NTIRE 2025 Challenge on Day and Night Raindrop Removal for Dual-Focused Images: Methods and Results

Xin Li† Yeying Jin† Xin Jin† Zongwei Wu† Bingchen Li† Yufei Wang† Wenhan Yang†
Yu Li† Zhibo Chen† Bihan Wen† Robby T. Tan† Radu Timofte† Qiyu Rong
Hongyuan Jing Mengmeng Zhang Jinglong Li Xiangyu Lu Yi Ren Yuting Liu
Meng Zhang Xiang Chen Qiyuan Guan Jiangxin Dong Jinshan Pan Conglin Gou
Qirui Yang Fangpu Zhang Yunlong Lin Sixiang Chen Guoxi Huang Ruirui Lin Yan Zhang
Jingyu Yang Huanjing Yue Jiyuan Chen Qiaosi Yi Hongjun Wang Chenxi Xie
Shuai Li Yuhui Wu Kaiyi Ma Jiakui Hu Juncheng Li Liwen Pan Guangwei Gao
Wenjie Li Zhenyu Jin Heng Guo Zhanyu Ma Yubo Wang Jinghua Wang Wangzhi Xing
Anjusree Karnavar Diqi Chen Mohammad Aminul Islam Hao Yang Ruikun Zhang
Liyuan Pan Qianhao Luo XinCao Han Zhou Yan Min Wei Dong Jun Chen
Taoyi Wu Weijia Dou Yu Wang Shengjie Zhao Yongcheng Huang Xingyu Han
Anyan Huang Hongtao Wu Hong Wang Yefeng Zheng Abhijeet Kumar
Aman Kumar Marcos V. Conde Paula Garrido Daniel Feijoo Juan C. Benito
Guanglu Dong Xin Lin Siyuan Liu Tianheng Zheng Jiayu Zhong Shouyi Wang
Xiangtai Li Lanqing Guo Lu Qi Chao Ren Shuaibo Wang Shilong Zhang
Wanyu Zhou Yunze Wu Qinzhong Tan Jieyuan Pei Zhuoxuan Li Jiayu Wang
Haoyu Bian Haoran Sun Subhajit Paul Ni Tang Junhao Huang Zihan Cheng
Hongyun Zhu Yuehan Wu Kaixin Deng Hang Ouyang Tianxin Xiao Fan Yang
Zhizun Luo Zeyu Xiao Zhuoyuan Li Pham Hoang Le Nguyen Dinh Thien An
Luu Thanh Son Kiet Van Nguyen Ronghua Xu Xianmin Tian Weijian Zhou
Jiacheng Zhang Yuqian Chen Yihang Duan Yujie Wu Suresh Raikwar Arsh Garg
Kritika Jianhua Zheng Xiaoshan Ma Ruolin Zhao Yongyu Yang Yongsheng Liang
Guiming Huang Qiang Li Hongbin Zhang Xiangyu Zheng A.N. Rajagopalan

Abstract

This paper reviews the NTIRE 2025 Challenge on Day and Night Raindrop Removal for Dual-Focused Images. This challenge received a wide range of impressive solutions, which are developed and evaluated using our collected real-world Raindrop Clarity dataset [33]. Unlike existing deraining datasets, our Raindrop Clarity dataset is more diverse and challenging in degradation types and contents, which includes day raindrop-focused, day background-focused, night raindrop-focused, and night background-focused degradations. This dataset is divided into three subsets for competition: 14,139 images for training, 240 images for validation, and 731 images for testing. The primary objective of this challenge is to establish a new and powerful benchmark for the task of removing raindrops under varying lighting and focus conditions. There are a total of 361 participants in the competition, and 32 teams submitting valid solutions and fact sheets for the final testing phase. These submissions achieved state-of-the-art (SOTA) performance on the Raindrop Clarity dataset. The project can be found at https://lixinustc.github.io/CVPR-NTIRE2025-RainDrop-Competition.github.io/.

†X. Li, Y. Jin, X. Jin, Z. Wu, B. Li, Y. Wang, W. Yang, Y. Li, Z. Chen, B. Wen, R. Tan and R. Timofte are the challenge organizers

The other authors are participants of the NTIRE 2025 Challenge on Day and Night Raindrop Removal for Dual-Focused Images.

The NTIRE2025 website: https://cvlai.net/ntire/2025/

The Competition website https://codalab.lisn.upsaclay.fr/competitions/21345

The Raindrop Clarity database: https://github.com/jinyeying/RaindropClarity

1. Introduction

Image deraining has been a long-standing research topic in low-level image processing [39, 42, 43, 55, 75, 76, 79, 81, 89, 92], aiming to remove visual artifacts caused by rain streaks [84] or raindrops [64] under adverse weather conditions. It plays a crucial role not only in enhancing the perceptual quality of images but also in improving the performance of high-level vision tasks, such as autonomous driving and pedestrian detection. With the rapid development of diverse backbone architectures in low-level image processing [36, 38, 42, 43, 49, 61, 75, 76, 79, 81, 89, 92], several deraining benchmarks have been introduced, incorporating architectures such as ResNet [42, 79, 90], Transformer [4, 82], MLP [75], and Mamba-based frameworks [92]. Moving beyond single-task benchmarks, recent studies have begun to explore multi-task deraining methods [60, 76] or unified frameworks capable of handling various adverse weather conditions simultaneously, such as rain, haze and snow removal. These multi-task deraining methods impose new challenge to the model design, which requires higher adaptability and generalization capabilities. Furthermore, to enhance subjective visual quality [59, 87, 91], diffusion-based deraining methods [7, 25, 44, 60, 70], such as WeatherDiff [60], have recently emerged, demonstrating promising results and opening new opportunities for advanced research on image deraining.

Datasets are essential for evaluating the effectiveness of deraining algorithms. In the early stages, rain degradation was typically synthesized [20, 23, 24, 27, 40, 41, 48, 85, 90], as capturing paired rainy and clean images simultaneously is extremely challenging, even with professional cameras and controlled environments. To address this limitation, recent works have proposed methods that generate real-world rainy degradation by utilizing the video temporal priors [79] or inserting the glass with adherent raindrops [62, 65, 71] and sprayed water [66]. Raindrops, as the typical rain degradation type, usually appear on camera lenses or windshields, which significantly reduces image visibility in human life, posing challenges for applications like surveillance and autonomous driving. Effective raindrop removal is crucial to ensuring reliable performance in these systems. However, most deraining datasets overlooked the complex environment in real life. In particular, there are few raindrop-focused datasets, and most existing datasets focus on capturing background scenes while the camera is focused on the background. Meanwhile, these datasets and related works primarily target daytime scenarios, with limited attention to nighttime conditions [30, 32, 52, 53].

To address the aforementioned challenges, we collaborate with the 2025 New Trends in Image Restoration and Enhancement (NTIRE 2025) workshop to organize the first Challenge on Day and Night Raindrop Removal for Dual-Focused Images. The challenge is based on our Raindrop Clarity Dataset [33], which contains 5,442 daytime and 4,712 nighttime raindrop image pairs or triplets. Its primary objective is to promote research on real-world rain removal under varying lighting conditions while simultaneously restoring fine image textures degraded by raindrop-induced defocus. This competition attracted a total of 361 participants, resulting in 32 teams submitting their methods and results, which are documented in this challenge report. We believe that this challenge will encourage further advancements of research on image deraining.

