Raindrop Clarity: A Dual-Focused Dataset for Day and Night Raindrop Removal

Yeying Jin[1], Xin Li[2], Jiadong Wang[1], Yan Zhang[1], and Malu Zhang[3]*

[1] National University of Singapore
[2] University of Science and Technology of China
[3] University of Electronic Science and Technology of China
jinyeying@u.nus.edu, xin.li@ustc.edu.cn, maluzhang@uestc.edu.cn

Abstract. Existing raindrop removal datasets have two shortcomings. First, they consist of images captured by cameras with a focus on the background, leading to the presence of blurry raindrops. To our knowledge, none of these datasets include images where the focus is specifically on raindrops, which results in a blurry background. Second, these datasets predominantly consist of daytime images, thereby lacking nighttime raindrop scenarios. Consequently, algorithms trained on these datasets may struggle to perform effectively in raindrop-focused or nighttime scenarios. The absence of datasets specifically designed for raindrop-focused and nighttime raindrops constrains research in this area. In this paper, we introduce a large-scale, real-world raindrop removal dataset called Raindrop Clarity. Raindrop Clarity comprises 15,186 high-quality pairs/triplets (raindrops, blur, and background) of images with raindrops and the corresponding clear background images. There are 5,442 daytime raindrop images and 9,744 nighttime raindrop images. Specifically, the 5,442 daytime images include 3,606 raindrop- and 1,836 background-focused images. While the 9,744 nighttime images contain 4,838 raindrop- and 4,906 background-focused images. Our dataset will enable the community to explore background-focused and raindrop-focused images, including challenges unique to daytime and nighttime conditions. [4]

Keywords: Raindrop-focused · Background-focused · Nighttime

1 Introduction

Adherent raindrops on lenses or windscreens can significantly reduce visibility. Raindrop removal is crucial for surveillance, self-driving cars, object detection, outdoor photography, filmmaking, augmented reality, *etc*. Recently, a few raindrop datasets (e.g, [24–26, 30]) have been proposed. In Fig. 1, these datasets primarily contain daytime background scenes with blurry raindrops, as the images are taken with the camera's focus on the background. They have impacts in the raindrop removal area, inspiring several successful methods [24, 25, 27].

---

[*] Corresponding author

[4] Our data and code are available at: https://github.com/jinyeying/RaindropClarity

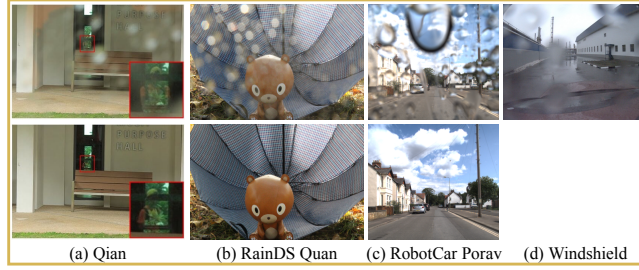(a) Qian   (b) RainDS Quan   (c) RobotCar Porav   (d) Windshield

Fig. 1. Existing raindrop datasets (e.g, [24–26, 30]) exhibit two limitations: first, they do not include raindrop-focused images (Fig. 2 right); and second, they lack night raindrops (Fig. 2 bottom).
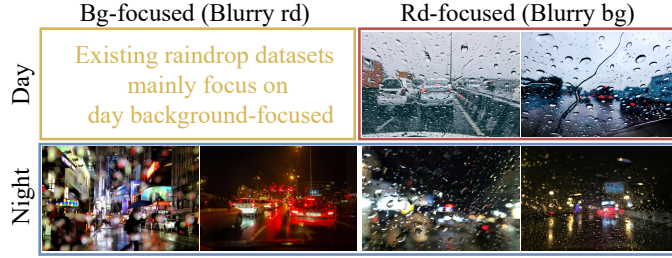

Fig. 2. Motivation for Raindrop Clarity: real images sourced from the Internet feature scenarios overlooked by existing datasets, including raindrop-focused (right) and nighttime raindrops (bottom). In this figure, Rd = Raindrop, Bg = Background.

While these datasets are valuable, they have two limitations. First, a significant issue arises when the camera is in automatic focus mode. In Fig. 2 (right), the camera may unintentionally focus on raindrops that are attached to the lens, windshield, or glass, resulting in images with sharp raindrops and a blurry background. Consequently, methods designed for images with blurry raindrops and a sharp background may not perform optimally. Second, existing datasets consist only of daytime images. Therefore, existing methods, especially those reliant on training data, may struggle to effectively handle nighttime raindrop images because of the inherent domain gap between day and night conditions.

In Fig. 2, we observe that besides daytime images with blurry raindrops and a sharp background, it is common to come across images with sharp raindrops and a blurry background (right), as well as nighttime raindrops (bottom). Thus, we collected varied raindrop images as shown in Fig. 3: the top-left with daytime blurry raindrops against a sharp background, the top-right with daytime sharp raindrops and a blurry background, the bottom-left with nighttime blurry raindrops and a clear background, and the bottom-right with nighttime sharp raindrops with a blurry background.

