UAV-Rain1k: A Benchmark for Raindrop Removal from UAV Aerial Imagery

Wenhui Chang[1]   Hongming Chen[1]   Xin He[2]   Xiang Chen[3*]   Liangduo Shen[4*]

[1]Shenyang Aerospace University   [2]Naval Aviation University
[3]Nanjing University of Science and Technology   [4]Zhejiang Ocean University

Abstract

Raindrops adhering to the lens of UAVs can obstruct visibility of the background scene and degrade image quality. Despite recent progress in image deraining methods and datasets, there is a lack of focus on raindrop removal from UAV aerial imagery due to the unique challenges posed by varying angles and rapid movement during drone flight. To fill the gap in this research, we first construct a new benchmark dataset for removing raindrops from UAV images, called UAV-Rain1k. In this paper, we provide a dataset generation pipeline, which includes modeling raindrop shapes using Blender, collecting background images from various UAV angles, random sampling of rain masks and etc. Based on the proposed benchmark, we further present a comprehensive evaluation of existing representative image deraining algorithms, and reveal future research opportunities worth exploring. The proposed dataset is publicly available at https://github.com/cschenxiang/UAV-Rain1k.

1. Introduction

As unmanned aerial vehicles (UAVs) are increasingly employed for tasks such as surveillance, mapping, and environmental monitoring, the quality of captured imagery becomes paramount. Aerial images are susceptible to degradation in clarity, color distortion, and loss of information due to adverse weather conditions and other factors during the acquisition process [20]. For example, during rainy weather, raindrops adhering to the lens of UAVs can significantly degrade image quality, hindering visibility and affecting the accuracy of downstream image processing tasks. Thus, developing effective raindrop removal techniques tailored to UAV imagery is essential for enhancing the reliability and usability of UAV-based applications [2, 18, 21].

In recent years, notable progress has been made in image deraining, propelled by the advent of effective image priors and deep learning networks [1, 4]. To better address the
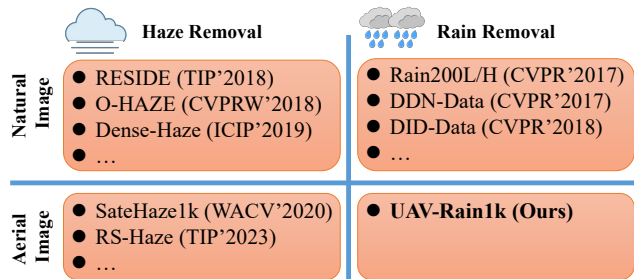

Figure 1. Classification diagram of publicly available datasets for haze removal and rain removal. Our proposed UAV-Rain1k fills the research gap.

problem of removing raindrops, Qian et al. [12] constructed the first raindrop removal dataset, comprising 1,119 pairs of images featuring raindrops against varied backgrounds. These images are captured using a camera equipped with two aligned pieces of glass, one sprayed with water while the other remained clean. Later, Quan et al. [13] presented a mixed rain dataset, namely RainDS, which contains synthetic and real-world raindrops created by manually mimicking rainfall with a sprinkler.

Although the above mentioned datasets [12, 13] have propelled research in raindrop removal, they are all tailored for autonomous driving scenarios. When revisiting research on UAV image processing under adverse weather conditions, most existing studies focus on haze removal, with little attention given to raindrop removal tasks. However, the reality is that raindrop interference on UAV aerial images is also a common scenario in rainy conditions. Compared to autonomous driving scenarios, raindrop effects on UAV aerial imagery pose greater challenges. On one hand, UAVs operate in dynamic environments, which can lead to unpredictable raindrop patterns on lenses or sensors. On the other hand, UAVs often capture images from varying distances and angles, resulting in inconsistent sizes and shapes of raindrops within the images.

To fill the gap in this research, we present a new benchmark for raindrop removal from UAV aerial imagery, called UAV-Rain1k. To the best of our knowledge, this is the first
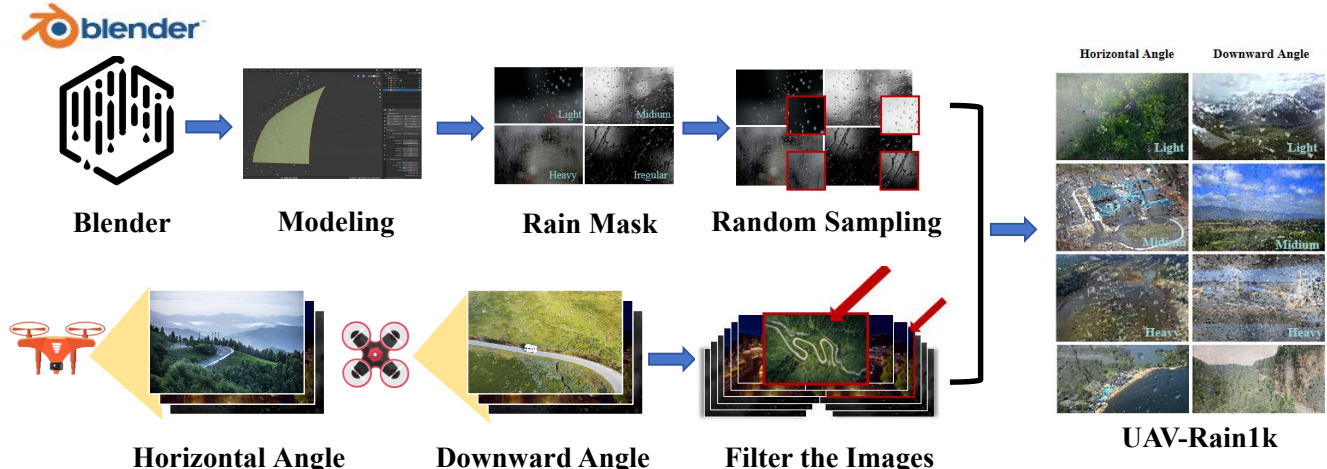
---

