UniRain: Unified Image Deraining with RAG-based Dataset Distillation and Multi-objective Reweighted Optimization
- Supplemental Material -

Qianfeng Yang[1,2] Qiyuan Guan[1,2] Xiang Chen[2] Jiyu Jin[1*] Guiyue Jin[1] Jiangxin Dong[2]
[1]Dalian Polytechnic University [2]Nanjing University of Science and Technology
https://lowlevelcv.com/ https://github.com/QianfengY/UniRain

Overview

In this document, we first introduce the deraining datasets in Section 1. Then, we present the details of the RAG-based dataset distillation pipeline in Section 2, followed by a validation of VLM reliability in Section 3. Next, we provide an overview of the constructed dataset in Section 4. The architecture and implementation details of UniRain are described in Section 5. Finally, we report additional evaluations on extended benchmarks and deraining models in Section 6, followed by more experimental results on both synthetic and real-world rainy images in Section 7.

1. Deraining Datasets

To better evaluate the image deraining methods, lots of image deraining datasets have been proposed. Table 1 presents an overview of the existing datasets for single image deraining, including synthetic and real-world datasets. In total, the publicly available deraining training and testing datasets contain 2,386,565 images, demonstrating that the community has already accumulated a very large amount of rainy data. However, despite the large quantity, the quality and realism vary significantly across datasets. As illustrated in Figure 1, synthetic datasets typically exhibit low realism, diverse scene coverage, and over-simplified rain patterns. These datasets are often generated using procedural streak rendering, patch-based raindrops, or simplified accumulation simulation, resulting in unrealistic photometric effects and domain gaps with real rainy scenes. In contrast, real-world datasets possess high realism and faithful rain appearance, yet their scene diversity is limited due to hardware constraints, restricted capture conditions, and the difficulty of collecting paired ground truth. Moreover, the rain in real datasets tends to be subtle, lacking sufficiently strong degradation variety. To address these limitations, we distill a new dataset from the entire corpus of existing synthetic and real data. The distilled dataset exhibits high realism, rich scene diversity, and complex rain degradations, effectively bridging the gap between synthetic controllability and real-world authenticity. This distilled collection provides a more challenging and representative benchmark for evaluating unified deraining models, and serves as a stronger training resource for robust rain removal across diverse environments.

2. Details of RAG-based Dataset Distillation

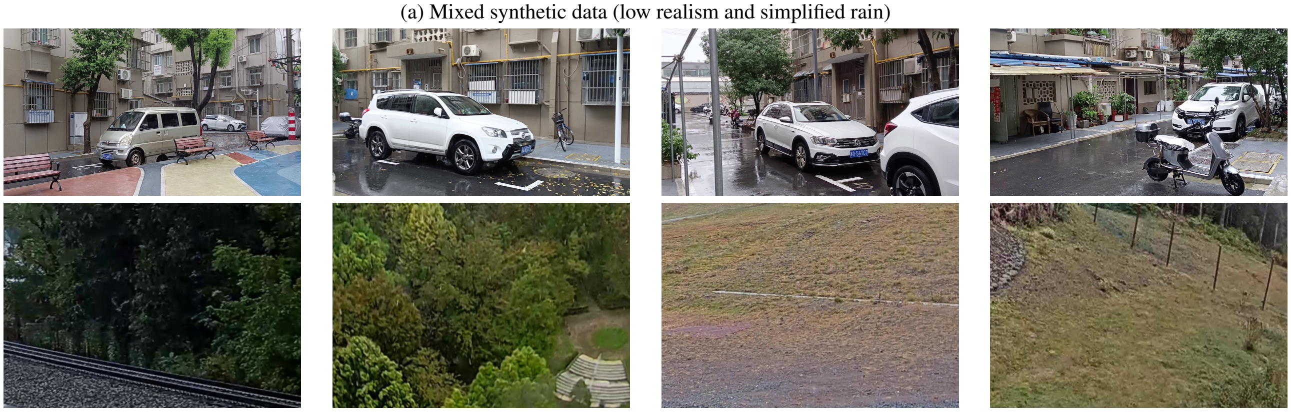
The RAG-based dataset distillation pipeline consists of two stages: a retrieval stage that identifies real-world reference images, and a generation stage that evaluates data quality and selects reliable samples for training. Figures 2a and 2b provide examples of these two stages. Before retrieval, we implement a proactive quality assurance step to ensure the reliability of reference data. Specifically, mixed real-world data is filtered using IQA models (e.g., DepictQA [39]) combined with manual selection to extract high-quality samples. This process ensures that the retrieval stage operates on a curated pool of reliable references, effectively serving as a high-quality proxy. For Figure 2a, each query image is first processed by BLIP [18] to generate an initial caption, which is then expanded with Qwen2.5 [13] for richer semantics. KeyBERT [7] extracts representative keywords to form the final caption for retrieval. CLIP [26] extracts visual features and keyword-based text features from the query image, while database descriptions are encoded with MiniLM [32] to achieve more consistent semantic alignment across different granularities.

*Corresponding author.