Physical solutions like rain covers or wipers can reduce the presence of raindrops, but their effectiveness is limited. In certain situations where wipers cannot
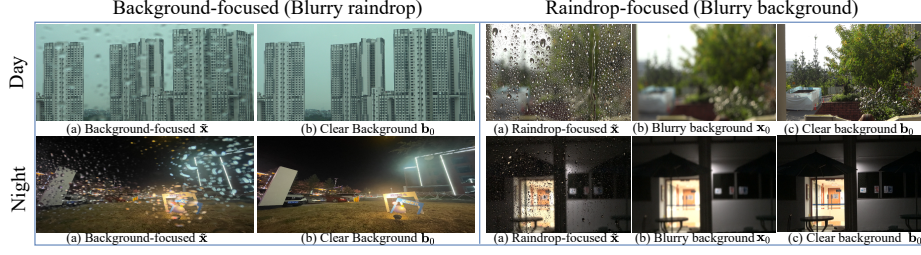
Fig. 3. Raindrop Clarity: Examples of our pairs $(\tilde{\mathbf{x}}, \boldsymbol{b}_0)$ and triplets $(\tilde{\mathbf{x}}, \boldsymbol{x}_0, \boldsymbol{b}_0)$ of a raindrop image $\tilde{\mathbf{x}}$, the corresponding blurry background image $\boldsymbol{x}_0$ and the corresponding clear background image $\boldsymbol{b}_0$, in day and night. For the pairs $(\tilde{\mathbf{x}}, \boldsymbol{b}_0)$, there is no $\boldsymbol{x}_0$, since in this case $\boldsymbol{x}_0 = \boldsymbol{b}_0$. The reason we have the blurry background image $\boldsymbol{x}_0$ in triplets is to have the raindrop difference maps $\tilde{\mathbf{m}}$ (see Fig. 8).

be used, or wind blows rain onto the lens, these physical measures become impractical. Physics-based methods [1, 8, 23, 44, 45] remove raindrops by leveraging the physics of raindrops or utilizing generated synthetic data. However, the effectiveness of methods trained on synthetic data relies on the quality of raindrop rendering models. Currently, raindrop rendering models commonly employ surface modeling and ray tracing [8, 28, 29]. Unfortunately, these methods may not fully consider factors such as scene depth, lighting conditions, and the 3D shapes of raindrops [23, 28], which are crucial for realistic raindrop rendering. Additionally, raindrops can introduce refraction effects [7, 44], distorting the shapes and positions of objects, which further complicates accurate raindrop rendering. To the best of our knowledge, developing a rendering approach that can produce physically accurate raindrops adaptable to various scenes is challenging.

Motivated by the aforementioned challenges and the observation that existing raindrop datasets predominantly consist of background-focused images while neglecting raindrop-focused images and nighttime raindrops, we collect a novel dataset, called "Raindrop Clarity". Our dataset comprises 6,742 (1,836 daytime, 4,906 nighttime) pairs and 8,444 (3,606 daytime, 4,838 nighttime) triplets of raindrop-focused images and background-focused images. All of them have the corresponding clear background images. The triplets include one more raindrop-free blurry background images. Fig. 3 shows samples and explanation on the meaning of our pairs and triplets. From the 15,186 raindrops pairs/triplet, 5,442 are captured during daytime and the 9,744 at nighttime. Raindrops Clarity exhibits diverse sizes and shapes, and raindrop appearance varies based on the scene and lighting conditions. To our knowledge, Raindrop Clarity is the largest real-world raindrop dataset that contains both the pair and triplets, day and night. Our main contributions are summarized as follows:

- We present the Raindrop Clarity dataset, a large-scale real-world dataset that encompasses both raindrop-focused triplets and background-focused pairs. For the pairs, we provide images with blurry raindrops and their cor-

responding clear backgrounds. For the triplets, we additionally provide the raindrop-free blurry background.
- Raindrop Clarity includes images collected under both daytime and nighttime. It aims to remove raindrops and recover clear backgrounds regardless of day or night, whether the raindrops or background are blurry.
- Our experimental results indicate that even with the application of state-of-the-art raindrop removal methods, there are instances where these methods fail. This suggests the existence of unresolved challenges in the field of raindrop removal, which await further attention and investigation.

2   Related Work

Raindrop Dataset. Qian *et al.* [25] introduce a raindrop dataset consisting of raindrop and corresponding ground truths. Porav *et al.* [24] propose a stereo RobotCar dataset, using a rain-making device that causes one lens to be affected by raindrops, while the other lens remains clear. Hao *et al.* [8] create a synthetic dataset of adherent raindrops with pixel-level masks for training deraining algorithms. Quan *et al.* [26] introduce RainDS, a dataset that includes raindrops, rain steaks, and their corresponding ground truths. Soboleva and Shipitko [30] introduce Windshields, a dataset containing raindrop images on both camera lenses and windshields, annotated with binary masks representing the raindrop areas. However, all of these datasets primarily consist of daytime images taken by cameras focusing on the background scenes and thus have blurry raindrops.

Single Image Raindrop Removal. Most image deraining methods have primarily focused on removing rain streaks rather than raindrops. However, recent learning-based methods have been proposed specifically for single-image raindrop removal. Eigen *et al.* [6] introduce a supervised CNN-based method to remove raindrops, using a shallow CNN. Qian *et al.* [25] propose an AttentiveGAN for raindrop removal, that relies on attention maps but has a limited focus on local spatial information. Quan *et al.* [27] propose RaindropAttention, which employs an attention mechanism to generate attention maps based on mathematical raindrop shapes, yet it lacks exploration of long-range information. Liu *et al.* [18] demonstrate a dual residual network, by leveraging the complementary nature of paired residual blocks. Hao *et al.* [8] propose a raindrop synthesis algorithm, used their synthetic dataset to train a multi-task network for raindrop detection and removal. Porav *et al.* [24] propose a Pix2PixHD-based deraining network architecture. Subsequent work by Quan *et al.* [26] propose a cascaded network architecture for simultaneous removal of rain streaks and raindrops, achieved through a neural architecture search. Yasarla and Patel [41] propose the uncertainty guided multi-scale residual learning network to effectively remove rain from images. Zhang *et al.* [47] propose the Dual Attention-in-Attention Model to remove rain streaks and raindrops.