This challenge is one of the NTIRE 2025 Workshop associated challenges on: ambient lighting normalization [78], reflection removal in the wild [83], shadow removal [77], event-based image deblurring [72], image denoising [73], XGC quality assessment [57], UGC video enhancement [69], night photography rendering [19], image super-resolution (x4) [14], real-world face restoration [15], efficient super-resolution [67], HR depth estimation [88], efficient burst HDR and restoration [35], cross-domain few-shot object detection [21], short-form UGC video quality assessment and enhancement [46, 47], text to image generation model quality assessment [22], day and night raindrop removal for dual-focused images [45], video quality assessment for video conferencing [26], low light image enhancement [58], light field super-resolution [80], restore any image model (RAIM) in the wild [50], raw restoration and super-resolution [17] and raw reconstruction from RGB on smartphones [18].

2. Challenge

The NTIRE 2025 Challenge on Day and Night Raindrop Removal for Dual-Focused Images is the first competition to be organized to advance the development of real-world image draining under different light conditions and focusing degrees. The details of the whole challenge are as follows:

2.1. Datasets

The dataset used in this challenge is the RainDrop Clarity dataset [33], which includes daytime and nighttime scenes for training, validation, and testing. The original dataset includes both daytime and nighttime scenes for training and testing: (i) the daytime raindrop dataset contains a total of 5,442 paired or triplet images, with 4,713 pairs/triplets used for training and the remaining 729 pairs/triplets for testing; Specifically, among the 4,713 daytime training image pairs/triplets, 1,575 are background-focused while 3,138 are raindrop-focused; (ii) the nighttime raindrop dataset consists of 9,744 paired or triplet images, where 8,655 pairs/triplets are allocated to the training set, and the remaining 1,089 pairs/triplets are reserved for the valida-

https://www.cvlai.net/ntire/2025/

tion and test set. Specifically, among the 8,655 nighttime training image pairs/triplets, 4,143 are background-focused while 4,512 are raindrop-focused.

In this challenge, the Raindrop Clarity dataset is divided into training, validation, and testing subsets, where the training set comprises 4,713 triplets, totaling 14,139 images, while the validation and testing sets contain 240 and 731 image, respectively. To ensure diversity and prevent distribution bias, the validation and testing sets are reorganized to maintain a balanced distribution in various rainy scenes. In particular, during the validation and testing phases, intermediate rain-free blurry images are withheld to better simulate real-world application scenarios. Additionally, to rigorously evaluate the robustness of the algorithms submitted, no explicit distinction is made between day- and night-time scenes in the validation and testing subsets. The training set for this challenge is exactly the same as that used in the Raindrop Clarity dataset paper [33]. The validation set of 240 images consists of 120 daytime and 120 nighttime images, each including 60 raindrop-focused and 60 background-focused samples. The test set of 731 images contains 381 daytime images and 350 nighttime images. Among the 381 daytime images, 294 are raindrop-focused and 87 are background-focused. For the 350 nighttime images, 230 are raindrop-focused and 120 are background-focused.

2.2. Evaluation Protocol

This challenge utilizes three metrics to measure the objective and subjective quality of restored images, *i.e.*, PSNR, SSIM, and LPIPS, respectively. The final score used for ranking is computed by reweighting the above metrics as:

$$\mathrm{Score} = 10\times\mathrm{PSNR(Y)}+10\times\mathrm{SSIM(Y)}-5\times\mathrm{LPIPS}, \quad (1)$$

where (Y) denotes the PSNR and SSIM are measured with the Y channel after coverting the image from RGB space to the YCbCr space. For LPIPS, we first normalize the image pixel values into the range $[-1, 1]$, then utilize the Alex network configuration for distance measurement between restored images and ground-truth images. PSNR and SSIM, two widely used metrics in deraining [5, 6, 8, 54, 86] and restoration tasks [28, 29, 31, 34].

2.3. Challenge Phases

There are two phases in this challenge, *i.e.*, the development and testing phases. The details are as follows.

Development Phase: In the development phase, we release 4,713 triplets, totaling 14,139 images in our Raindrop Clarity dataset, including daytime and nighttime rainy images, rain-free blur images, and their corresponding ground-truth images, to support each team in developing their algorithms. Moreover, we release 240 daytime and nighttime rainy images without their ground truth for validation. Each participant can upload their restored images to the challenge platform by removing the raindrop with their developed algorithm. Then they can obtain the corresponding final score, PSNR, SSIM, and LPIPS. In the development phases, we received 1264 submissions from 74 teams in total.

Testing Phases: In the testing phases, we release 731 daytime and nighttime rainy images without their ground truth for testing. To ensure the fairness of this challenge, we hide the leaderboard, where each team cannot find the performance of other teams. The final ranking is achieved with the score in Eq. 1. In the test stage, 75 teams submitted their final results to the challenge platform. At the end of this competition, we received the fact sheets and source codes from 32 teams, which are utilized for final ranking.

3. Challenge Results

We have summarized the challenge results in Table 1. Among all submissions, Team Miracle ranked 1st with the highest final score of 34.3518, achieving the top performance across PSNR (27.6619) and SSIM (0.8087), as well as a competitive LPIPS score (0.2794), despite using a relatively moderate parameter count (26.89M) and no ensemble or extra data. The 2nd place, EntroVision, delivered a final score of 34.2145. It achieved high PSNR (27.6629) and SSIM (0.8071) values, along with a relatively low LPIPS (0.3038). The model benefited from the use of ensemble and extra data. The 3rd place, IIRLab, achieved comparable accuracy with a final score of 33.4940, PSNR of 26.9564, SSIM of 0.7993, and LPIPS of 0.2910. Notably, it accomplished this while maintaining a lightweight design—using only 11.69M parameters and 12.62GFlops, despite employing an ensemble strategy, demonstrating an excellent tradeoff between accuracy and efficiency. Overall, the table reveals a wide diversity in model complexity, ranging from lightweight solutions to highly complex networks, reflecting different trade-offs between performance and resource consumption.

4. Teams and Methods

Please find the details of all **32 teams and their methods** in the **Supplementary**.

4.1. Miracle

This team proposes the STRRNet [68], which is developed based on Restormer [89], as shown in Fig 1. They categorize the training images into four classes based on lighting conditions and raindrop types: night_bg_focus, night_raindrop_focus, day_bg_focus, and day_raindrop_focus. As shown in Fig 2, a text embedder is first trained on the labeled training set using these four categories. Then, they design a semantic guidance module, which is added at the end of the Restormer

Table 1. Quantitative results of the NTIRE 2025 on Day and Night Raindrop Removal for Dual-Focused Images. The best and second results are in red and blue, respectively.