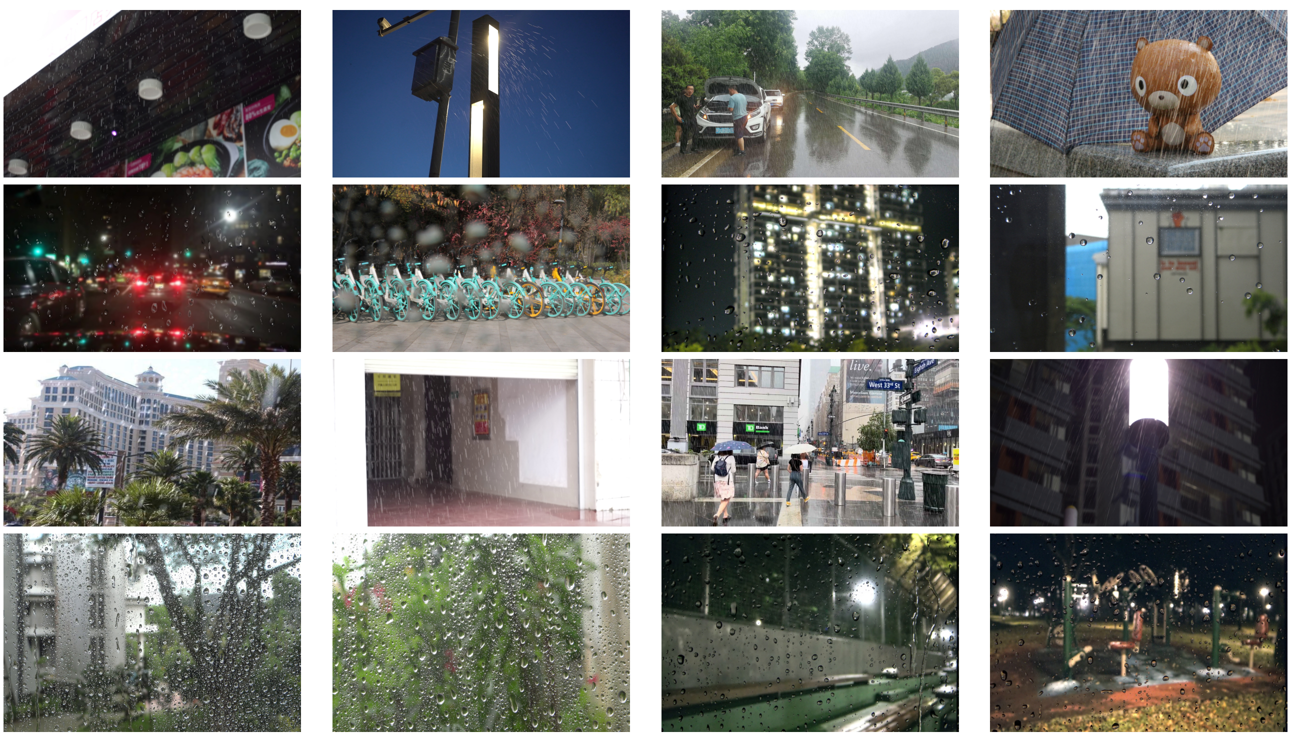
Table 1. Summarization of public datasets for single image deraining task. “Syn”and “Real” is the synthetic and real-world rainy datasets.

Datasets	Number	Syn/Real	Category	Venue	Paired
Rain12 [22]	0/12	Syn	Day Rain Streaks	CVPR 2016	✔
Rain100L [38]	1,800/100	Syn	Day Rain Streaks	CVPR 2017	✔
Rain100H [38]	1,800/100	Syn	Day Rain Streaks	CVPR 2017	✔
Rain200L [38]	1,800/200	Syn	Day Rain Streaks	CVPR 2017	✔
Rain200H [38]	1,800/200	Syn	Day Rain Streaks	CVPR 2017	✔
DDN-Data [6]	12,600/1,400	Syn	Day Rain Streaks	CVPR 2017	✔
DID-Data [45]	12,000/1,200	Syn	Day Rain Streaks	CVPR 2018	✔
RainDrop [24]	861/249	Syn	Day Raindrops	CVPR 2018	✔
Rain800 [20]	700/100	Syn	Day Rain Streaks	TCSVT 2019	✔
SPA-Data [31]	638,492/1,000	Real	Day Rain Streaks/Day Raindrops	CVPR 2019	✔
MPID [20]	3,961/419	Syn+Real	Day Rain Streaks	CVPR 2019	✔
RainCityscapes [12]	9,432/1,188	Syn	Day Rain Streaks	CVPR 2019	✔
Outdoor-Rain [19]	9,000/1,500	Syn	Day Rain Streaks	CVPR 2019	✔
Rain13K [14]	13,712/4,298	Syn	Day Rain Streaks	CVPR 2020	✔
RainKITTI2012 [46]	4,062/4,085	Syn	Day Rain Streaks	ECCV 2020	✔
RainKITTI2015 [46]	4,200/4,189	Syn	Day Rain Streaks	ECCV 2020	✔
IVIPC-DQA [35]	206	Real	Day Rain Streaks	TCSVT 2020	✗
RainDS [25]	3,450/900	Syn+Real	Day Rain Streaks/Day Raindrops	CVPR 2021	✔
RainDirection [23]	2,920/430	Syn	Day Rain Streaks	CVPR 2021	✔
Real3000 [23]	2,700/300	Real	Day Rain Streaks	ICCV 2021	✗
GT-RAIN [1]	28,217/2,100	Real	Day Rain Streaks	ECCV 2022	✔
RealRain-1K [21]	1,568/224	Real	Day/Night Rain Streaks	arXiv 2022	✔
SynRain-13K [21]	13,711/1200	Syn	Day Rain Streaks	arXiv 2022	✔
GTAV-NightRain [43]	11,860/1186	Syn	Night Rain Streaks	arXiv 2022	✔
VRDS [34]	7,200/3,000	Syn	Day Rain Streaks/Day Raindrops	ACM MM 2023	✔
LWDDS [33]	67,500/600	Syn	Day Raindrops	ICRA 2023	✔
LHP-Rain [11]	1,000,000/1,000	Real	Day Rain Streaks	ICCV 2023	✔
UAV-Rain1k [2]	800/220	Syn	Day Raindrops	CVPR 2024	✔
RoadScene-rain [29]	181/40	Syn	Night Raindrops	TMM 2024	✔
RaindropClarity [15]	13,368/1,818	Syn	Day/Night Raindrops	ECCV 2024	✔
HQ-NightRain [8]	10,000/300	Syn	Night Rain Streaks/Night Raindrops	NeurIPS 2025	✔
WeatherBench-Rain [9]	14,729/200	Real	Day/Night Rain Streaks	ACM MM 2025	✔
FoundIR-Rain [16]	204,789/300	Syn+Real	Day Rain Streaks/Raindrops	ICCV 2025	✔
GenDeg-Rain [27]	239,441/1298	Syn+Real	Day Rain Streaks/Raindrops	CVPR 2025	✔
CDD-11-Rain [10]	3,549/800	Syn	Day/Night Rain Streaks	CVPR 2025	✔
URIR-8K [37]	7,200/800	Syn+Real	Day Rain Streaks/Raindrops	AAAI 2025	✔

