nanyang Technological University, Singapore. Currently, he is a professor with the School of Information Management, Jiangxi University of Finance and Economics, Nanchang, China. He serves as an associate editor of the *IEEE Access* and is on the editorial board of the *Signal Processing: Image Communication*. His research interests include visual quality assessment, computer vision, 3D image/video processing, etc.



**Jiaying Liu** (Senior Member, IEEE) received the PhD (hons.) degree in computer science from Peking University, Beijing, China, in 2010. She is currently an associate professor with the Institute of Computer Science and Technology, Peking University. She has authored more than 100 technical articles in refereed journals and proceedings, and holds 42 granted patents. Her current research interests include multimedia signal processing, compression, and computer vision. She is a senior member of the CSIG and CCF. She was a visiting scholar with the University of Southern California, Los Angeles, from 2007 to 2008. She was a visiting researcher with the Microsoft Research Asia in 2015 supported by the Star Track Young Faculties Award. She has served as a member of Membership Services Committee in IEEE Signal Processing Society, a member of Multimedia Systems & Applications Technical Committee (MSA TC), Visual Signal Processing and Communications Technical Committee (VSPC TC) in IEEE Circuits and Systems Society, a member of the Image, Video, and Multimedia (IVM) Technical Committee in APSIPA. She has also served as the associate editor of the *IEEE Transactions on Image Processing*, and the Elsevier *Journal of Visual Communication and Image Representation*, the technical program chair of IEEE VCIP-2019/ACM ICMR-2021, the publicity chair of IEEE ICME-2020/ICIP-2019/VCIP-2018/MIPR 2020, and the area chair of ECCV-2020/ICCV-2019. She was the APSIPA distinguished lecturer (2016-2017).

▷ For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/csdl](http://www.computer.org/csdl).