Teams	Leader	Final Score ↑	PSNR↑	SSIM↑	LPIPS ↓	Params. (M)	GFlops (G)	Ensemble	Extra Data	Rank
Miracle	Qiyu Rong	34.3518	27.6619	0.8087	0.2794	26.89	42.33	⊗	⊗	1
EntroVision	Xiang Chen	34.2145	27.6629	0.8071	0.3038	16.6	129.9	✓	✓	2
IIRLab	Conglin Gou	33.4940	26.9564	0.7993	0.2910	11.69	12.62	✓	⊗	3
PolyRain	Jiyuan Chen	32.7910	26.1945	0.7701	0.2210	26	183	✓	⊗	4
H3FC2Z	Kaiyi Ma	32.5652	26.3954	0.7737	0.3133	36.2	120	✓	⊗	5
IIC Lab	Juncheng Li	32.5367	26.1606	0.7517	0.2281	45.5	230	⊗	⊗	6
BUPT CAT	Wenjie Li	32.4091	26.2216	0.7774	0.3172	46.1	139	✓	⊗	7
WIRTeam	Yubo Wang	32.2750	26.0379	0.7714	0.2954	35.84	265.88	⊗	⊗	8
GURain	Wangzhi Xing	32.1223	25.9473	0.7671	0.2993	25.31	87.7	⊗	⊗	9
BIT_ssvgg	Hao Yang	31.9508	25.7955	0.7665	0.3019	68	742	⊗	⊗	10
CisdiInfo-MFDehazNet	Qianhao Luo	31.9346	25.7951	0.7670	0.3062	5	132	⊗	⊗	11
McMaster-CV	Han Zhou	31.8097	25.7445	0.7533	0.2936	11.15	158.32	⊗	⊗	12
Falconi	Taoyi Wu	31.6836	25.4967	0.7506	0.2638	109.59	522	⊗	⊗	13
Dfusion	Yongcheng Huang	31.6098	25.5081	0.7469	0.2735	-	105	⊗	⊗	14
RainMamba	Hongtao Wu	31.4608	25.4251	0.7509	0.2946	34.49	119	⊗	⊗	15
RainDropX	AbhijeetKumar	31.1777	25.1864	0.7518	0.3053	26.1	141	⊗	⊗	16
Cidaut AI	Marcos V. Conde	31.1667	25.3108	0.7456	0.3201	2.38	64.28	⊗	⊗	17
DGL_DeRainDrop	Guanglu Dong	31.0267	24.9668	0.7397	0.2674	129.99	91.03	⊗	⊗	18
xdu_720	Shuaibo Wang	31.0051	25.1503	0.7535	0.3360	148	205	⊗	⊗	19
EdgeClear-DNSST Team	Jieyuan Pei	30.8227	25.0620	0.7384	0.3246	23	175	⊗	⊗	20
MPLNet	Jiayu Wang	30.7730	24.6601	0.7178	0.2130	10.7	82.3296	⊗	⊗	21
Singularity	Subhajit Paul	30.4006	24.4232	0.7256	0.2556	46	760	⊗	⊗	22
VIPLAB	Ni Tang	30.3678	24.3905	0.7203	0.2452	5.97	147	⊗	⊗	23
2077Agent	Kaixin Deng	30.3581	24.5354	0.7209	0.2773	28	-	⊗	⊗	24
X-L	Zeyu Xiao	30.2202	24.5935	0.7474	0.3694	12.63	120	✓	⊗	25
UIT-SHANKS	Pham Hoang Le Nguyen	30.0521	24.5982	0.7361	0.3814	-	-	✓	⊗	26
One Go Go	Ronghua Xu	30.0126	24.1069	0.7241	0.2671	25.31	33	✓	⊗	27
DualBranchDerainNet	Yuqian Chen	29.7762	24.6438	0.7247	0.4229	2.5	-	⊗	⊗	28
QWE	Yihang Duan	29.5937	24.2767	0.6987	0.3340	-	-	⊗	⊗	29
Visual and Signal Information Processing Team	Suresh Raikwar	24.5878	21.8319	0.6018	0.6523	-	-	⊗	⊗	30
The Zheng family group	Jianhua Zheng	23.9370	21.0757	0.5524	0.5326	-	-	-	-	31
RainClear Pioneers	Qiang Li	22.7072	18.3433	0.6521	0.4314	-	-	-	-	32

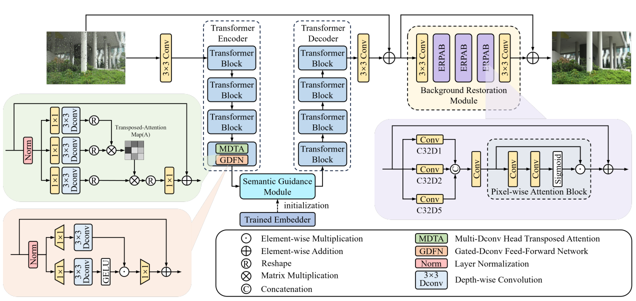

Figure 1. The framework of STRRNet, proposed by Team Miracle

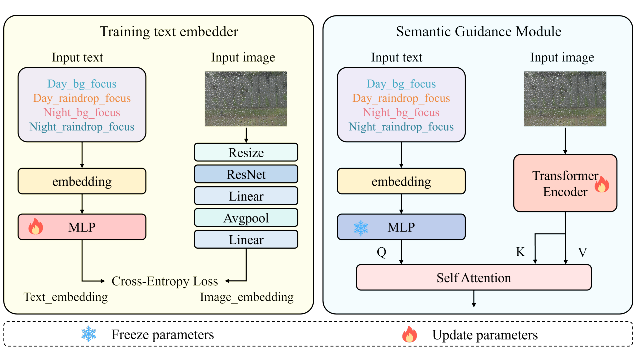

Figure 2. Semantic Guidance Module of STRRNet, proposed by Team Miracle

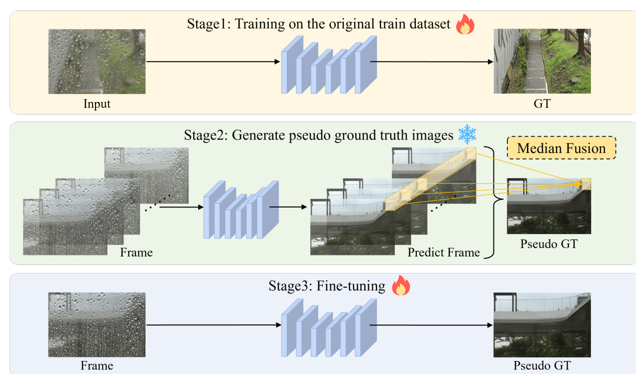

Figure 3. Training Strategy of STRRNet, proposed by Team Miracle

encoder. This module utilizes the encoded image features from Restormer to guide the decoder in performing distinct image restoration operations for the four different types of images. Additionally, they introduce a background restoration subnetwork at the output of Restormer, which consists of multiple convolutional layers to enhance image details. The training strategy of STRRNet is illustrated in the Fig 3. First, they train a pre-trained model on the original training dataset. This model is then used to perform inference on all frames within the same scene in the test set. The inferred images may still contain residual raindrops and artifacts.

Since the background remains consistent across different time frames while raindrop positions vary temporally, this motivates them to perform median fusion across multiple frames from the same scene to obtain a median-fused image. Due to the dynamic nature of raindrop artifacts, their inconsistent locations across frames typically prevent them from appearing in the median values, whereas the stable background information is preserved. They then treat the median-fused image as a pseudo ground truth and use it in a semi-supervised fine-tuning phase to enhance the model's raindrop removal capability on unlabeled images.

Training description The Adam optimizer is used for training, with a total of 500,000 iterations. The learning rate is set to 0.0003 for the first 9,2000 iterations and then gradually decays from 0.0003 to 0.000001 for the remaining iterations. Images are randomly cropped to a fixed size of 128×128 for network training, and geometric image augmentation is applied. The network is optimized using the L1 loss function and the multi-scale SSIM loss function, with weights of 1 and 0.2, respectively. All their experiments were conducted on an RTX 4090.

Testing description Use a sliding window to move across the image, applying the model for rain removal on each window. Set the sliding window size to 128×128 with an overlap of 32. Then, obtain a median image by applying median fusion to the images of the same scene. Finally, perform a weighted sum of the median image and the original image to obtain the final output.