(a) Mixed synthetic data (low realism and simplified rain)



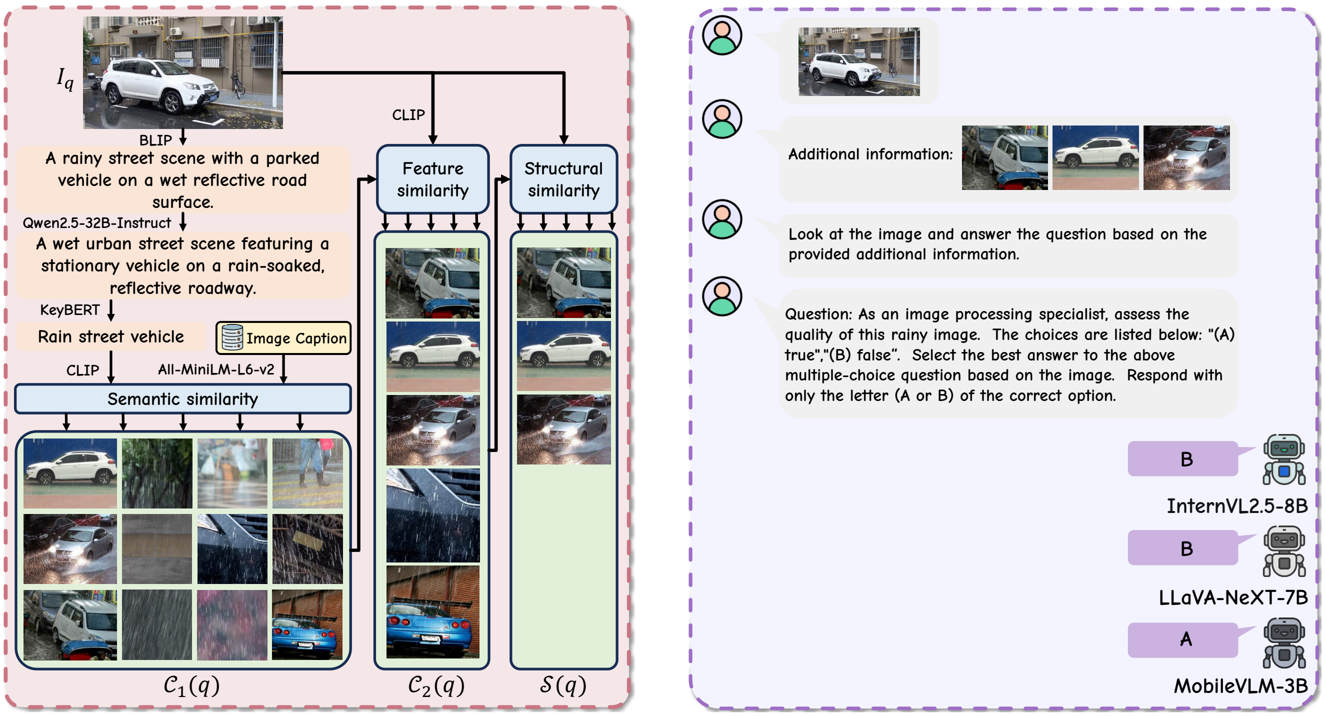
(b) Mixed real-world data (limited scene and subtle rain)



(c) Our distilled data (high realism, diverse scene, and complex rain)

Figure 1. Visualization of rainy images from different mixed datasets.

Our distillation pipeline adopts a coarse-to-fine hierarchical reference retrieval strategy. Specifically, we first leverage text-label similarity for efficient semantic-level filtering, where L2 distance is used to obtain the top-K candidate samples. This is followed by CLIP-based cosine similarity to further enforce visual consistency, particularly in terms of rain patterns. Finally, structural similarity (SSIM) is employed for pixel-level refinement, selecting the most consistent reference images. This progressive matching strategy ensures accurate and reliable retrieval of high-quality reference samples. For the generation stage shown in Figure 2b, we follow the same procedure described in the main paper. After obtaining the retrieved set, we combine it with the query image and a predefined prompt template as input to a vision–language model (VLM) to assess sample reliability. To improve robustness, we employ an ensemble of three VLMs (InternVL2.5-8B, LLaVA-NeXT-7B, and MobileVLM-3B), each producing a binary decision, with the final result determined by majority voting.



(a) Retrieval stage (b) Generation stage

Figure 2. Example of the retrieval stage and the generation stage.