*Corresponding author.

Figure 2. Illustration of the dataset generation pipeline.

dataset specifically curated for raindrop removal tasks in the field of aerial image processing (see Figure 1). In addition, we further conduct a comprehensive evaluation of several state-of-the-art image deraining methods, which are quantitatively and qualitatively evaluated on our new dataset. Our evaluation and analysis demonstrate the performance and limitations of existing approaches, providing promising insights. The proposed UAV-Rain1k dataset is publicly available for research purposes, and we intend to periodically update our benchmarking results for noticeable new algorithms.

The rest of this paper is structured as follows. In Section 2, we provide a detailed description of the pipeline used to construct the dataset. We analyze the performance of existing algorithms on our benchmark in Section 3. Finally, the concluding remarks will be given in Section 4.

2. Dataset Construction

In this section, we describe the process of dataset construction. Since obtaining paired real-world raindrop and rain-free aerial images for the same scene and field of view is not feasible, we synthesize a more diverse dataset of UAV aerial images with raindrops. Due to the diverse scenes, complex objects, and vast spatial coverage in aerial images, the previous methods [9] for synthesizing rainy images are not suitable for aerial imagery. For instance, in the previously synthesized data [12, 13], raindrops are consistently simulated as being captured at a horizontal angle over land. However, in reality, the distribution and size of raindrop adhesions vary with different aerial angles. Therefore, we construct a dataset that takes into account the realistic distribution of raindrops in aerial scenes. Figure 2 shows the dataset generation pipeline. The details will be discussed below.

2.1. Raindrop Generation

As a complex atmospheric process, rainfall can lead to various forms of visibility reduction due to environmental factors such as raindrop size, rain density, and wind speed. The fidelity and intricacy of rain play a crucial role in the synthesis of raindrops [14]. To achieve higher quality and more realistic synthetic aerial images with raindrops, we employ an open-source 3D graphics engine (Blender) to simulate and generate images of raindrops within real aerial scenes for training purposes. This 3D graphics engine is capable of rendering raindrops using a physical motion model, allowing us to set depth information and color values separately in the RGB channels [7], which is facilitated by the Blender plugin known as the Rain Generator.

Inspired by Garg [5], we model the generation of raindrop layers as a motion blur process. The instantaneous shape of a water droplet at time t is represented by the function $r[t, \vartheta, \phi]$, where r is the distance from a point on the droplet's surface to its center, θ is defined as the angle opposite to the direction of the droplet's fall at that point, and ϕ is defined as the angle between that point and the projection of the line of sight onto any plane perpendicular to the direction of fall. As raindrops fall, the influence of aerodynamics and surface tension acting on the raindrop causes rapid shape distortion over time. Mathematically, this motion process can be represented based on the oscillatory characteristics of the falling raindrop:

$$r[t, \vartheta, \phi] = r_0(1 + \sum_m \cos(m\phi)p_m(\theta)), \qquad (1)$$

where r_0 is the undeformed radius, and the factor $\cos(m\phi)$ depends on the droplet size r_0. The function $p_m(\theta)$ describes the time dependent variation of modal shape and amplitude with respect to θ. The droplet's shape at any
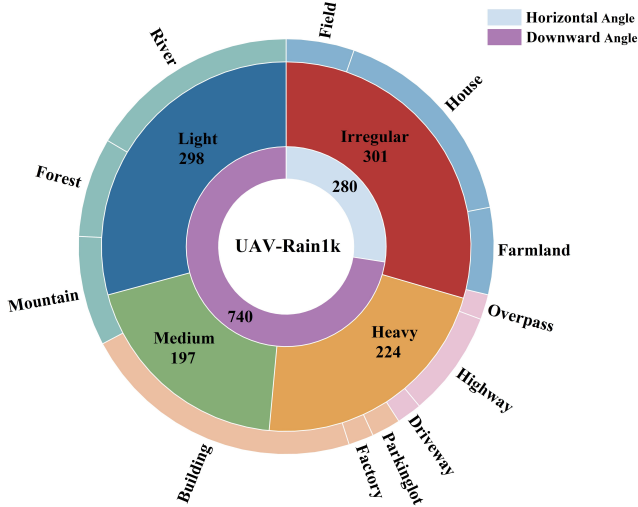
Figure 3. Distribution of rain and scene of the proposed benchmark.