Restoration Backbone. There are image restoration and enhancement tasks that utilize transformer-based models, taking advantage of the long-range dependencies, including IPT [2] (vanilla transformer), Uformer [36], SwinIR [17]
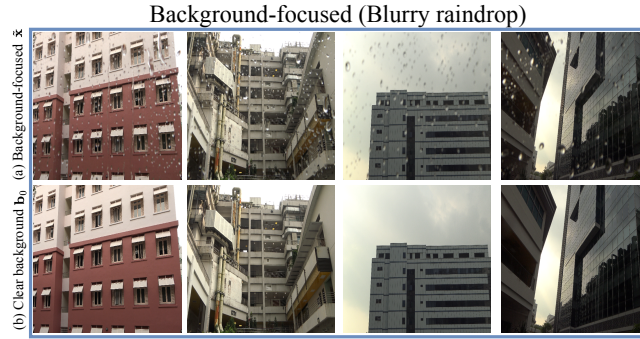
Background-focused (Blurry raindrop)

Fig. 4. Pairs $(\tilde{\mathbf{x}}, \boldsymbol{b}_0)$ show various sizes, shapes, and densities.

Raindrop-focused (Blurry background)

Fig. 5. The triplets $(\tilde{\mathbf{x}}, \boldsymbol{x}_0, \boldsymbol{b}_0)$ cover diverse scenes and shapes.

and Restormer [46]. Xiao *et al.* [40] propose an image deraining transformer (IDT) for rain streak and raindrop removal with window-based and spatial-based dual Transformer. RaindropDiffusion [21] develops a UNet-based diffusion framework, which demonstrates robust generation capabilities for raindrops, rain streaks, and snow. RainDiffusion [38], an unsupervised diffusion model, utilizes unpaired real-world data instead of weakly adversarial training. DiT [22] is a transformer-based diffusion model.

3   Raindrop Clarity Dataset

3.1   Data Collection

To create Raindrop Clarity, we employed a sphere pan-tilt platform to keep our camera stationary. The images were captured using Sony FDR-AX33 4K Ultra HD Handycam Camcorder, Sony alpha 7R III digital camera, the back camera of an iPhone 14 Pro and iPhone 15 Pro Max, ensuring the highest quality possible. The 4K Camcorder captured RGB images at a resolution 3840×2164.

Using the optical refraction model (raindrops appear upside-down), we projected raindrops by spraying water on glass and adjusting the distance, angle, speed, and patterns of the nozzle/wipers. We also captured in rain conditions.

Our collection procedure was designed to ensure diversity in the raindrop (such as focuses, shapes and sizes), background scenes and lighting conditions. First, we fixed the camera position on a tripod and installed a glass plate. The distance between the camera and the glass plate ranged from 5cm to 25cm. Second, we sprayed waterdrops/raindrops onto the glass plate and focused the camera on the raindrops. Third, we sprayed the waterdrops/raindrops to capture images of them under different shapes and sizes, ranging from sparse to dense. The captured images include elliptical raindrops and raindrop flow traces ($\tilde{\mathbf{x}}$). Fourth, we removed the glass plate, allowing the camera to focus on the very close plane. The duration is kept very short, within milliseconds, to minimize lighting changes. In this way, we obtained raindrop-free blurry background images ($\boldsymbol{x}_0$). Finally, we adjusted the camera focus to the background, directing the camera to focus on the distant background ($\boldsymbol{b}_0$).

3.2   Characteristics of Raindrop Clarity

Raindrop Clarity is a new dataset consisting of around 15,186 high-quality image raindrop pairs and triplets. In Fig. 3, our dataset introduces a unique element of triplet data, which includes raindrop-focused images, corresponding blurry raindrop-free background images, and their clear background equivalents, distinguishing it from existing datasets. The second feature that distinguishes the Raindrop Clarity from the existing ones is the collection of night raindrops.

Objectives. Our dataset is designed to support research in the elimination of raindrops and the restoration of a clear background, irrespective of whether the camera's focus is on the raindrops (leading to a blurry background) or on the background (resulting in blurry raindrops). This holds true regardless of whether it is daytime or nighttime.

Raindrops and Diversity. Our Raindrop Clarity dataset offers an extensive coverage of diverse raindrop sizes, shapes, and occlusion types in various scenes and lighting conditions, making it an exceptionally comprehensive resource for advancing raindrop removal research. In Fig. 4 and Fig. 5, the dataset captures raindrops and rain flows of different sizes and shapes, including heavy raindrops
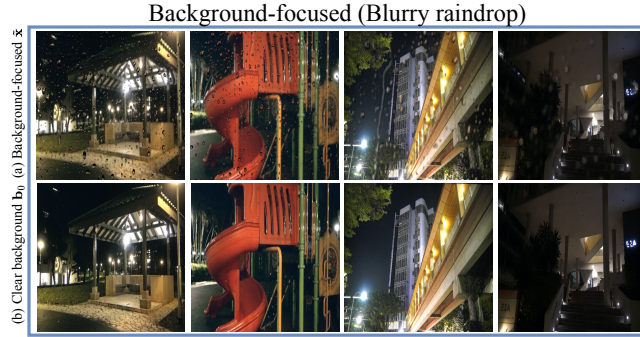
Background-focused (Blurry raindrop)

Fig. 6. The pairs $(\tilde{\mathbf{x}}, \boldsymbol{b}_0)$ under different lighting conditions.

Raindrop-focused (Blurry background)

Fig. 7. Triplets $(\tilde{\mathbf{x}}, \boldsymbol{x}_0, \boldsymbol{b}_0)$ under different lighting conditions.

that result in unpredictable water flows and blurry backgrounds when they hit window panes or car windshields. Two primary types of occlusions caused by raindrops, namely partial and complete occlusions, are included in our dataset. Both types present unique challenges for raindrop removal algorithms.