3. Validation of VLMs Reliability

We conduct a user study to measure human–VLMs agreement on a representative sample subset. The results in the Table 2 show the consistency between VLMs scores and human preferences, showing the reliability of the VLMs evaluation metric.

Table 2. Human–VLMs agreement.

Pairs (Distilled/Discarded)	Users	Human–VLMAgreement (%)
20	43	90

4. Dataset Overview

Table 3 provides a detailed overview of our proposed RainRAG dataset, including the number of images in each subset. The dataset covers four types of rainy conditions: daytime rain streaks (DRS), daytime raindrops (DRD), nighttime rain streaks (NRS), and nighttime raindrops (NRD). For each subset, we report the number of paired images in both the training and testing sets, offering a comprehensive summary of the dataset composition.

Table 3. Overview of our proposed RainRAG dataset. The dataset includes daytime rain streaks (DRS),daytime raindrops (DRD), nighttime rain streaks (NRS), and nighttime raindrops (NRD).

Subset	Number			
	DRS	DRD	NRS	NRD
Training set	24,864 (pairs)	6,692 (pairs)	10,249 (pairs)	11,087 (pairs)
Testing set	100 (pairs)	100 (pairs)	100 (pairs)	100 (pairs)

5. Details of the UniRain

As illustrated in Figure 3a, the overall pipeline adopts a U-shaped architecture with an asymmetric encoder-decoder design [40]. The degraded input image is first processed by a 3×3 convolution to extract shallow features, which are subsequently propagated through four hierarchical encoding and decoding stages. Although both stages employ Transformer blocks, we integrate the soft-MoE into the encoder and deploy the hard-MoE within the decoder, as shown in Figure 3b. The structures of both the soft-MoE and hard-MoE modules are detailed in Figure 3c. The encoder consists of four levels containing $[2, 4, 4, 6]$ Transformer blocks [41] from the top layer to the bottom. Each MoE layer contains a set of n nested experts with different capacities, following $R = C/2^i$ for $i \in \{1, \ldots, n\}$. In addition, the number of experts decreases across layers, reducing computational cost while preserving representational diversity. The network's initial embedding dimension is set to $C = 48$, providing a compact yet expressive feature space for subsequent multi-scale processing.

6. Evaluationon Additional Benchmarks and Deraining Models

Table 4 further conducts an evaluation on an additional recent benchmark, FoundIR-R [17] (ICCV 2025), and compare with existing deraining models (i.e., DRSformer [4], MSDT [3], and CST-Net [8] (NeurIPS 2025)) trained on the original SPA-Data dataset versus our RainRAG dataset. The PSNR/SSIM results in the table below indicate that our RainRAG leads to consistent performance gains across diverse benchmarks and models, underscoring its generalizability.

Table 4. Evaluationon additional benchmarks and deraining models.

Training Datasets	MSDT	DRSformer	CST-Net	UniRain (ours)
SPA-Data	24.12/0.7920	22.57/0.7587	26.28/0.8404	**30.41/0.9031**
RainRAG (ours)	30.23/0.9097	31.29/0.9162	30.02/0.9163	**32.56/0.9203**

7. More Experimental Results

This section presents additional experimental results to further validate the effectiveness of our method. Figure 4 shows results on the daytime rain streaks (DRS) subset, where UniRain produces cleaner structures and preserves fine textures across different streak densities. Figure 5 illustrates the daytime raindrop (DRD) subset, demonstrating that UniRain effectively removes adherent droplets while maintaining sharp edges and avoiding over-smoothing. For nighttime degradations, Figure 6 presents comparisons on the nighttime rain streaks (NRS) subset. Despite low illumination, UniRain restores clearer contours and more natural luminance than prior methods. Figure 7 shows the nighttime raindrop (NRD) subset, where UniRain successfully removes raindrop occlusions and recovers structural details in dark environments. Finally, Figure 8 presents results on real-world rain mist scenes. UniRain removes rain artifacts while preserving fine structures, such as the chair legs, demonstrating generalization to complex real-world rain mist conditions.