4.2. EntroVision

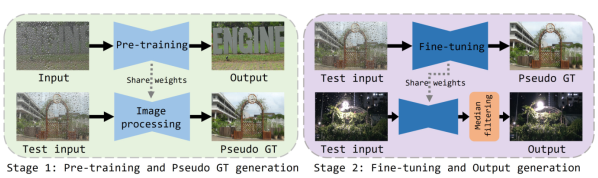

Figure 4. Overview of the technique proposed by Team EntroVision for raindrop removal.

This team utilizes a two-stage approach to achieve raindrop removal in this challenge. The technical overview is shown in the Fig. 4. Given the impact of multi-scale features in image deraining [9, 12], they pre-train the deraining model in the first stage using MSDT [3] on the RainDrop Clarity [33] training set. To enhance the model's generalization capability, additional pre-training is conducted on the UAV-Rain1k [2] dataset. Then, the pre-trained deraining model is utilized to obtain the pseudo-ground-truth images for testing images for the test-time learning of the second stage.

In the second stage of the process, they performed fine-tuning using the test samples and the generated paired pseudo ground-truth images. This approach provides a clear direction for transferring pretrained knowledge, rather than simply relying on the model's generalization ability. Subsequently, they process the testing inputs using this fine-tuned deraining model to obtain the final output results. Notably, they designed the deraining network specifically according to the characteristics of the dataset used. To address scenarios where blurry and clear backgrounds might coexist, they employ median filtering to preserve edge information and avoid excessive blurring. Through the above two-stage processing strategies and the specially designed median filtering technique, they obtain clear deraining results.

4.3. IIRLab

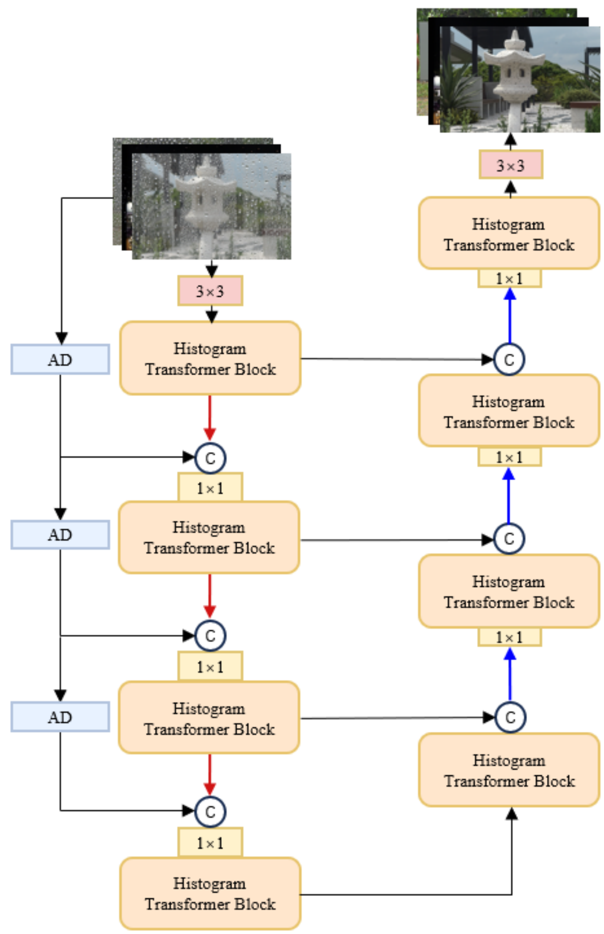

Figure 5. The pipeline of the method proposed by Team IIRLab

This team observed that the raindrop dataset used in this challenge differs significantly from existing raindrop removal datasets. Traditional datasets primarily focus on background clarity during image shooting, resulting in clean backgrounds and blurry raindrops in the foreground.

In such cases, simply removing the raindrops is sufficient to recover a clear background. However, the dataset in this challenge includes a notable portion of images where the camera is focused on the raindrops during shooting, leading to clean raindrops and blurred backgrounds. This introduces a new challenge: removing not only the raindrops themselves but also mitigating background blur caused by defocus, to ultimately recover a clean draining image.

Based on this new question, this team has chosen the Histoformer [74] network as their raindrop removal model, as shown in Fig 5. Histoformer is a transformer-based model designed to restore images degraded by severe weather conditions. It incorporates a histogram self-attention mechanism, which sorts and segments spatial features into intensity-based bins and applies self-attention within each bin. This enables the model to focus on spatial features across dynamic intensity ranges and handle long-range dependencies between similarly degraded pixels. Built upon Restormer [89], Histoformer is well-suited for addressing both raindrop artifacts and defocus blur, making it a strong candidate for this task. Based on its architecture and prior performance, the team chose Histoformer as the core model for this challenge.

Training Details. This team utilized all image pairs provided in the training set, consisting of raindrop-degraded images (Drop) and their corresponding clean background images (Clean). From this, they extracted 1,200 image pairs for validation, including 400 daytime and 800 nighttime samples. For training, they employed a two-stage training strategy that combines regular training with subsequent fine-tuning. In the first stage, the draining model was trained for 300,000 iterations using the default Histoformer configuration. In the second stage, the model was fine-tuned for an additional 13,000 iterations using only the $\mathcal{L}_1$ loss function. This progressive training approach effectively enhanced model performance, leading to improved final results.

Implememtation Details. The implementation is based on PyTorch and was conducted on an NVIDIA RTX 3090 GPU. The network was trained for a total of 300,000 iterations, with an initial batch size of 6 and a patch size of 128×128, following a progressive learning strategy. The team employed the AdamW optimizer with an initial learning rate of 3×10^{-4} for the first 92,000 iterations, which was then gradually reduced to 1×10^{-6} using a cosine annealing schedule over the remaining 208,000 iterations. The number of blocks at each stage was set as $L_{i \in 1,2,3,4} = 4, 4, 6, 8$, and the channel dimension was fixed at $C = 36$. The channel expansion factor in the DGFF module was set to $r = 2.667$. The number of self-attention heads at each stage was configured as $1, 2, 4, 8$, respectively. For data augmentation, horizontal and vertical flips were applied randomly during training.

4.4. PolyRain

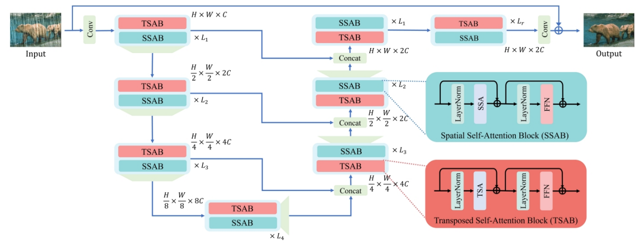

Figure 6. The pipeline of the method proposed by Team PolyRain

This team initialized and trained a dense X-Restormer model with the given dataset based on Restormer [89]. After the first training stage, they finetuned the model on a larger patch size with different loss functions. The whole framework of this team is shown in Fig. 6.

Training Details. The training process consists of two phases. In the first phase, the patch size of the training image is set to 256, with a batch size of 8, and a total of 3×105 training iterations. The learning rate is set to $3e^{-4}$ and $\mathcal{L}_1$ Loss is used as the loss function. In the second phase, the model is fine-tuned with a $5e^{-5}$ learning rate and the image patch with 448×448. During this phase, the model is trained simultaneously using $\mathcal{L}_2$ Loss, LPIPS Loss, and SSIM Loss with weights of 1, 0.1, and 0.1, respectively, for 5×104 iterations.