Scenes and Large-scale. The dataset encompasses a wide range of background scenes, such as urban, countryside, campus, wilderness, road, and aerial views, covering both day and night. This variation allows for the exploration of different lighting and weather conditions in different scenes. With approximately 38,816 images in total, including 15,186 high-quality pairs/triplets of images, the Raindrop Clarity dataset is one of the largest datasets of its kind.

Lighting Conditions. The appearance of raindrops highly depends on the surrounding environment. A distinctive feature of our dataset is its comprehen-

Table 1. Summary of comparisons between our Raindrop Clarity and existing datasets. Bg-focus = Background-focused, Rd-focus = Raindrop-focused, RD = Raindrop. Raindrop Clarity uniquely features raindrop-focused triplets and the nighttime raindrops.

| Datasets	Components	N. Images	Syn/ Real	Bg-focus		Rd-focus	
				Day	Night	Day	Night
Our RD Clarity (Day)	3,606 Triplets+1,836 Pairs	14,490	Real	✓	×	✓	×
Our RD Clarity (Night)	4,838 Triplets+4,906 Pairs	24,326	Real	×	✓	×	✓
Raindrop Qian [25]	919 Pairs	1,838	Real	✓	×	×	×
RainDS-Real [26]	248 Quadruplets	992	Real	✓	×	×	×
RainDS-Syn [26]	1,200 Quadruplets	4,800	Syn	✓	×	×	×
RobotCar Porav [24]	4,818 Pairs	9,636	Real	✓	×	×	×
Windshield [30]	3,390	3,390	Real	✓	×	×	×
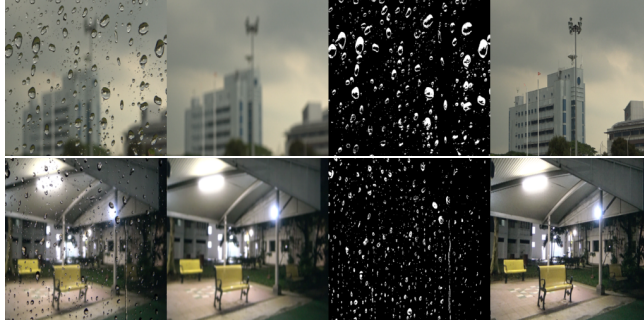
Fig. 8. We show $(\tilde{\mathbf{x}}, \boldsymbol{x}_0, \tilde{\mathbf{m}}, \boldsymbol{b}_0)$, difference map $\tilde{\mathbf{m}}$ is derived by distinguishing raindrop-focused $\tilde{\mathbf{x}}$ from blurry background $\boldsymbol{x}_0$.

sive coverage of various environmental conditions, setting it apart from other datasets. The dataset's variation in day and night conditions makes it valuable for developing raindrop removal algorithms.

During the daytime, raindrop images exhibit a diverse array of appearances influenced by multiple factors such as sunlight direction, intensity, and shadowing. In Figs. 4-5, our dataset captures this variety, with images taken at different times of the day and under varying weather conditions (such as light rain, heavy rain, and rain combined with wind), providing rich information for algorithms to learn.

Unlike daytime, nighttime raindrops present unique challenges due to artificial lighting conditions and low light. The appearance of raindrops can drastically change under street lights, headlights, or neon lights. In Fig. 6 and Fig. 7, our dataset includes night images with varied artificial lights, contributing to the complexity and comprehensiveness of the dataset. These night images are challenging due to the higher contrast between raindrops and the background, and potential overexposure or underexposure issues.

Background- and Raindrop-focused. When the raindrop image $\tilde{\mathbf{x}}$ is background-focused, we decompose a raindrop degraded image $\tilde{\mathbf{x}}$ into a clear background
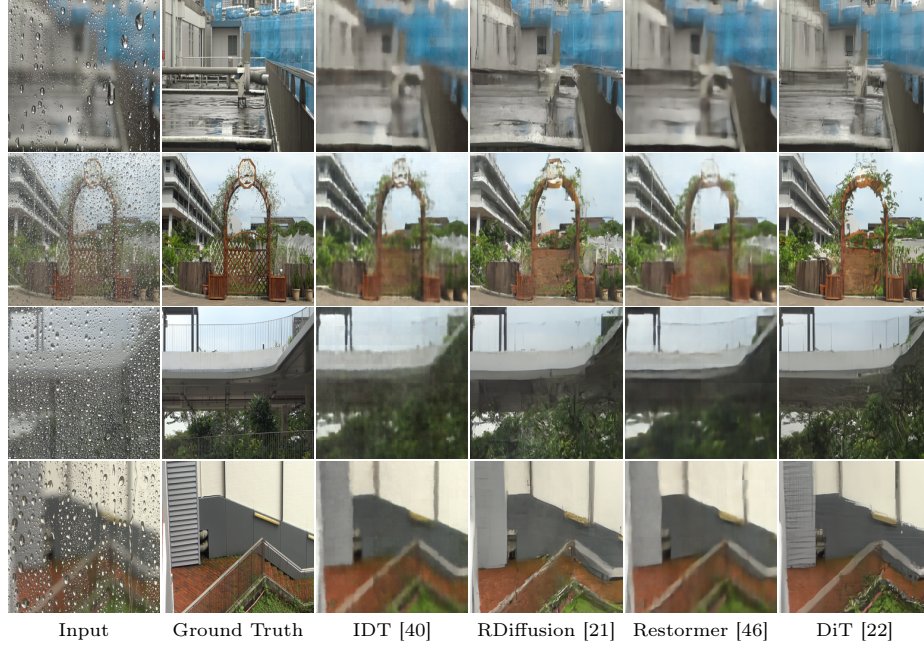
| Input | Ground Truth | IDT [40] | RDiffusion [21] | Restormer [46] | DiT [22] |

Fig. 9. Comparison of existing methods on daytime raindrop-focused images, covering both raindrop algorithms and restoration backbones.