Figure 4. User study results. The ratings given by all participants on different raindrop datasets.

given moment is determined by the cumulative impact of falling time on both modal shape and amplitude.

2.2. Background Collection

The quality of the clear background image, also known as the ground-truth image, is equally crucial when constructing a paired dataset [4, 6, 10]. These images should have higher resolution and rich object details. Unfortunately, existing datasets have placed significant emphasis on rain simulation but have overlooked the need for high-quality background images. Regrettably, the realistic rainy weather images used in previous synthetic datasets were consistently captured from a flat angle on the ground. In reality, there is a significant disparity between the lighting conditions in the air and those on the ground, especially in aerial imagery. Aerial images predominantly feature ground structures as the primary compositional elements in the background, which aligns with human visual perception of rainy scenes. Obtaining rain-free background images that accurately represent the human perception of real rainy conditions is a challenging task, particularly when searching for such images on the Internet. These images also suffer from various issues, including compression artifacts, watermarks, low resolution, defocused blurriness, absence of objects, and etc. These challenges can present difficulties for advanced vision-based remote sensing applications.

Our collection of background images encompasses a wide range of scenes, including urban, natural, and rural environments. Furthermore, we deliberately sought variations in aerial shooting angles during the photography process to capture different scene effects in real scenarios. Consequently, the complexity of these scenes and backgrounds presents both challenges and value to our subset of real data.
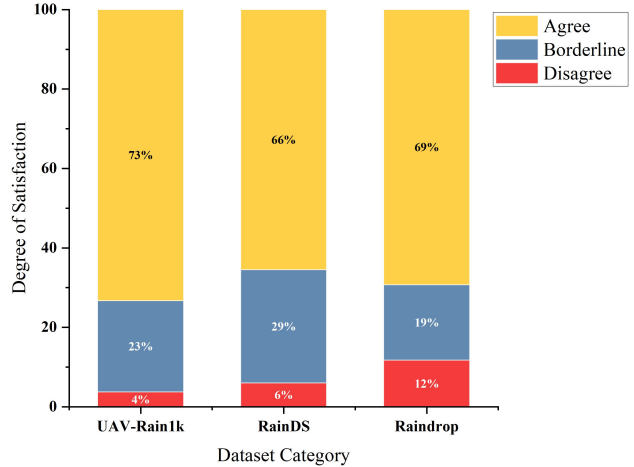
To generate synthetic data, we employ the SMRC-ct13 model unmanned aerial vehicle for on-site shooting. To ensure that the subsequently synthesized rainy images are more realistic and cohesive, we deliberately select overcast and rainy weather conditions for capturing background materials. Additionally, during the shooting process, we diversify our approach by flying the drone in different directions to cover scenes from various angles. Due to constraints in selecting actual shooting locations for aerial images, we complement our dataset by acquiring copyright-free and non-commercial aerial images and videos through Google searches to synthesize rainy weather images. In total, we capture and collect 327 aerial video sequences, comprising approximately 25,000 frames.

To maintain the richness and diversity of our dataset, we carefully select frames from each video sequence that met our criteria, filtering out low-quality images containing issues such as low resolution, website watermarks, compression artifacts, and blurriness. We also exclude background images with excessive lighting or high brightness to ensure appropriate background composition. In summary, we capture and collect footage from over 12 typical scenes, including parking lots, streets, alleys, playgrounds, courtyards, and forests. The solar storm diagram as shown in Figure 3.