Testing Details. To enhance the robustness of the model, the self-ensemble technique is employed. The implementation references the BasicSR library .

Implementation Details. This method is implemented based on the famous BasicSR framework [1, 11] written in Python. They utilized the AdamW optimizer with an initial learning rate of $3e^{-4}$. Eight A100 GPUs were used for the model training, lasting for about 72 hours for 300000 iterations. In addition, the CosineAnnealingRestart-CyclicLR scheduler was chosen to restart the learning rate at a setting of $[92000, 208000]$. They did not use any efficient optimization strategy or extra datasets.

4.5. H3FC2Z

This team utilizes the RDDM [56] as the deraining backbone and trained it with a patch size of 256×256. Subsequently, the self-ensemble strategy used in the EDSR [51] was improved by replacing the average ensemble with their "Dual Kmeans fusion" in the Fig. 7. Experimental results indicate that employing "Dual Kmeans fusion" [13] increases the score of the RDDM baseline from 31.72 to 32.56.

Dual Kmeans Fusion. Given a clean image $\mathcal{I}_{gt}$, several raindrop and blur degradations are added to $\mathcal{I}_{gt}$, obtaining

https://github.com/XPixelGroup/BasicSR

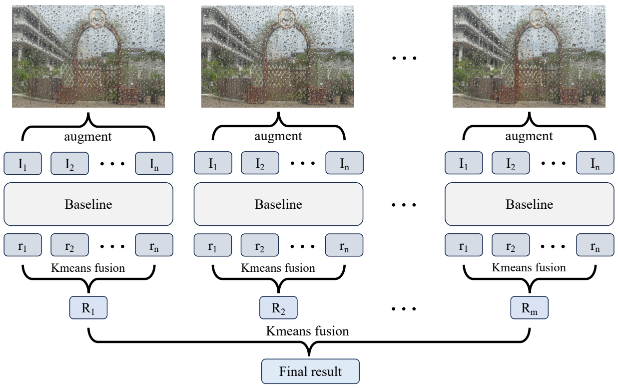

Figure 7. Dual kmeans fusion for RainDrop task proposed by Team H3FC2Z.

m images $(\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_m)$. As shown in Fig. 7, taking $\mathcal{I}_1$ as an example, they serve flipping and rotating as augmentations, generating n images $(I_1, I_2, \cdots I_n)$. **Stage-1**: The baseline model processes these images and obtains preliminary results $r_i = \text{Baseline}(I_i)$, where $i = 1, 2, \cdots, n$. The images r are flipped or rotated to a fixed angle and Kmeans fusion is performed to obtain R_1. **Stage-2**: After performing the above fusion on images $\mathcal{I}$, they get m results $R_1, R_2, \cdots, R_m$. they perform Kmeans fusion on these m results to get the final result image.

Training and Testing Details. The training dataset provided in this challenge is used for model training. To improve generalization, data augmentation techniques, including rotation and flipping are applied. The model is trained for 100,000 iterations using the AdamW optimizer with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.95$, on a single NVIDIA A6000 GPU. Training is conducted with a batch size of 8, a learning rate of 3×10^{-4}, and a patch size of 256×256. During inference, the authors adopt a Dual K-means fusion strategy to further enhance the model's performance. Experimental results indicate that employing "Dual Kmeans fusion" increases the score of RDDM [56] baseline from 31.72 to 32.56. Specifically, Kmeans fusion in stage one increases the score of baseline from 31.72 to 32.26, and Kmeans fusion in stage two increases the score of baseline from 32.26 to 32.56.

4.6. IIC Lab

This team developed an effective frequency-aware and Mamba-based network for image deraining, named FA-Mamba, as shown in Fig. 8. Specifically, the key component of the proposed framework is the Wavelet Domain Restoration Module (WDRM) which contains a Dual-branch Feature Extraction Block (DFEB) that has superior local perception and global modeling capabilities and a Prior-Guided Module (PGM) that provides refined texture detail guidance for feature extraction. It is worth mentioning that the re-

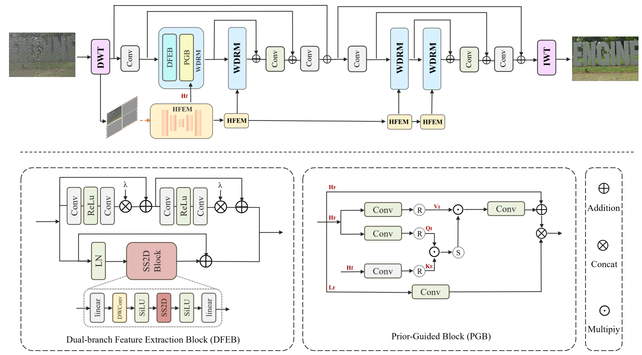

Figure 8. The pipeline of the FA-Mamba proposed by Team IIC Lab.

fined texture details are obtained by enhancing the input high-frequency information through the High-Frequency Enhancement Module (HFEM).

Training Details. During training, they utilized the Adam optimizer with a batch size of 1 and a patch size of 256 for a total of 80 iterations. The initial learning rate is fixed at $1e^{-4}$ for 60 iterations, and then decreased to $5e^{-5}$ for 20 iterations. No data augmentation techniques were applied. The entire framework is performed on PyTorch with an NVIDIA GeForce RTX 3090 GPU, which works in an end-to-end learning fashion without costly large-scale pertaining.

4.7. BUPT CAT

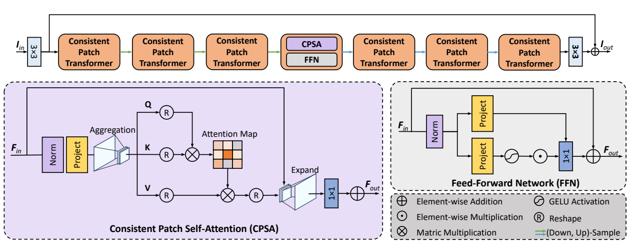

Figure 9. The pipeline of the method proposed by Team BUPT CAT.

As shown in Fig. 9, this team introduces a novel Consistent Patch Transformer (CPT) for dual-focused day and night raindrop removal task, which leverages a UNet-based architecture designed to enhance both spatial consistency and feature representation capability. The framework comprises multiple Consistent Patch Transformer blocks, each consisting of two key components: Consistent Patch Self-Attention (CPSA) and a Feed-Forward Network (FFN). The model utilizes the Test Time Local Converter (TLC) mechanism [16] to effectively revisit global information aggregation and ensure robust and consistent feature learning with different patch sizes at training and testing. The CPSA module is responsible for capturing both long-range dependencies and spatially consistent local details. Instead of using traditional window-based attention mechanisms, the CPSA module integrates a TLC-based feature aggregation and scaling strategy that maintains consistent patch sizes during training and testing, reducing spatial inconsistencies between training and testing.

Training Details. To reduce the training GPU memory, this team augments the input data by randomly cropping the input image into patches of the same size and performing strategies such as random rotation.

Testing Details. During the testing stage, in their self-attention part, they use the TLC strategy to segment and aggregate the full image into a series of patches of the same size as the training patch, and the rest of the model is the full image. This setup can effectively improve the inconsistency of the model's patch size between training and testing, especially in the self-attention part.