image b_0 by following this model [25]: $\tilde{\mathbf{x}} = (1 - \mathbf{M}) \odot x_0 + \mathbf{D}$, where $\mathbf{D}$ is the effect of the blurry raindrops, referring to the localized blurry imagery formed by the light reflected from the surroundings. The operator $\odot$ denotes element-wise multiplication. When binary mask $\mathbf{M}(\mathbf{x})$ equals 1, it signifies that the pixel x belongs to the area affected by raindrops. Conversely, if $\mathbf{M}(\mathbf{x})$ does not equal 1, it indicates that the pixel is within the background areas. In the case where $\tilde{\mathbf{x}}$ is background-focused raindrop images, x_0 is identical to the clear, raindrop-free image b_0, namely $x_0 = b_0$. Therefore, we use $(\tilde{\mathbf{x}}, b_0)$ to stand for background-focused (resulting in blurry raindrops) pairs, shown in the left part of Fig. 3.

In the right part of Fig. 3 and Fig. 8, when $\tilde{\mathbf{x}}$ is the raindrop-focused image, x_0 is the corresponding raindrop-free blurry background image, while b_0 is the clear background image. In this case, $x_0 \neq b_0$. We can calculate the pixel-level difference map $\tilde{\mathbf{m}}$ between the raindrop-free blurry background x_0 and the raindrop-focused $\tilde{\mathbf{x}}$ as: $\tilde{\mathbf{m}} = \tilde{\mathbf{x}} - x_0$. Therefore, we use $(\tilde{\mathbf{x}}, x_0, b_0)$ to stand for raindrop-focused (resulting in blurry background) triplets.

Since the camera is stationary, neither the pairs $(\tilde{\mathbf{x}}, b_0)$ nor the triplets $(\tilde{\mathbf{x}}, x_0, b_0)$ have alignment problems. This is also evident in the raindrop difference maps $\tilde{\mathbf{m}}$, which accurately assign values to raindrop regions across multiple images. Specifically, for each pixel $\mathbf{x}$ in the input image, any discrepancies in the map's assigned values for the static background would indicate alignment issues; however, these are not observed, as $\tilde{\mathbf{m}}$ indicates only the raindrop regions.

Table 2. Quantitative evaluation of daytime and nighttime raindrop data with a benchmark for raindrop algorithms and restoration backbones.

Data\Method		Input	Raindrop Algorithms						Restoration Backbones		
			AtGAN [25]	Robot [24]	RainDS [26]	IDT [40]	RDdiff [21]	Restor. [46]	Uform. [36]	DiT [22]
Day	PSNR↑	21.92	23.62	23.82	25.39	26.05	25.52	**26.08**	25.78	26.03
	SSIM↑	0.560	0.658	0.670	0.717	0.736	0.734	0.748	0.727	**0.752**
	LPIPS↓	0.247	0.200	0.191	0.147	0.141	0.111	0.131	0.148	**0.106**
Night	PSNR↑	24.78	24.38	24.57	25.39	26.81	26.48	**26.87**	25.26	26.23
	SSIM↑	0.726	0.773	0.787	0.820	**0.851**	0.831	0.848	0.816	0.826
	LPIPS↓	0.209	0.185	0.182	0.139	0.125	0.112	0.123	0.142	**0.111**

| Input | Ground Truth | IDT [40] | RDiffusion [21] | Restormer [46] | DiT [22] |
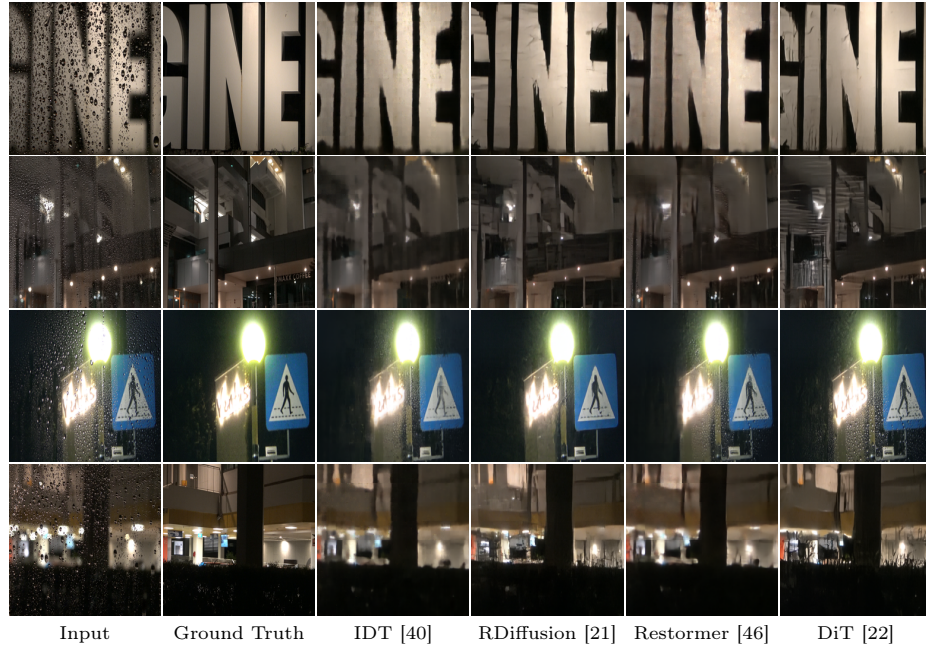
Fig. 10. Comparison of methods on nighttime raindrop-focused images, covering both raindrop algorithms and restoration backbones.

Since the pairs/triplets of images are captured in quick succession, there are no problems with changes in illumination.