2.3. Image Composition

In [12], a method is employed where a piece of glass with attached water droplets is inserted in front of the drone camera lens to simulate the presence of raindrops. However, during the process of capturing aerial rain images, due to the complexity and specificity of capturing scenes with moving perspectives, we are unable to obtain perfectly consistent rainy and rain-free images through aerial
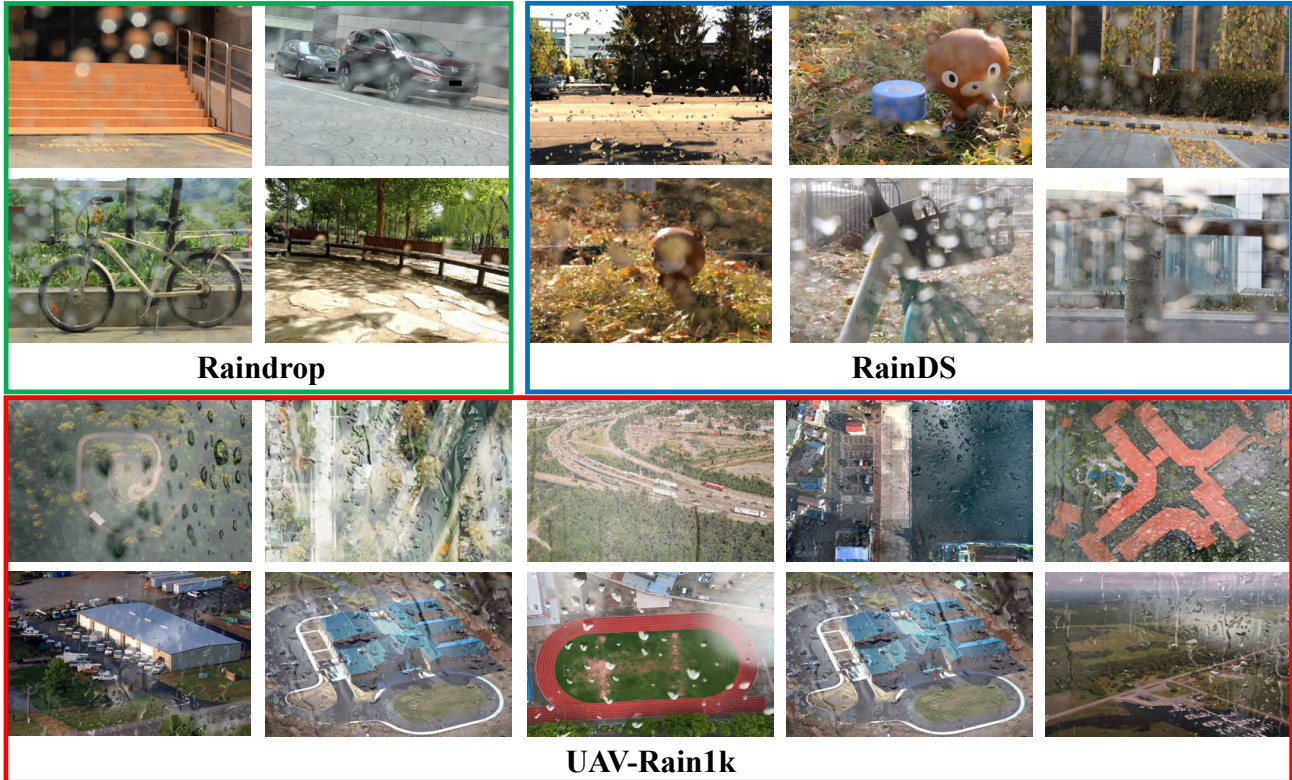
Figure 5. Example images from previous representative datasets (Raindrop [12] and RainDS [13]) and our proposed more challenging UAV-Rain1k.

Table 1. Quantitative comparisons on the UAV-Rain1k benchmark dataset. "#FLOPs" and "#Params" represent FLOPs (in G) and the number of trainable parameters (in M), respectively.

Methods	Input	DSC [11]	RCDNet [16]	SPDNet [22]	Restormer [23]	IDT [19]	DRSformer [3]
Category	-	Prior	CNN	CNN	Transformer	Transformer	Transformer
PSNR	16.80	16.68	22.48	22.54	24.78	22.47	24.93
SSIM	0.7196	0.7142	0.8753	0.8594	0.9054	0.8957	0.9155
#Params	-	-	3.17	3.04	26.12	16.41	33.65
#FLOPs	-	-	21.2	89.3	174.7	61.9	242.9

images. This method also provides some insights for verifying the authenticity of synthetic datasets. We collect real aerial images captured during rainy weather and gather drone footage of raining scenes at different angles and motion states as a reference baseline for synthetic images. This ensures that the collected images exhibit variations in raindrop density, size, and positions, allowing us to filter out low-quality and misleading synthetic data.

Our goal is to achieve visual realism and consistency in synthetic rainy images, with the aim of minimizing the domain gap between synthetic and real images. To achieve this, we do not rely on simple copy-and-paste methods. Instead, we employ a random sampling approach for raindrop masks to ensure that each synthesized rainy image is

unique. To ensure the synthesized images exhibit visual realism and reduce the domain gap between synthesized and real images, inspired by Qian et al. [12], we model the blurred or occluded effects D of raindrops in dispersed small-sized locally coherent regions, combined with a clean background image B, to form the degraded raindrop image R_d. Mathematically, the synthesis process can be expressed as follows:

$$R_d = (1 - M) \bigodot B + D, \qquad (2)$$

where D represents the occlusion or blurring effects caused by raindrops, and B is a binary mask. If $M(x) = 1$, then the pixel x in the mask belongs to the raindrop region, otherwise, it is part of the background image. As a result, we

seamlessly integrate the raindrop mask with the background image. Consequently, compared to other dataset generation methods, our proposed dataset tends to be more complex and diverse, enhancing its ability to generalize across various scenarios.

2.4. Benchmark Statistics

As a result, we propose a new benchmark for raindrop removal from UAV aerial imagery, called UAV-Rain1k. In total, the training and testing set of the UAV-Rain1k contains 800 and 220 synthetic images, respectively. The average resolution of all images is 1500×1000. Our dataset comprises four rain density labels (e.g., light, moderate, heavy, and irregular). In Figure 3, the inner ring displays labels for each rain density, corresponding to a specific quantity of synthesized images. Furthermore, we categorize the background images according to aerial shooting angles and aerial scenes. Consequently, our dataset demonstrates greater diversity and a larger scale in terms of both rain conditions and ground truth (GT) images. Our emphasis goes beyond just the sequence number and total frames of real rain. We also prioritize the diversity of real rain patterns and the authenticity of background images. This focus has a significant impact on the model's ability to generalize to real rain scenarios.