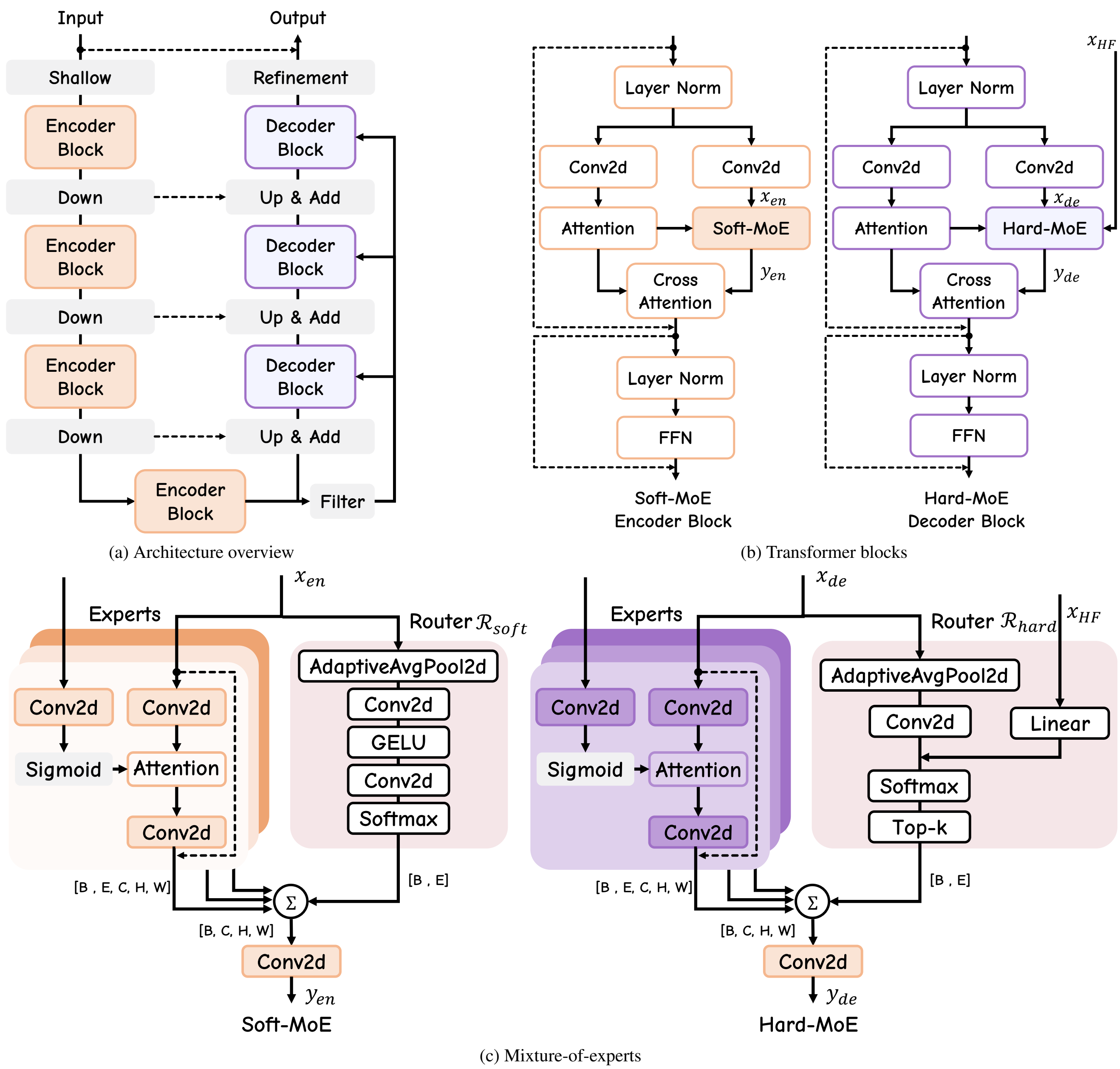


Figure 3. Detailed structure of the proposed UniRain framework.

