From Sky to the Ground: A Large-scale Benchmark and Simple Baseline Towards Real Rain Removal

Yun Guo[1,†], Xueyao Xiao[1,†], Yi Chang[1,*], Shumin Deng[1], Luxin Yan[1]
[1]National Key Laboratory of Science and Technology on Multispectral Information Processing,
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, China
{guoyun, xiaoxueyao, yichang, shumindeng, yanluxin}@hust.edu.cn

ICCV 2023

WeatherStream: Light Transport Automation of Single Image Deweathering

Howard Zhang[1*] Yunhao Ba[1*] Ethan Yang[1] Varan Mehra[1] Blake Gella[1]
Akira Suzuki[1] Arnold Pfahnl[1] Chethan Chinder Chandrappa[1]
Alex Wong[2] Achuta Kadambi[1]
[1]University of California, Los Angeles [2]Yale University

CVPR 2023

Presenter: Hao Wang

Advisor: Prof. Chia-Wen Lin

Outline

- Introduction
- LHP-Rain
  - Method
  - Experiment
- WeatherStream
  - Method
  - Experiment
- Conclusion

Outline

- Introduction
- LHP-Rain
  - Method
  - Experiment
- WeatherStream
  - Method
  - Experiment
- Conclusion

Introduction

- Challenge
  - Previous state-of-the-art methods that have attempted the all-weather removal task train on synthetic pairs, and are thus limited by the Sim2Real domain gap.

- Method
  - construct a large-scale high-quality paired real benchmark for real single image

Introduction

- Ways to construct the pairs
  - video-based generation (LHP-Rain)
  - time-interval acquisition (WeatherStream)
- Source
  - Smartphone (LHP-Rain)
  - Youtube stream (WeatherStream)
- Categories
  - Diverse Rain, including streak, veiling, occlusion, splashing (LHP-Rain)
  - Rain and Snow (WeatherStream)
- Amount
  - 1.0M pairs (LHP-Rain)
  - 202K pairs (WeatherStream)

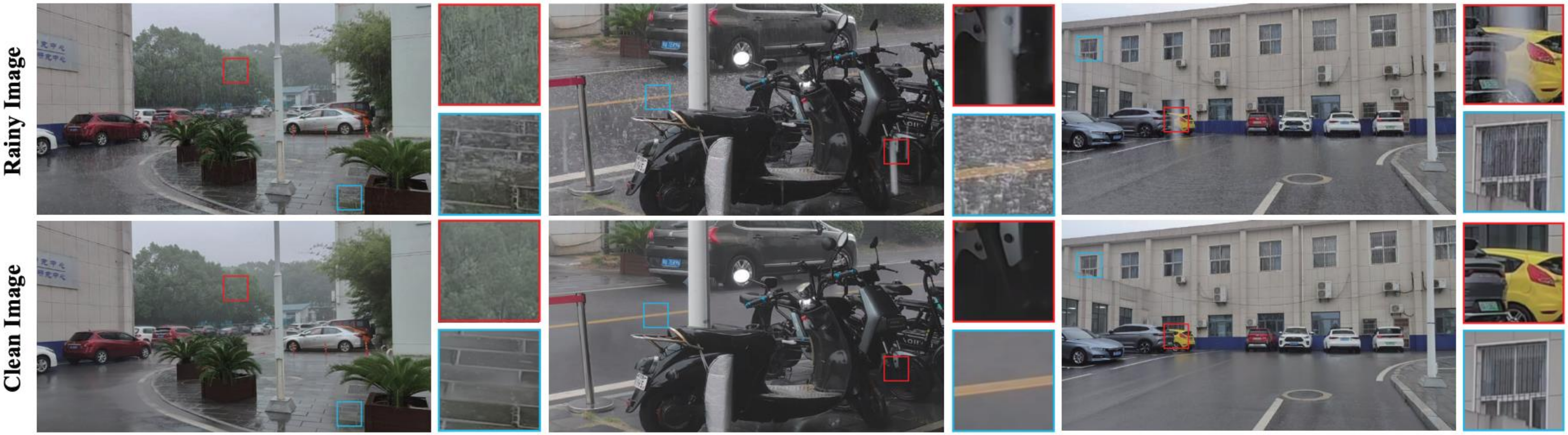
Outline

- Introduction
- LHP-Rain
  - Method
  - Experiment
- WeatherStream
  - Method
  - Experiment
- Conclusion