Based on the aforementioned work, we have summarized three innovative aspects of the UAV-Rain1k dataset:

1) A more comprehensive and reliable background collection effort. Encompassing a diverse range of scenes, while simultaneously considering the complexity and uniqueness of capturing scenes from dynamic perspectives.

2) More authentic raindrop shapes. We model the generation of the raindrop layer as a motion blur process and employed an open source 3D rendering engine for the depiction of raindrops, achieving higher quality and more realistic raindrop masks.

3) A more natural synthesis effect. Integrating authentic and diverse raindrop masks seamlessly with background images covering various scenes to achieve a more realistic visual effect. Figure 5 provides some visual examples.

2.5. Subjective Assessment

We conduct an online user study to evaluate the quality of the synthesis rainy images. Here, we prepare for 90 rainy images, randomly chosen from 3 datasets (Raindrop [12], RainDS [13] and UAV-Rain1k) with 30 samples from each dataset. We recruit 40 participants, consisting of 20 males and 20 females. Each participant is shown a random set of 150 rainy images. Then, employing a 3-point Likert scale (i.e., agree, neutral, disagree), all participants assess the perceived realism of each image. In the end, each category receives 600 ratings, which are used to assess the quality of the synthetic dataset. Subjective scores are shown in Figure 4, where our UAV-Rain1k consistently outperforms the other two benchmarks. This also indicates that our synthesized rain is assessed as more realistic compared to previous datasets. Note that we do not standardize the scores. Therefore, what matters here is the score ranking rather than the absolute score values. Despite the relatively small number of evaluators, we observe good consensus and small inter-rater differences in the same paired comparison results, making the scores reliable.

3. Algorithm Benchmarking

In this section, based on the proposed new benchmark, we evaluate 6 representative image deraining algorithms: DSC [11], RCDNet [16], SPDNet [22], Restormer [23], IDT [19], and DRSformer [3]. To ensure fair comparisons, we utilize the officially released codes of these approaches. Each method undergoes retraining on the UAV-Rain1k benchmark, conducted on servers equipped with NVIDIA GeForce RTX 3090 GPUs.

3.1. Quantitative Evaluation

Table 1 presents the quantitative performance evaluation results of various algorithms on the UAV-Rain1k dataset. Here, we utilize two widely referenced image quality assessment metrics, PSNR (Peak Signal-to-Noise Ratio) [8] and SSIM (Structural Similarity Index) [17], to measure the restoration quality of different methods. Additionally, we assess the model efficiency of each approach by considering parameters and FLOPs (Floating Point Operations per Second).

It is evident from the results that DRSformer [3] and Restormer [23] achieve the top two quantitative outcomes in rain removal performance. However, it is noteworthy that their model complexity is comparatively higher than traditional CNN methods [16, 22]. To comprehensively evaluate the performance and efficiency of different algorithms, future research endeavors could delve deeper into strategies for maintaining high performance while reducing model complexity. This would address the requirements of speed and resource consumption in real-world remote sensing applications [2].