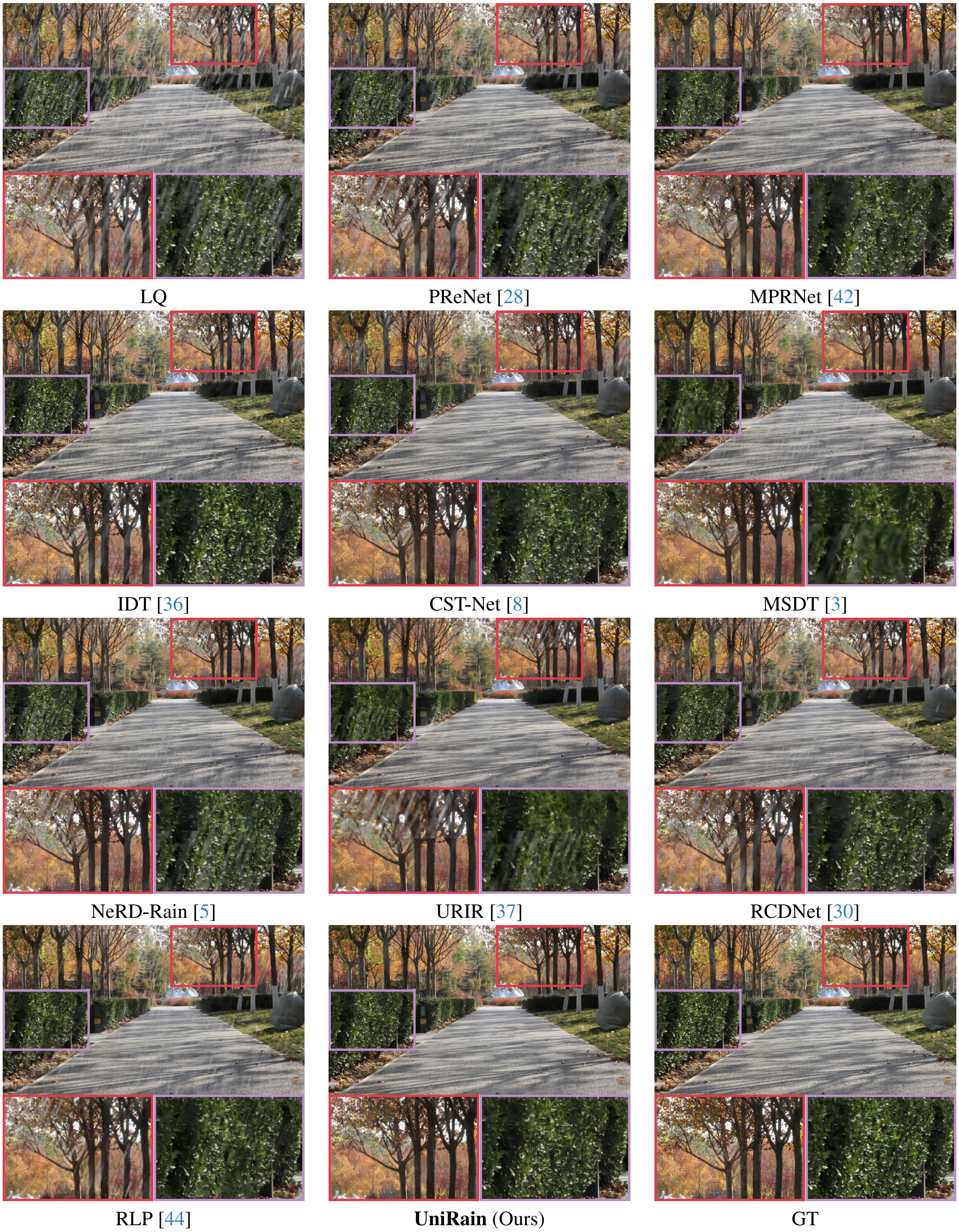
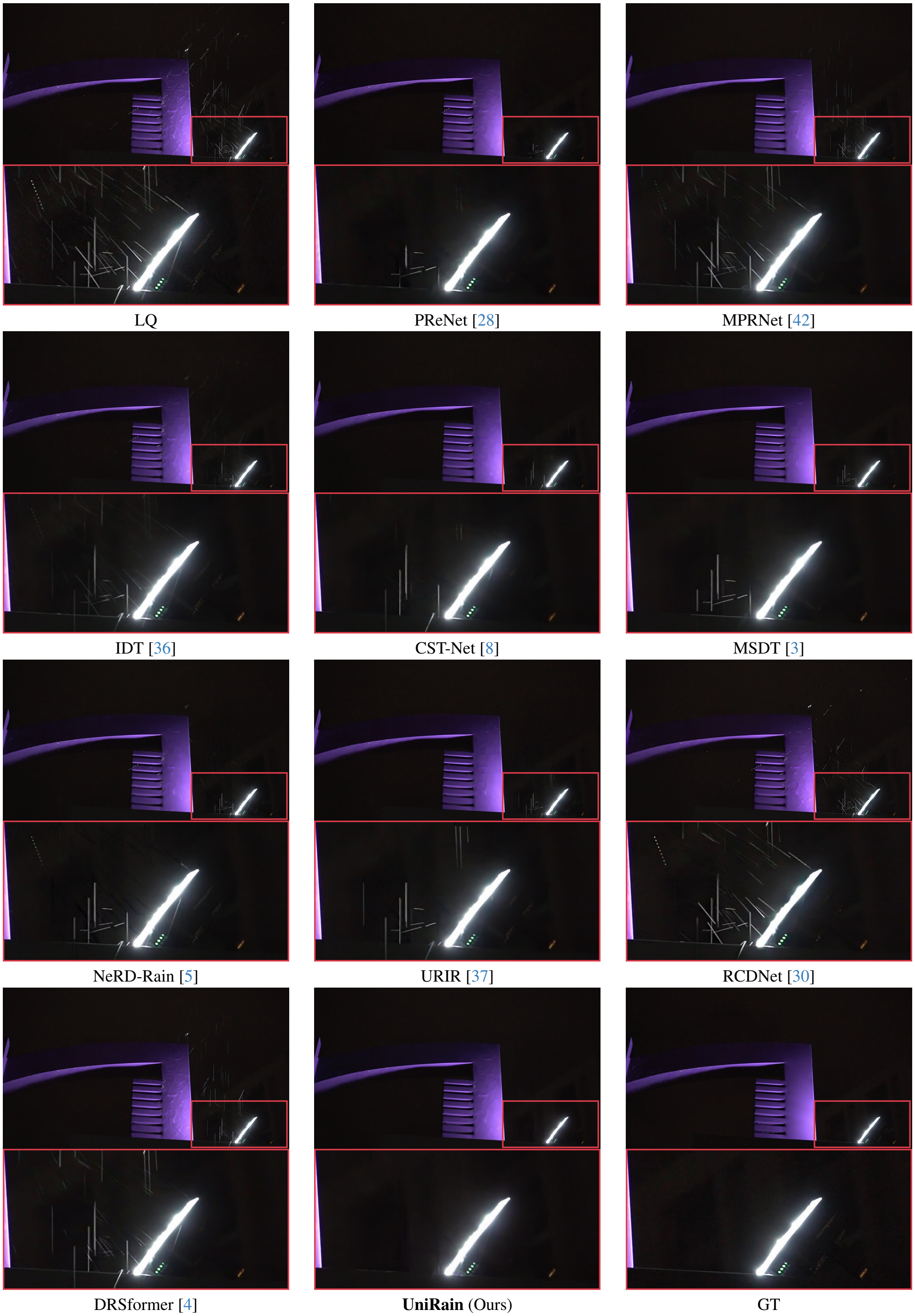


Figure 4. Visual results on the daytime rain streaks subset (DRS). UniRain produces cleaner structures and more faithful texture restoration.

Figure 5. Visual results on the daytime raindrop subset (DRD). UniRain removes adherent raindrops while preserving fine image details.



LQ | PReNet [28] | MPRNet [42]

IDT [36] | CST-Net [8] | MSDT [3]

NeRD-Rain [5] | URIR [37] | RCDNet [30]

DRSformer [4] | **UniRain** (Ours) | GT

Figure 6. Visual results on the nighttime rain streaks subset (NRS). UniRain achieves cleaner removal of rain streaks with fewer artifacts.