Implementation Details. This team utilizes the Pytorch framework with the NVIDIA GeForce RTX 4090. During training, they set the total batch size to 8, the initial learning rate from $5e^{-4}$ to $1e^{-5}$ with a scheduler in 500K iterations, and the patch size is set to 192×192. For the loss function, they use $\mathcal{L}_1$loss and Fourier loss to constrain their model with weights of 1 and 0.1, respectively. They train their framework using the Adam optimizer with $\beta 1 = 0.9, \beta 2 = 0.99$. They set the number of channels to 64 in their network. The total training duration is approximately 80 hours. The training and test sets are official datasets provided by the dual-focused day and night raindrop removal challenge.

4.8. WIRTeam

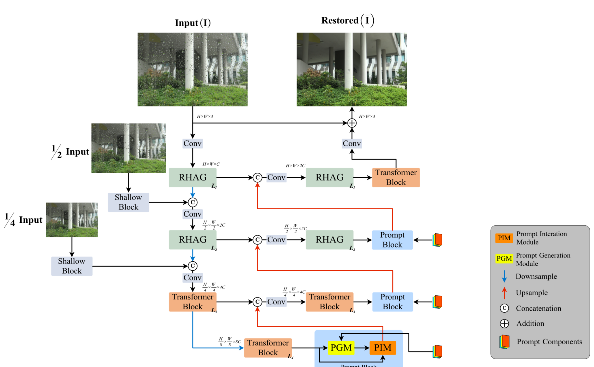

Figure 10. The pipeline of the method proposed by Team WIRteam.

Inspired by recent advancements [10, 37, 63, 89] in image restoration and image deraining, this team propose a novel multi-scale prompt-based image deraining approach (MPID) that incorporates both local and global attention mechanisms, as shown in Fig. 10. Specifically, given a degraded image $I \in R^{H \times W \times 3}$, their model produces a corresponding clear image $\bar{I} \in R^{H \times W \times 3}$ through a 4-level encoder-decoder framework. In the shadow two layers, they employ Residual Hybrid Attention Groups (RHAG) [10] to capture detailed local features. For deeper layers, they integrate Transformer Blocks [89] to facilitate cross-channel global feature extraction. In the encoding phase, acknowledging the advantages of utilizing images at various resolutions for deraining tasks [12], they additionally incorporate multi-scale image information (at 1/2 and 1/4 of the original resolution) to enhance auxiliary information during the encoding process. During decoding, considering the dual focus on raindrop-focus and background-focus images within the Raindrop Clarity dataset [45]—which introduces both raindrop occlusions and background blurring—they introduce specialized prompt mechanism [37]. These components are designed to address and decouple different types of degradation factors. Notably, Prompt Block, which integrates Prompt Generation Module (PGM) and Prompt Interaction Module (PIM), utilizes image-specific cues to effectively guide the image reconstruction process, thereby improving the clarity and quality of the restored images.

Training Details. They employ the end-to-end training methodology, training their model for 400 epochs. The training procedure is based on the AdamW optimizer with the decay parameters $\beta_1 = 0.9$ and $\beta_2 = 0.99$. The initial learning rate is 2e-4 and gradually reduces to 1e-6 with the cosine annealing strategy. Horizontal and vertical flips are adopted for data augmentation. Furthermore, their training is merely based on the Raindrop Clarity dataset provided by the competition organizers, without the use of any extra datasets or pretrained models.

Testing Details. During the testing phase, they preprocess the input images by padding them to the multiple of 32. This ensures that the dimensions of the input images are compatible with the architecture of their method. They don't resort to any other means of Test-Time Augmentation during the testing phase.

Acknowledgments

This work was partially supported by NSFC under Grant 623B2098 and the China Postdoctoral Science Foundation-Anhui Joint Support Program under Grant Number 2024T017AH. We thank the challenge sponsor: Eastern Institute for Advanced Study, Ningbo. This work was partially supported by the Humboldt Foundation. We thank the NTIRE 2025 sponsors: ByteDance, Meituan, Kuaishou, and University of Wurzburg (Computer Vision Lab).

References

[1] Kelvin CK Chan, Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Basicvsr: The search for essential components in video super-resolution and beyond. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4947–4956, 2021. 6

[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*, pages 15–22, 2024. 5

[3] Hongming Chen, Xiang Chen, Jiyang Lu, and Yufeng Li. Rethinking multi-scale representations in deep deraining transformer. In *AAAI*, pages 1046–1053, 2024. 5

[4] Hongming Chen, Xiang Chen, Jiyang Lu, and Yufeng Li. Rethinking multi-scale representations in deep deraining transformer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1046–1053, 2024. 2

[5] Sixiang Chen, Tian Ye, Jinbin Bai, Erkang Chen, Jun Shi, and Lei Zhu. 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*, pages 13106–13117, 2023. 3

[6] Sixiang Chen, Tian Ye, Kai Zhang, Zhaohu Xing, Yunlong Lin, and Lei Zhu. Teaching tailored to talent: Adverse weather restoration via prompt pool and depth-anything constraint. In *European Conference on Computer Vision*, pages 95–115. Springer, 2024. 3

[7] Sixiang Chen, Tian Ye, Kai Zhang, Zhaohu Xing, Yunlong Lin, and Lei Zhu. Teaching tailored to talent: Adverse weather restoration via prompt pool and depth-anything constraint. In *European Conference on Computer Vision*, pages 95–115. Springer, 2024. 2

[8] Tingting Chen, Beibei Lin, Yeying Jin, Wending Yan, Wei Ye, Yuan Yuan, and Robby T Tan. Dual-rain: Video rain removal using assertive and gentle teachers. In *European Conference on Computer Vision*, pages 127–143. Springer, 2024. 3

[9] 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. 5

[10] Xiangyu Chen, Xintao Wang, Jiantao Zhou, Yu Qiao, and Chao Dong. Activating more pixels in image super-resolution transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 22367–22377, 2023. 8

[11] Xiangyu Chen, Zheyuan Li, Yuandong Pu, Yihao Liu, Jiantao Zhou, Yu Qiao, and Chao Dong. A comparative study of image restoration networks for general backbone network design. In *European Conference on Computer Vision*, pages 74–91. Springer, 2024. 6

[12] Xiang Chen, Jinshan Pan, and Jiangxin Dong. Bidirectional multi-scale implicit neural representations for image deraining. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25627–25636, 2024. 5, 8

[13] Zheng Chen, Zongwei Wu, Eduard Zamfir, Kai Zhang, Yulun Zhang, Radu Timofte, Xiaokang Yang, Hongyuan Yu, Cheng Wan, Yuxin Hong, et al. Ntire 2024 challenge on image super-resolution (x4): Methods and results. In *CVPR*, pages 6108–6132, 2024. 6

[14] Zheng Chen, Kai Liu, Jue Gong, Jingkai Wang, Lei Sun, Zongwei Wu, Radu Timofte, Yulun Zhang, et al. NTIRE 2025 challenge on image super-resolution (×4): Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[15] Zheng Chen, Jingkai Wang, Kai Liu, Jue Gong, Lei Sun, Zongwei Wu, Radu Timofte, Yulun Zhang, et al. NTIRE 2025 challenge on real-world face restoration: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[16] Xiaojie Chu, Liangyu Chen, Chengpeng Chen, and Xin Lu. Improving image restoration by revisiting global information aggregation. In *European Conference on Computer Vision*, pages 53–71. Springer, 2022. 7

[17] Marcos Conde, Radu Timofte, et al. NTIRE 2025 challenge on raw image restoration and super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[18] Marcos Conde, Radu Timofte, et al. Raw image reconstruction from RGB on smartphones. NTIRE 2025 challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[19] Egor Ershov, Sergey Korchagin, Alexei Khalin, Artyom Panshin, Arseniy Terekhin, Ekaterina Zaychenkova, Georgiy Lobarev, Vsevolod Plokhotnyuk, Denis Abramov, Elisey Zhdanov, Sofia Dorogova, Yasin Mamedov, Nikola Banic, Georgii Perevozchikov, Radu Timofte, et al. NTIRE 2025 challenge on night photography rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[20] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3855–3863, 2017. 2