The captured images do not contain any personal information such as phone numbers, or human faces. We will release the dataset publicly with car license plates obscured, as we believe that our Raindrop Clarity dataset will benefit the progress of the raindrop removal field.

| Input | Ground Truth | IDT [40] | RDiffusion [21] | Restormer [46] | DiT [22] |

Fig. 11. Comparison of methods on daytime background-focused images, covering both raindrop algorithms and restoration backbones.

4 Experiments

Experimental Setup. Our benchmark includes two parts: Daytime and Nighttime Raindrop Removal. As shown in Table 1, our daytime raindrop dataset contains a total of 5,442 paired/triplets images, where 4,713 paired images are used for the training set, and the remaining 729 paired images are regarded as the test set. Our nighttime raindrop dataset consists of 9,744 paired/triplets images, where 8,655 paired images are selected as the training set and the rest 1,089 paired images are used for evaluation.

Baselines. To create the benchmark, we retrain existing state-of-the-art methods on our daytime and nighttime raindrop datasets. Our benchmark consists of five raindrop algorithms [21, 24–26, 40] and three restoration methods [22, 36, 46]. Specifically, IDT [40], RDiffusion [21], RainDS [26], Robot Car [24] and AtGAN [25] are raindrop algorithms, while Restormer [46], Uformer [36] and DiT [22] are restoration backbones. For a fair comparison, we conduct experiments using the original implementation of these methods. All experiments are implemented in PyTorch and are trained on four NVIDIA RTX A5000 GPUs.

Comparison with Existing Datasets. The difference between our Raindrop Clarity dataset and the existing datasets is shown in Table. 1 and Fig. 2.

The stereo RobotCar dataset [24] includes 4,818 image pairs, randomly selected for training, validation, and testing. However, the raindrops in this dataset are disproportionately large and lack realism.

The raindrop dataset proposed by Qian *et al.* [25] consists of 861 training pairs and 58 testing pairs. While it features common raindrop shapes, it lacks the

Input   Ground Truth   IDT [40]   RDiffusion [21]   Restormer [46]   DiT [22]

Fig. 12. Comparison of methods on nighttime background-focused images, covering both raindrop algorithms and restoration backbones.

diversity of real-world scenes [44]. Also, the backgrounds between the raindrop images and the ground truth are not perfectly aligned [39].

The RainDS-Real dataset, proposed by Quan *et al.* [26], comprises 248 quadruplets, including raindrops, rain streaks, and their corresponding ground truths. The RainDS-Synthetic dataset has a domain gap with the real raindrop images. Moreover, the real-world raindrops are consistently small, obscuring only a fraction of the image details, making recovery less challenging.

The Windshield dataset, proposed by Soboleva and Shipitko [30], was captured by a camera attached to the vehicle during its movement in urban areas and highway environments. The Windshield does not provide the corresponding clean pairs, making this dataset not available for robust raindrop removal method testing. Additionally, frequent moving objects further complicate the dataset.

Existing raindrop datasets, though valuable, have certain limitations. They do not include raindrop-focused and nighttime raindrops. In addition, we have noticed reflection errors in the datasets from Qian [25] and RainDS-Real [26].

Comparison with Synthetic Raindrops. In Fig. 13, we show the existing state-of-the-art synthetic physical model [23], although it can generate raindrops on blurry and clear backgrounds (Fig. 13 d generated on b, Fig. 13 e generated on c). It still has a gap with real-world collected raindrops. Real-world collected raindrops display diverse appearances, including various shapes such as round, elliptical, and long strips. Also, the transparency and glare effects of raindrops collected in the real-world varies.

Quantitative Comparison. To effectively evaluate the results on our proposed datasets, we use three popular metrics PSNR, SSIM [37], and LPIPS [48] in generation tasks [3–5, 9–16, 19, 20, 31–35, 42, 43].

The qualitative results are shown in Table 2. For the daytime dataset, the PSNR and SSIM of the original input images are 21.92 and 0.560. Raindrop algorithm IDT [40] achieves the best performance (26.05 PSNR of and 0.736 of SSIM) compared with other raindrop algorithms. For the restoration backbones, DiT [22] achieves a PSNR of 26.03 and a SSIM of 0.752. Restormer [46] achieves a PSNR of 26.08 and a SSIM of 0.748, which also performs competitively.

It can be observed that, for the nighttime dataset, PSNR and SSIM of the input images are 24.78 and 0.726. Among the raindrop algorithms, IDT [40] achieves better performance than other methods, with a PSNR of 26.81 and a SSIM of 0.851. For the restoration backbones, Restormer [46] achieves the highest PSNR of 26.87, while DiT [22] achieves the lowest LPIPS of 0.111, which outperforms other backbones.

Qualitative Comparison. Figs. 9-12 show the qualitative results of state-of-the-art methods. To be specific, Fig. 9 and Fig. 11 show the results of daytime raindrop removal. The input images in Fig. 9 and Fig. 11 are daytime raindrop-focused and background-focused, respectively. Fig. 10 and Fig. 12 show the results of nighttime raindrop-focused and background-focused removal, respectively.

It can be found that, for daytime and nighttime raindrop-focused removal in Fig. 9 and Fig. 10, existing state-of-the-art methods fail to remove raindrops and recover the background. This is because most of them are designed for background-focused scenes and thus neglect the blurry problem. To our knowledge, Restormer [46] and Uformer [36] are specifically designed to address deblurring problem. However, they still struggle with recovering details and textures. As a result, these methods are struggling to simultaneously remove raindrops and recover the background.