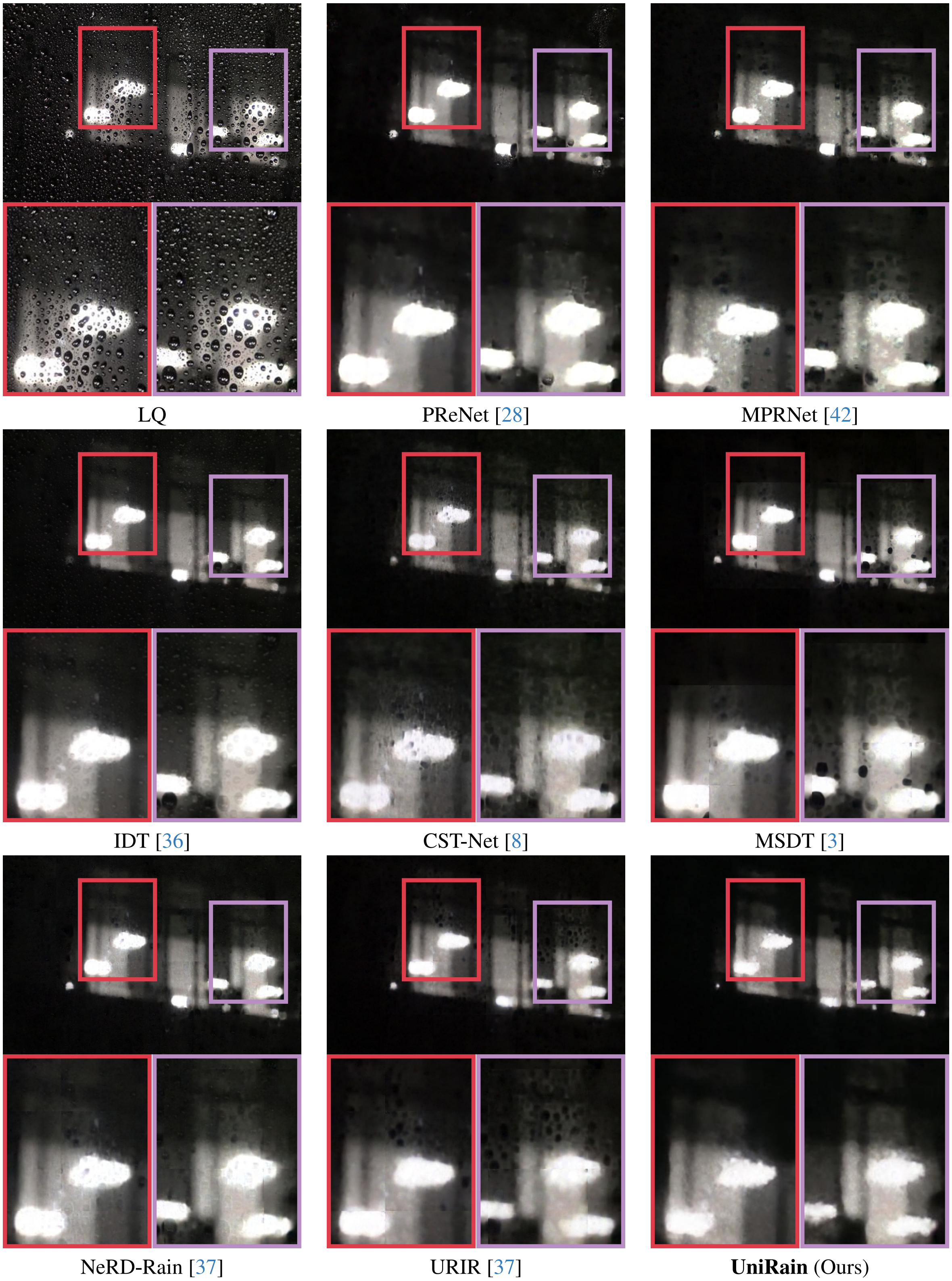


Figure 7. Visual results on the nighttime raindrop subset (NRD). UniRain removes raindrop occlusions in dark environments.

Figure 8. Visual results on the rain mist from MPID [20]. UniRain effectively removes rain artifacts while preserving fine structural details.

References

[1] Yunhao Ba, Howard Zhang, Ethan Yang, Akira Suzuki, Arnold Pfahnl, Chethan Chinder Chandrappa, Celso M De Melo, Suya You, Stefano Soatto, Alex Wong, et al. Not just streaks: Towards ground truth for single image deraining. In *ECCV*, 2022. 2

[2] Wenhui Chang, Hongming Chen, Xin He, Xiang Chen, and Liangduo Shen. Uav-rain1k: A benchmark for raindrop removal from uav aerial imagery. In *CVPR*, 2024. 2

[3] Hongming Chen, Xiang Chen, Jiyang Lu, and Yufeng Li. Rethinking multi-scale representations in deep deraining transformer. In *AAAI*, 2024. 5, 7, 8, 9, 10, 11

[4] Xiang Chen, Hao Li, Mingqiang Li, and Jinshan Pan. Learning a sparse transformer network for effective image deraining. In *CVPR*, 2023. 5, 9

[5] Xiang Chen, Jinshan Pan, and Jiangxin Dong. Bidirectional multi-scale implicit neural representations for image deraining. In *CVPR*, 2024. 7, 8, 9

[6] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *CVPR*, 2017. 2

[7] Maarten Grootendorst. Keybert: Minimal keyword extraction with bert., 2020. 1

[8] Qiyuan Guan, Xiang Chen, Guiyue Jin, Jiyu Jin, Shumin Fan, Tianyu Song, and Jinshan Pan. Rethinking nighttime image deraining via learnable color space transformation. *NeurIPS*, 2025. 2, 5, 7, 8, 9, 10, 11

[9] Qiyuan Guan, Qianfeng Yang, Xiang Chen, Tianyu Song, Guiyue Jin, and Jiyu Jin. Weatherbench: A real-world benchmark dataset for all-in-one adverse weather image restoration. In *ACM MM*, 2025. 2

[10] Yu Guo, Yuan Gao, Yuxu Lu, Ryan Wen Liu, and Shengfeng He. Onerestore: A universal restoration framework for composite degradation. In *ECCV*, 2024. 2