[21] Yuqian Fu, Xingyu Qiu, Bin Ren Yanwei Fu, Radu Timofte, Nicu Sebe, Ming-Hsuan Yang, Luc Van Gool, et al. NTIRE 2025 challenge on cross-domain few-shot object detection: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[22] Shuhao Han, Haotian Fan, Fangyuan Kong, Wenjie Liao, Chunle Guo, Chongyi Li, Radu Timofte, et al. NTIRE 2025 challenge on text to image generation model quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[23] Zhixiang Hao, Shaodi You, Yu Li, Kunming Li, and Feng Lu. Learning from synthetic photorealistic raindrop for single image raindrop removal. In *ICCV Workshops*, pages 0–0, 2019. 2

[24] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 8022–8031, 2019. 2

[25] Wenfeng Huang, Guoan Xu, Wenjing Jia, Stuart Perry, and Guangwei Gao. Revivediff: A universal diffusion model for restoring images in adverse weather conditions. *arXiv preprint arXiv:2409.18932*, 2024. 2

[26] Varun Jain, Zongwei Wu, Quan Zou, Louis Florentin, Henrik Turbell, Sandeep Siddhartha, Radu Timofte, et al. NTIRE 2025 challenge on video quality enhancement for video conferencing: Datasets, methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[27] 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 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8346–8355, 2020. 2

[28] Yeying Jin, Aashish Sharma, and Robby T Tan. 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*, pages 5027–5036, 2021. 3

[29] Yeying Jin, Wending Yan, Wenhan Yang, and Robby T Tan. Structure representation network and uncertainty feedback learning for dense non-uniform fog removal. In *Proceedings of the Asian Conference on Computer Vision*, pages 2041–2058, 2022. 3

[30] Yeying Jin, Wenhan Yang, and Robby T Tan. Unsupervised night image enhancement: When layer decomposition meets light-effects suppression. In *European Conference on Computer Vision*, pages 404–421. Springer, 2022. 2

[31] Yeying Jin, Ruoteng Li, Wenhan Yang, and Robby T Tan. Estimating reflectance layer from a single image: Integrating reflectance guidance and shadow/specular aware learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1069–1077, 2023. 3

[32] Yeying Jin, Beibei Lin, Wending Yan, Yuan Yuan, Wei Ye, and Robby T Tan. Enhancing visibility in nighttime haze images using guided apsf and gradient adaptive convolution. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 2446–2457, 2023. 2

[33] Yeying Jin, Xin Li, Jiadong Wang, Yan Zhang, and Malu Zhang. Raindrop clarity: A dual-focused dataset for day and night raindrop removal. In *European Conference on Computer Vision*, pages 1–17. Springer, 2024. 1, 2, 3, 5

[34] Yeying Jin, Wei Ye, Wenhan Yang, Yuan Yuan, and Robby T Tan. Des3: Adaptive attention-driven self and soft shadow removal using vit similarity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2634–2642, 2024. 3

[35] Sangmin Lee, Eunpil Park, Angel Canelo, Hyunhee Park, Youngjo Kim, Hyungju Chun, Xin Jin, Chongyi Li, Chun-Le Guo, Radu Timofte, et al. NTIRE 2025 challenge on efficient burst hdr and restoration: Datasets, methods, and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[36] Bingchen Li, Xin Li, Yiting Lu, Sen Liu, Ruoyu Feng, and Zhibo Chen. Hst: Hierarchical swin transformer for compressed image super-resolution. In *European conference on computer vision*, pages 651–668. Springer, 2022. 2

[37] Bingchen Li, Xin Li, Yiting Lu, Ruoyu Feng, Mengxi Guo, Shijie Zhao, Li Zhang, and Zhibo Chen. Promptcir: blind compressed image restoration with prompt learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6442–6452, 2024. 8

[38] Bingchen Li, Xin Li, Hanxin Zhu, Yeying Jin, Ruoyu Feng, Zhizheng Zhang, and Zhibo Chen. Sed: Semantic-aware discriminator for image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 25784–25795, 2024. 2

[39] Bingchen Li, Xin Li, Yiting Lu, and Zhibo Chen. Hybrid agents for image restoration. *arXiv preprint arXiv:2503.10120*, 2025. 2

[40] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1633–1642, 2019. 2

[41] 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 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3838–3847, 2019. 2

[42] Xin Li, Xin Jin, Jianxin Lin, Sen Liu, Yaojun Wu, Tao Yu, Wei Zhou, and Zhibo Chen. Learning disentangled feature representation for hybrid-distorted image restoration. In *ECCV*, pages 313–329. Springer, 2020. 2

[43] Xin Li, Bingchen Li, Xin Jin, Cuiling Lan, and Zhibo Chen. Learning distortion invariant representation for image restoration from a causality perspective. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1714–1724, 2023. 2

[44] Xin Li, Yulin Ren, Xin Jin, Cuiling Lan, Xingrui Wang, Wenjun Zeng, Xinchao Wang, and Zhibo Chen. Diffusion models for image restoration and enhancement–a comprehensive survey. *arXiv preprint arXiv:2308.09388*, 2023. 2

[45] Xin Li, Yeying Jin, Xin Jin, Zongwei Wu, Bingchen Li, Yufei Wang, Wenhan Yang, Yu Li, Zhibo Chen, Bihan Wen, Robby Tan, Radu Timofte, et al. NTIRE 2025 challenge on day and night raindrop removal for dual-focused images: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2, 8

[46] Xin Li, Xijun Wang, Bingchen Li, Kun Yuan, Yizhen Shao, Suhang Yao, Ming Sun, Chao Zhou, Radu Timofte, and Zhibo Chen. NTIRE 2025 challenge on short-form ugc video quality assessment and enhancement: Kwaisr dataset and study. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[47] Xin Li, Kun Yuan, Bingchen Li, Fengbin Guan, Yizhen Shao, Zihao Yu, Xijun Wang, Yiting Lu, Wei Luo, Suhang

Yao, Ming Sun, Chao Zhou, Zhibo Chen, Radu Timofte, et al. NTIRE 2025 challenge on short-form ugc video quality assessment and enhancement: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[48] Yu Li, Robby T Tan, Xiaojie Guo, Jiangbo Lu, and Michael S Brown. Rain streak removal using layer priors. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2736–2744, 2016. 2

[49] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1833–1844, 2021. 2

[50] Jie Liang, Radu Timofte, Qiaosi Yi, Zhengqiang Zhang, Shuaizheng Liu, Lingchen Sun, Rongyuan Wu, Xindong Zhang, Hui Zeng, Lei Zhang, et al. NTIRE 2025 the 2nd restore any image model (RAIM) in the wild challenge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[51] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144, 2017. 6

[52] Beibei Lin, Yeying Jin, Wending Yan, Wei Ye, Yuan Yuan, and Robby T Tan. Nighthaze: Nighttime image dehazing via self-prior learning. *arXiv preprint arXiv:2403.07408*, 2024. 2

[53] Beibei Lin, Yeying Jin, Wending Yan, Wei Ye, Yuan Yuan, Shunli Zhang, and Robby T Tan. Nightrain: Nighttime video deraining via adaptive-rain-removal and adaptive-correction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3378–3385, 2024. 2