In Fig. 11, for daytime background-focused inputs, we can observe that existing state-of-the-art methods can handle most raindrops. However, they are struggling to address long strips of raindrops. This is because their methods may not detect long strips of raindrops and thus fail to remove them. Furthermore, in Fig. 12, we find that existing state-of-the-art methods are struggling to handle nighttime background-focused samples. To be specific, raindrops in these samples cannot be totally removed and details of the restored images are not clear. This is because nighttime raindrops usually reflect artificial light and thus

Various Shapes, Glare, Transparency, Blur, Upside-down

a Raindrop-focused $\tilde{\mathbf{x}}$ b Blurry background $\tilde{\mathbf{x}}$ c Clear background $\mathbf{b}_0$ d Synthetic Rd on $\tilde{\mathbf{x}}$ e Synthetic Rd on $\mathbf{b}_0$

Fig. 13. We show the differences between the synthetic raindrops [23] and real-world raindrops from Raindrop Clarity.

are more complex than daytime raindrops. Meanwhile, existing state-of-the-art methods are designed for daytime raindrop removal and neglect the effect of complex nighttime conditions. Thus, existing state-of-the-art methods cannot achieve promising results.

5 Conclusion

In this paper, we propose a new raindrop dataset, named Raindrop Clarity. It includes a total of 15,186 high-quality pairs/triplets of images with raindrops and the corresponding clear background images. Moreover, our dataset is collected under various conditions, including different scenes, times (daytime and nighttime), and focuses (raindrop- and background-focused). To be specific, 3,606 daytime raindrop-focused, 1,836 daytime background-focused, 4,838 nighttime raindrop-focused and 4,906 nighttime background-focused images are collected. With the proposed dataset, we re-analyze existing state-of-the-art methods, including raindrop algorithms and restoration backbones, and point out several open problems: raindrop-focus raindrop removal and nighttime raindrop removal. We wish to draw the attention of the community to the open problems.

6 Social Impacts

Existing raindrop datasets focus on daytime background-focused raindrop removal. However, in the real world, raindrop images may focus on raindrops, resulting in a blurry background. Also, raindrops may occur in nighttime images. To draw the attention of the community to the open problems, we propose a new raindrop dataset, named Raindrop Clarity. It mainly collected from four conditions: daytime raindrop-focused, daytime background-focused, nighttime raindrop-focused and nighttime background-focused. Experimental results on our proposed dataset show that existing state-of-the-art methods cannot effectively remove raindrops under complex conditions. As a result, we believe that further exploration of our Raindrop Clarity dataset can help the community to address more complex raindrops.

Acknowledgments

References

1. [EB/OL], `http://liuyunfei.net/Projs/PBRR/`, title = PBRR: Physically Based Raindrop Rendering, author = Yunfei Liu, Zhixiang Hao, Shadi You, Yu Li, and Feng Lu
2. Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., Ma, S., Xu, C., Xu, C., Gao, W.: Pre-trained image processing transformer. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 12299–12310 (2021)
3. Chen, S., Ye, T., Bai, J., Chen, E., Shi, J., Zhu, L.: Sparse sampling transformer with uncertainty-driven ranking for unified removal of raindrops and rain streaks. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 13106–13117 (2023)
4. Chen, S., Ye, T., Liu, Y., Chen, E.: Snowformer: Context interaction transformer with scale-awareness for single image desnowing. arXiv preprint arXiv:2208.09703 (2022)
5. Chen, S., Ye, T., Liu, Y., Liao, T., Jiang, J., Chen, E., Chen, P.: Msp-former: Multi-scale projection transformer for single image desnowing. In: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5. IEEE (2023)
6. Eigen, D., Krishnan, D., Fergus, R.: Restoring an image taken through a window covered with dirt or rain. In: Proceedings of the IEEE international conference on computer vision. pp. 633–640 (2013)
7. Halimeh, J.C., Roser, M.: Raindrop detection on car windshields using geometric-photometric environment construction and intensity-based correlation. In: 2009 IEEE Intelligent Vehicles Symposium. pp. 610–615. IEEE (2009)
8. Hao, Z., You, S., Li, Y., Li, K., Lu, F.: Learning from synthetic photorealistic raindrop for single image raindrop removal. In: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. pp. 0–0 (2019)
9. Jin, Y., Lin, B., Yan, W., Yuan, Y., Ye, W., Tan, R.T.: Enhancing visibility in nighttime haze images using guided apsf and gradient adaptive convolution. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 2446–2457 (2023)
10. Jin, Y., Sharma, A., Tan, R.T.: Dc-shadownet: Single-image hard and soft shadow removal using unsupervised domain-classifier guided network. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5027–5036 (2021)
11. Jin, Y., Yang, W., Tan, R.T.: Unsupervised night image enhancement: When layer decomposition meets light-effects suppression. In: European Conference on Computer Vision. pp. 404–421. Springer (2022)

12. Jin, Y., Ye, W., Yang, W., Yuan, Y., Tan, R.T.: Des3: Adaptive attention-driven self and soft shadow removal using vit similarity. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 2634–2642 (2024)

13. Li, B., Li, X., Zhu, H., Jin, Y., Feng, R., Zhang, Z., Chen, Z.: Sed: Semantic-aware discriminator for image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2024)

14. Li, X., Jin, X., Lin, J., Liu, S., Wu, Y., Yu, T., Zhou, W., Chen, Z.: Learning disentangled feature representation for hybrid-distorted image restoration. In: ECCV. pp. 313–329. Springer (2020)

15. Li, X., Li, B., Jin, X., Lan, C., Chen, Z.: Learning distortion invariant representation for image restoration from a causality perspective. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1714–1724 (2023)

16. Li, X., Ren, Y., Jin, X., Lan, C., Wang, X., Zeng, W., Wang, X., Chen, Z.: Diffusion models for image restoration and enhancement–a comprehensive survey. arXiv preprint arXiv:2308.09388 (2023)

17. Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R.: Swinir: Image restoration using swin transformer. arXiv preprint arXiv:2108.10257 (2021)