3.2. Qualitative Evaluation

Figure 6 and Figure 7 illustrates the visual comparison results of different baseline algorithms on our proposed benchmark (In Figure 6, the effect corresponds to a composite image taken from a horizontal angle, while Figure 7 represents a composite image captured from an downward angle). In comparison to CNN-based methods (RCDNet
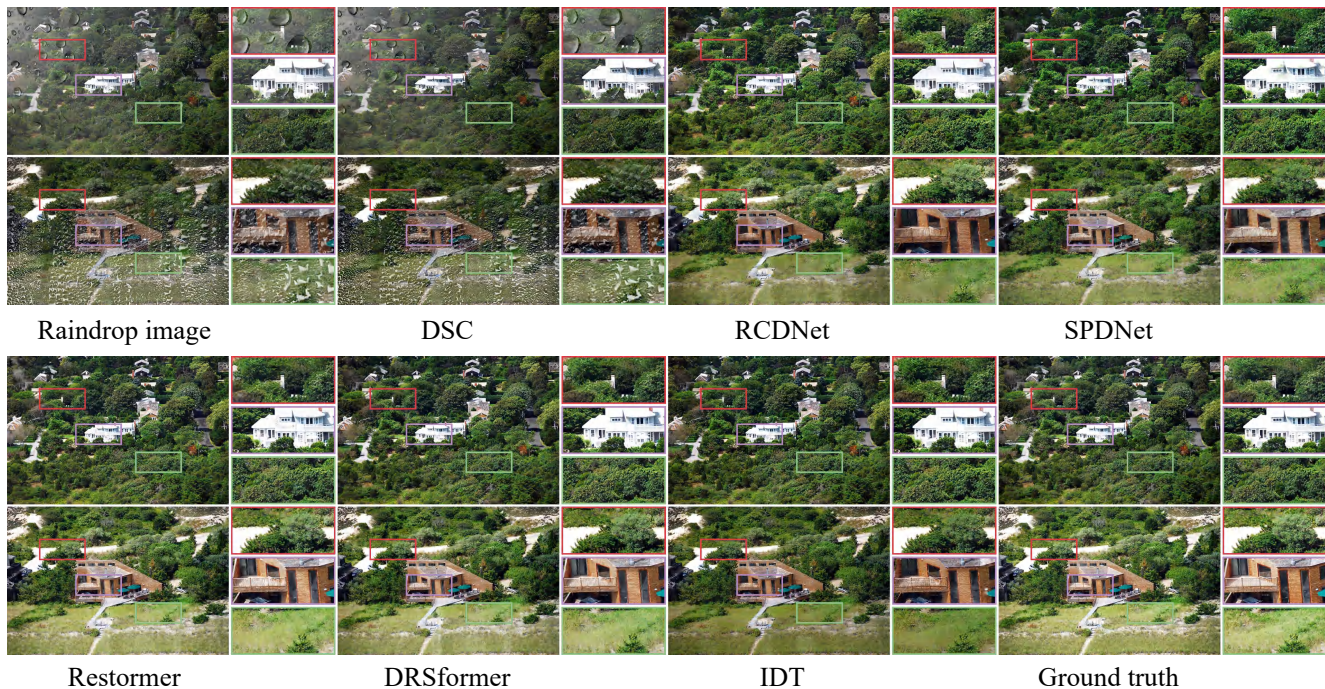
Figure 6. Visual quality comparison on the UAV-Rain1k dataset (horizontal angle). Zooming in the figures offers a better view at the deraining capability.
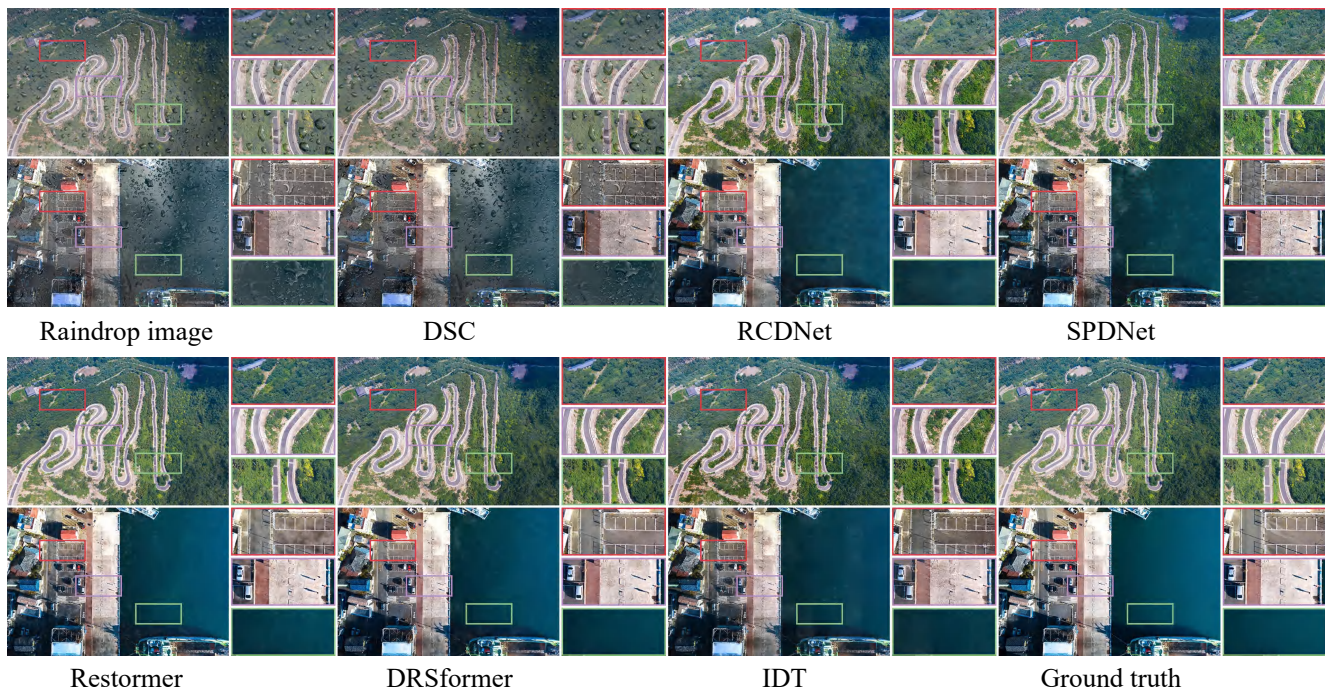

Figure 7. Visual quality comparison on the UAV-Rain1k dataset (downward angle). Zooming in the figures offers a better view at the deraining capability.