[54] Yunlong Lin, Zixu Lin, Haoyu Chen, Panwang Pan, Chenxin Li, Sixiang Chen, Yeying Jin, Wenbo Li, and Xinghao Ding. Jarvisir: Elevating autonomous driving perception with intelligent image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. 3

[55] Jianzhao Liu, Jianxin Lin, Xin Li, Wei Zhou, Sen Liu, and Zhibo Chen. Lira: Lifelong image restoration from unknown blended distortions. In *European Conference on Computer Vision*, pages 616–632. Springer, 2020. 2

[56] Jiawei Liu, Qiang Wang, Huijie Fan, Yinong Wang, Yandong Tang, and Liangqiong Qu. Residual denoising diffusion models. In *CVPR*, pages 2773–2783, 2024. 6, 7

[57] Xiaohong Liu, Xiongkuo Min, Qiang Hu, Xiaoyun Zhang, Jie Guo, et al. NTIRE 2025 XGC quality assessment challenge: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[58] Xiaoning Liu, Zongwei Wu, Florin-Alexandru Vasluianu, Hailong Yan, Bin Ren, Yulun Zhang, Shuhang Gu, Le Zhang, Ce Zhu, Radu Timofte, et al. NTIRE 2025 challenge on low light image enhancement: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[59] Yiting Lu, Xin Li, Bingchen Li, Zihao Yu, Fengbin Guan, Xinrui Wang, Ruling Liao, Yan Ye, and Zhibo Chen. Aigcvqa: A holistic perception metric for aigc video quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6384–6394, 2024. 2

[60] Ozan Özdenizci and Robert Legenstein. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8):10346–10357, 2023. 2

[61] Yingxue Pang, Xin Li, Xin Jin, Yaojun Wu, Jianzhao Liu, Sen Liu, and Zhibo Chen. Fan: Frequency aggregation network for real image super-resolution. In *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, pages 468–483. Springer, 2020. 2

[62] Horia Porav, Tom Bruls, and Paul Newman. I can see clearly now: Image restoration via de-raining. In *ICRA*, pages 7087–7093. IEEE, 2019. 2

[63] Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36:71275–71293, 2023. 8

[64] Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2482–2491, 2018. 2

[65] 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. 2

[66] Ruijie Quan, Xin Yu, Yuanzhi Liang, and Yi Yang. Removing raindrops and rain streaks in one go. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9147–9156, 2021. 2

[67] Bin Ren, Hang Guo, Lei Sun, Zongwei Wu, Radu Timofte, Yawei Li, et al. The tenth NTIRE 2025 efficient super-resolution challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[68] Qiyu Rong, Hongyuan Jing, Mengmeng Zhang, Jinlong Li, and Mengfei Han. STRRNet: Semantics-guided two-stage raindrop removal network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 3

[69] Nickolay Safonov, Alexey Bryntsev, Andrey Moskalenko, Dmitry Kulikov, Dmitriy Vatolin, Radu Timofte, et al. NTIRE 2025 challenge on UGC video enhancement: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[70] Yiyang Shen, Mingqiang Wei, Yongzhen Wang, Xueyang Fu, and Jing Qin. Rethinking real-world image deraining via an unpaired degradation-conditioned diffusion model. *arXiv preprint arXiv:2301.09430*, 2023. 2

[71] Vera Soboleva and Oleg Shipitko. Raindrops on windshield: Dataset and lightweight gradient-based detection algorithm. In *IEEE SSCI*, pages 1–7. IEEE, 2021. 2

[72] Lei Sun, Andrea Alfarano, Peiqi Duan, Shaolin Su, Kaiwei Wang, Boxin Shi, Radu Timofte, Danda Pani Paudel, Luc Van Gool, et al. NTIRE 2025 challenge on event-based image deblurring: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[73] Lei Sun, Hang Guo, Bin Ren, Luc Van Gool, Radu Timofte, Yawei Li, et al. The tenth ntire 2025 image denoising challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[74] Shangquan Sun, Wenqi Ren, Xinwei Gao, Rui Wang, and Xiaochun Cao. Restoring images in adverse weather conditions via histogram transformer. In *ECCV*, pages 111–129. Springer, 2024. 6

[75] Zhengzhong Tu, Hossein Talebi, Han Zhang, Feng Yang, Peyman Milanfar, Alan Bovik, and Yinxiao Li. Maxim: Multi-axis mlp for image processing. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5769–5780, 2022. 2

[76] Jeya Maria Jose Valanarasu, Rajeev Yasarla, and Vishal M Patel. Transweather: Transformer-based restoration of images degraded by adverse weather conditions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2353–2363, 2022. 2

[77] Florin-Alexandru Vasluianu, Tim Seizinger, Zhuyun Zhou, Cailian Chen, Zongwei Wu, Radu Timofte, et al. NTIRE 2025 image shadow removal challenge report. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[78] Florin-Alexandru Vasluianu, Tim Seizinger, Zhuyun Zhou, Zongwei Wu, Radu Timofte, et al. NTIRE 2025 ambient lighting normalization challenge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[79] 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 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12270–12279, 2019. 2

[80] Yingqian Wang, Zhengyu Liang, Fengyuan Zhang, Lvli Tian, Longguang Wang, Juncheng Li, Jungang Yang, Radu Timofte, Yulan Guo, et al. NTIRE 2025 challenge on light field image super-resolution: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[81] Hongtao Wu, Yijun Yang, Huihui Xu, Weiming Wang, Jinni Zhou, and Lei Zhu. Rainmamba: Enhanced locality learning with state space models for video deraining. In *ACMMM*, pages 7881–7890, 2024. 2

[82] Jie Xiao, Xueyang Fu, Aiping Liu, Feng Wu, and Zheng-Jun Zha. Image de-raining transformer. *IEEE transactions on pattern analysis and machine intelligence*, 45(11):12978–12995, 2022. 2

[83] Kangning Yang, Jie Cai, Ling Ouyang, Florin-Alexandru Vasluianu, Radu Timofte, Jiaming Ding, Huiming Sun, Lan Fu, Jinlong Li, Chiu Man Ho, Zibo Meng, et al. NTIRE 2025 challenge on single image reflection removal in the wild: Datasets, methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[84] 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 *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. workshops*, pages 1357–1366, 2017. 2

[85] 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 *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1357–1366, 2017. 2

[86] Tian Ye, Sixiang Chen, Jinbin Bai, Jun Shi, Chenghao Xue, Jingxia Jiang, Junjie Yin, Erkang Chen, and Yun Liu. Adverse weather removal with codebook priors. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 12653–12664, 2023. 3

[87] Zihao Yu, Fengbin Guan, Yiting Lu, Xin Li, and Zhibo Chen. Sf-iqa: Quality and similarity integration for ai generated image quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6692–6701, 2024. 2

[88] Pierluigi Zama Ramirez, Fabio Tosi, Luigi Di Stefano, Radu Timofte, Alex Costanzino, Matteo Poggi, Samuele Salti, Stefano Mattoccia, et al. NTIRE 2025 challenge on hr depth from images of specular and transparent surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025. 2

[89] 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 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5728–5739, 2022. 2, 3, 6, 8

[90] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 695–704, 2018. 2

[91] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, pages 586–595, 2018. 2

[92] Zhen Zou, Hu Yu, Jie Huang, and Feng Zhao. Freqmamba: Viewing mamba from a frequency perspective for image deraining. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 1905–1914, 2024. 2