18. Liu, X., Suganuma, M., Sun, Z., Okatani, T.: Dual residual networks leveraging the potential of paired operations for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7007–7016 (2019)

19. Luo, J., Ceylan, D., Yoon, J.S., Zhao, N., Philip, J., Frühstück, A., Li, W., Richardt, C., Wang, T.: Intrinsicdiffusion: Joint intrinsic layers from latent diffusion models. In: ACM SIGGRAPH 2024 Conference Papers. SIGGRAPH '24, Association for Computing Machinery, New York, NY, USA (2024). `https://doi.org/10.1145/3641519.3657472`, `https://doi.org/10.1145/3641519.3657472`

20. Luo, J., Zhao, N., Li, W., Richardt, C.: Crefnet: Learning consistent reflectance estimation with a decoder-sharing transformer. IEEE Transactions on Visualization and Computer Graphics (2023)

21. Özdenizci, O., Legenstein, R.: Restoring vision in adverse weather conditions with patch-based denoising diffusion models. IEEE Transactions on Pattern Analysis and Machine Intelligence (2023)

22. Peebles, W., Xie, S.: Scalable diffusion models with transformers. arXiv preprint arXiv:2212.09748 (2022)

23. Pizzati, F., Cerri, P., de Charette, R.: Physics-informed guided disentanglement in generative networks. IEEE Transactions on Pattern Analysis and Machine Intelligence (2023)

24. Porav, H., Bruls, T., Newman, P.: I can see clearly now: Image restoration via deraining. In: 2019 International Conference on Robotics and Automation (ICRA). pp. 7087–7093. IEEE (2019)

25. Qian, R., Tan, R.T., Yang, W., Su, J., Liu, J.: Attentive generative adversarial network for raindrop removal from a single image pp. 2482–2491 (2018)

26. Quan, R., Yu, X., Liang, Y., Yang, Y.: Removing raindrops and rain streaks in one go. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9147–9156 (2021)

27. Quan, Y., Deng, S., Chen, Y., Ji, H.: Deep learning for seeing through window with raindrops. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2463–2471 (2019)

28. Roser, M., Geiger, A.: Video-based raindrop detection for improved image registration. In: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops. pp. 570–577. IEEE (2009)
29. Roser, M., Kurz, J., Geiger, A.: Realistic modeling of water droplets for monocular adherent raindrop recognition using bezier curves. In: Asian conference on computer vision. pp. 235–244. Springer (2010)
30. Soboleva, V., Shipitko, O.: Raindrops on windshield: Dataset and lightweight gradient-based detection algorithm. In: 2021 IEEE Symposium Series on Computational Intelligence (SSCI). pp. 1–7. IEEE (2021)
31. Wang, C., Pan, J., Lin, W., Dong, J., Wang, W., Wu, X.M.: Selfpromer: Self-prompt dehazing transformers with depth-consistency. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 5327–5335 (2024)
32. Wang, C., Pan, J., Wang, W., Dong, J., Wang, M., Ju, Y., Chen, J.: Promptrestorer: A prompting image restoration method with degradation perception. Advances in Neural Information Processing Systems **36**, 8898–8912 (2023)
33. Wang, C., Pan, J., Wang, W., Fu, G., Liang, S., Wang, M., Wu, X.M., Liu, J.: Correlation matching transformation transformers for uhd image restoration. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 5336–5344 (2024)
34. Wang, J., Pan, Z., Zhang, M., Tan, R.T., Li, H.: Restoring speaking lips from occlusion for audio-visual speech recognition. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 38, pp. 19144–19152 (2024)
35. Wang, J., Qian, X., Zhang, M., Tan, R.T., Li, H.: Seeing what you said: Talking face generation guided by a lip reading expert. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14653–14662 (2023)
36. Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., Li, H.: Uformer: A general u-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 17683–17693 (2022)
37. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
38. Wei, M., Shen, Y., Wang, Y., Xie, H., Wang, F.L.: Raindiffusion: When unsupervised learning meets diffusion models for real-world image deraining. arXiv preprint arXiv:2301.09430 (2023)
39. Wen, Q., Wu, Y., Chen, Q.: Video waterdrop removal via spatio-temporal fusion in driving scenes. arXiv preprint arXiv:2302.05916 (2023)
40. Xiao, J., Fu, X., Liu, A., Wu, F., Zha, Z.J.: Image de-raining transformer. IEEE Transactions on Pattern Analysis and Machine Intelligence pp. 1–18 (2022)
41. Yasarla, R., Patel, V.M.: Uncertainty guided multi-scale residual learning-using a cycle spinning cnn for single image de-raining. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8405–8414 (2019)
42. Ye, T., Chen, S., Bai, J., Shi, J., Xue, C., Jiang, J., Yin, J., Chen, E., Liu, Y.: Adverse weather removal with codebook priors. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 12653–12664 (2023)
43. Ye, T., Zhang, Y., Jiang, M., Chen, L., Liu, Y., Chen, S., Chen, E.: Perceiving and modeling density for image dehazing. In: European conference on computer vision. pp. 130–145. Springer (2022)
44. You, S., Tan, R.T., Kawakami, R., Ikeuchi, K.: Adherent raindrop detection and removal in video. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1035–1042 (2013)

45. You, S., Tan, R.T., Kawakami, R., Mukaigawa, Y., Ikeuchi, K.: Adherent raindrop modeling, detection and removal in video. IEEE transactions on pattern analysis and machine intelligence **38**(9), 1721–1733 (2015)
46. Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H.: Restormer: Efficient transformer for high-resolution image restoration. In: CVPR (2022)
47. Zhang, K., Li, D., Luo, W., Ren, W.: Dual attention-in-attention model for joint rain streak and raindrop removal. IEEE Transactions on Image Processing **30**, 7608–7619 (2021)
48. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: CVPR (2018)