[16] and SPDNet [22]), Transformer-based approaches exhibit superior capability in removing undesired raindrop

and recovering clearer visual outcomes, it achieves remarkable visual effects from any aerial perspective, eliminat-

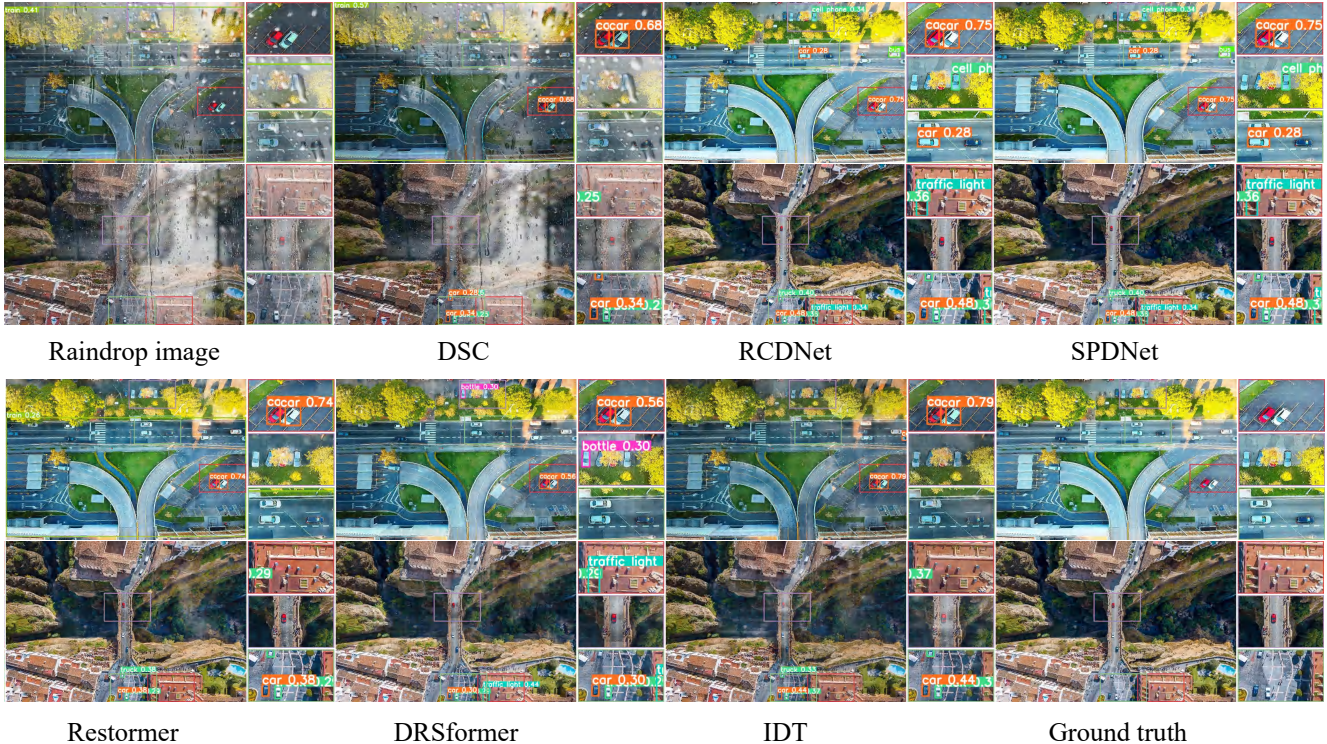| Raindrop image	DSC	RCDNet	SPDNet
Restormer	DRSformer	IDT	Ground truth

Figure 8. Visualization of the object detection results after applying different deraining algorithms.

ing raindrops and obscuring effects without rain residues and pseudo artifacts. owing to the evident advantage of Transformers in modeling non-local information. Additionally, we observe grid artifacts in the restoration results of IDT [19], indicating their limited capacity in handling high-resolution UAV imagery. Furthermore, nearly all methods exhibit poor performance in local detail regions, as they tend to remove image details while removing raindrop, suggesting that there are still rooms for improvement.

3.3. Application-Based Evaluation

To investigate whether the raindrop removal process benefits downstream vision-based remote sensing applications, we apply the popular object detection pre-trained model (YOLOv8 [15]) to evaluate the deraining results. Figure 8 shows the visualization of the object detection results after applying different deraining methods. As one can see, the detection precision of the deraining results by all approaches demonstrates varying degrees of improvement compared to the input raindrop images, especially in the recognition of vehicles on the road. While removing the raindrop regions, some existing methods may also remove the semantic information, disturbing human understanding and degrading high-level vision algorithms. Future works could integrate low-level and high-level tasks to ensure the former contributes to the latter.

4. Conclusion

In this paper, we have proposed a new benchmark dataset for raindrop removal from UAV aerial imagery. We have provided a detailed overview of the synthesis process for the UAV-Rain1k dataset, covering background collection, raindrop modeling, image synthesis, results demonstration, and subjective evaluation. Based on the proposed benchmark, we further experimentally validated representative image deraining algorithms in three aspects: quantitative evaluation, qualitative evaluation, and application-based evaluation. Experimental results shed light on the limitations of existing methods and point towards promising future directions. This paper not only pioneers a new research direction in UAV image processing but also advocates for the community to propose more effective algorithms.