[11] 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*, 2023. 2

[12] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. In *CVPR*, 2019. 2

[13] Binyuan Hui, Jian Yang, Zeyu Cui, Jiaxi Yang, Dayiheng Liu, Lei Zhang, Tianyu Liu, Jiajun Zhang, Bowen Yu, Keming Lu, et al. Qwen2. 5-coder technical report. *arXiv preprint arXiv:2409.12186*, 2024. 1

[14] Kui Jiang, Zhongyuan Wang, Peng Yi, Chen Chen, Baojin Huang, Yimin Luo, Jiayi Ma, and Junjun Jiang. Multi-scale progressive fusion network for single image deraining. In *CVPR*, 2020. 2

[15] Yeying Jin, Xin Li, Jiadong Wang, Yan Zhang, and Malu Zhang. Raindrop clarity: A dual-focused dataset for day and night raindrop removal. In *ECCV*, 2024. 2

[16] Hao Li, Xiang Chen, Jiangxin Dong, Jinhui Tang, and Jinshan Pan. Foundir: Unleashing million-scale training data to advance foundation models for image restoration. In *ICCV*, 2025. 2

[17] Hao Li, Xiang Chen, Jiangxin Dong, Jinhui Tang, and Jinshan Pan. Foundir: Unleashing million-scale training data to advance foundation models for image restoration. In *ICCV*, 2025. 5

[18] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *ICML*, 2022. 1

[19] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *CVPR*, 2019. 2

[20] 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*, 2019. 2, 11

[21] 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. 2

[22] Yu Li, Robby T Tan, Xiaojie Guo, Jiangbo Lu, and Michael S Brown. Rain streak removal using layer priors. In *CVPR*, 2016. 2

[23] Yang Liu, Ziyu Yue, Jinshan Pan, and Zhixun Su. Unpaired learning for deep image deraining with rain direction regularizer. In *ICCV*, 2021. 2

[24] 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*, 2018. 2

[25] Ruijie Quan, Xin Yu, Yuanzhi Liang, and Yi Yang. Removing raindrops and rain streaks in one go. In *CVPR*, 2021. 2

[26] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *ICML*, 2021. 1

[27] Sudarshan Rajagopalan, Nithin Gopalakrishnan Nair, Jay N Paranjape, and Vishal M Patel. Gendeg: Diffusion-based degradation synthesis for generalizable all-in-one image restoration. *arXiv preprint arXiv:2411.17687*, 2024. 2

[28] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. In *CVPR*, 2019. 7, 8, 9, 10, 11

[29] Cidan Shi, Lihuang Fang, Han Wu, Xiaoyu Xian, Yukai Shi, and Liang Lin. Nitedr: Nighttime image de-raining with cross-view sensor cooperative learning for dynamic driving scenes. *IEEE TMM*, 2024. 2

[30] Hong Wang, Qi Xie, Qian Zhao, and Deyu Meng. A model-driven deep neural network for single image rain removal. In *CVPR*, 2020. 7, 8, 9

[31] Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson WH Lau. Spatial attentive single-image deraining with a high quality real rain dataset. In *CVPR*, 2019. 2

[32] Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. *NeurIPS*, 2020. 1

[33] Qiang Wen, Yue Wu, and Qifeng Chen. Video waterdrop removal via spatio-temporal fusion in driving scenes. In *ICRA*, 2023. 2

[34] Hongtao Wu, Yijun Yang, Haoyu Chen, Jingjing Ren, and Lei Zhu. Mask-guided progressive network for joint raindrop and rain streak removal in videos. In *ACM MM*, 2023. 2

[35] Qingbo Wu, Lei Wang, King Ngi Ngan, Hongliang Li, Fanman Meng, and Linfeng Xu. Subjective and objective de-raining quality assessment towards authentic rain image. *IEEE TCSVT*, 2020. 2

[36] Jie Xiao, Xueyang Fu, Aiping Liu, Feng Wu, and Zheng-Jun Zha. Image de-raining transformer. *IEEE TPAMI*, 2022. 7, 8, 9, 10, 11

[37] Hujie Yan. Towards universal rainy image restoration: Benchmark and baseline. In *AAAI*, 2025. 2, 7, 8, 9, 10, 11

[38] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. In *CVPR*, 2017. 2

[39] Zhiyuan You, Zheyuan Li, Jinjin Gu, Zhenfei Yin, Tianfan Xue, and Chao Dong. Depicting beyond scores: Advancing image quality assessment through multi-modal language models. In *ECCV*, 2024. 1

[40] Eduard Zamfir, Zongwei Wu, Nancy Mehta, Yuedong Tan, Danda Pani Paudel, Yulun Zhang, and Radu Timofte. Complexity experts are task-discriminative learners for any image restoration. In *CVPR*, 2025. 5

[41] 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*, 2022. 5

[42] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *CVPR*, 2021. 7, 8, 9, 10, 11

[43] Fan Zhang, Shaodi You, Yu Li, and Ying Fu. Gtav-nightrain: Photometric realistic large-scale dataset for night-time rain streak removal. *arXiv preprint arXiv:2210.04708*, 2022. 2

[44] Fan Zhang, Shaodi You, Yu Li, and Ying Fu. Learning rain location prior for nighttime deraining. In *ICCV*, 2023. 7, 8, 11

[45] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In *CVPR*, 2018. 2

[46] Kaihao Zhang, Wenhan Luo, Wenqi Ren, Jingwen Wang, Fang Zhao, Lin Ma, and Hongdong Li. Beyond monocular deraining: Stereo image deraining via semantic understanding. In *ECCV*, 2020. 2