References

[1] Xiang Chen, Yufeng Huang, and Lei Xu. Multi-scale attentive residual dense network for single image rain removal. In *ACCV*, 2020. 1

[2] Xiang Chen, Jinshan Pan, Kui Jiang, Yufeng Li, Yufeng Huang, Caihua Kong, Longgang Dai, and Zhentao Fan. Unpaired deep image deraining using dual contrastive learning. In *CVPR*, pages 2017–2026, 2022. 1, 5

[3] Xiang Chen, Hao Li, Mingqiang Li, and Jinshan Pan. Learn-

ing a sparse transformer network for effective image deraining. In *CVPR*, pages 5896–5905, 2023. 4, 5

[4] Xiang Chen, Jinshan Pan, Jiangxin Dong, and Jinhui Tang. Towards unified deep image deraining: A survey and a new benchmark. *arXiv preprint arXiv:2310.03535*, 2023. 1, 3

[5] Kshitiz Garg and Shree K Nayar. Photorealistic rendering of rain streaks. *ACM TOG*, 25(3):996–1002, 2006. 2

[6] Yun Guo, Xueyao Xiao, Yi Chang, Shumin Deng, and Luxin Yan. From sky to the ground: A large-scale benchmark and simple baseline towards real rain removal. In *ICCV*, pages 12097–12107, 2023. 3

[7] Yufeng Huang, Xiang Chen, Lei Xu, and Kaiyuan Li. Single image desmoking via attentive generative adversarial network for smoke detection process. *Fire Technology*, pages 1–20, 2021. 2

[8] Quan Huynh-Thu and Mohammed Ghanbari. Scope of validity of psnr in image/video quality assessment. *Electronics Letter*, 44(13):800–801, 2008. 5

[9] Siyuan Li, Iago Breno Araujo, Wenqi Ren, Zhangyang Wang, Eric K Tokuda, Roberto Hirata Junior, Roberto Cesar-Junior, Jiawan Zhang, Xiaojie Guo, and Xiaochun Cao. Single image deraining: A comprehensive benchmark analysis. In *CVPR*, pages 3838–3847, 2019. 2

[10] Wei Li, Qiming Zhang, Jing Zhang, Zhen Huang, Xinmei Tian, and Dacheng Tao. Toward real-world single image deraining: A new benchmark and beyond. *arXiv preprint arXiv:2206.05514*, 2022. 3

[11] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *ICCV*, pages 3397–3405, 2015. 4, 5

[12] Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image. In *CVPR*, pages 2482–2491, 2018. 1, 2, 3, 4, 5

[13] Ruijie Quan, Xin Yu, Yuanzhi Liang, and Yi Yang. Removing raindrops and rain streaks in one go. In *CVPR*, pages 9147–9156, 2021. 1, 2, 4, 5

[14] Yuhui Quan, Shijie Deng, Yixin Chen, and Hui Ji. Deep learning for seeing through window with raindrops. In *ICCV*, pages 2463–2471, 2019. 2

[15] Gang Wang, Yanfei Chen, Pei An, Hanyu Hong, Jinghu Hu, and Tiange Huang. Uav-yolov8: a small-object-detection model based on improved yolov8 for uav aerial photography scenarios. *Sensors*, 23(16):7190, 2023. 7

[16] Hong Wang, Qi Xie, Qian Zhao, and Deyu Meng. A model-driven deep neural network for single image rain removal. In *CVPR*, pages 3103–3112, 2020. 4, 5, 6

[17] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004. 5

[18] Yanyan Wei, Zhao Zhang, Yang Wang, Mingliang Xu, Yi Yang, Shuicheng Yan, and Meng Wang. Deraincyclegan: Rain attentive cyclegan for single image deraining and rain-making. *IEEE TIP*, 30:4788–4801, 2021. 1

[19] Jie Xiao, Xueyang Fu, Aiping Liu, Feng Wu, and Zheng-Jun Zha. Image de-raining transformer. *CVPR*, 2022. 4, 5, 7

[20] Mang Ye, Jianbing Shen, Gaojie Lin, Tao Xiang, Ling Shao, and Steven CH Hoi. Deep learning for person re-identification: A survey and outlook. *IEEE TPAMI*, 44(6): 2872–2893, 2021. 1

[21] Yuntong Ye, Yi Chang, Hanyu Zhou, and Luxin Yan. Closing the loop: Joint rain generation and removal via disentangled image translation. In *CVPR*, pages 2053–2062, 2021. 1

[22] Qiaosi Yi, Juncheng Li, Qinyan Dai, Faming Fang, Guixu Zhang, and Tieyong Zeng. Structure-preserving deraining with residue channel prior guidance. In *ICCV*, pages 4238–4247, 2021. 4, 5, 6

[23] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *CVPR*, pages 5728–5739, 2022. 4, 5