From Sky to the Ground: A Large-scale Benchmark and Simple Baseline Towards Real Rain Removal

Yun Guo[1,†], Xueyao Xiao[1,†], Yi Chang[1,*], Shumin Deng[1], Luxin Yan[1]
[1]National Key Laboratory of Science and Technology on Multispectral Information Processing,
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, China
{guoyun, xiaoxueyao, yichang, shumindeng, yanluxin}@hust.edu.cn

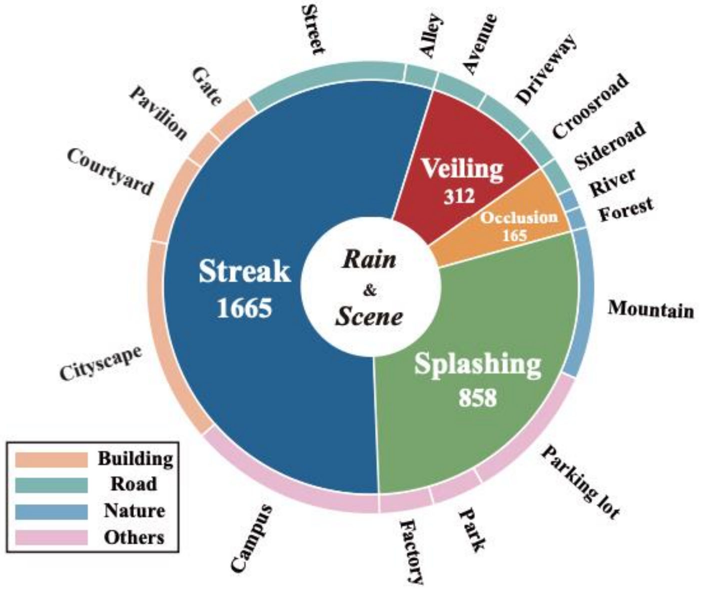
Example



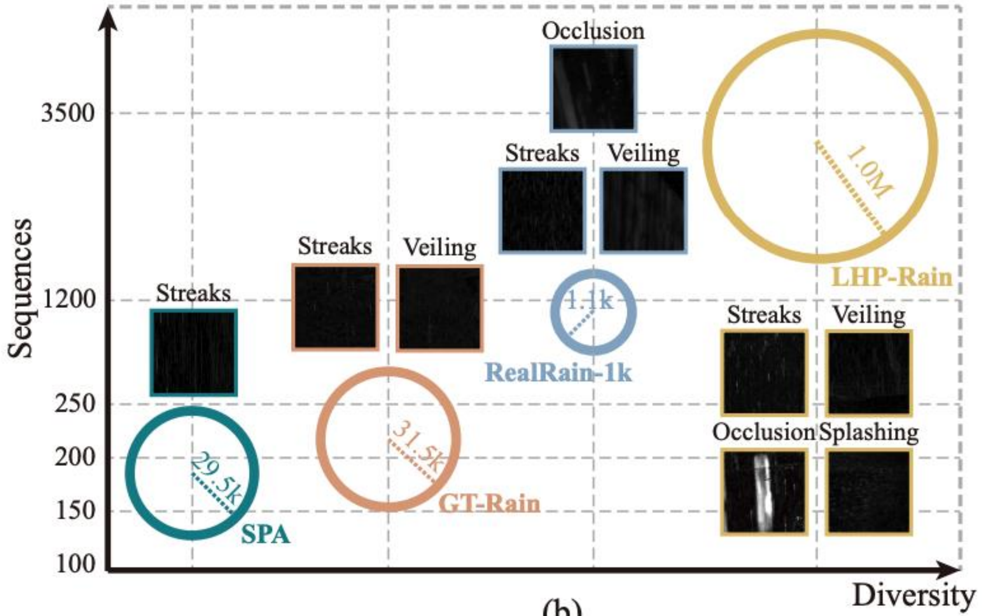
Existing dataset

Table 1: Summary of existing real rain datasets.

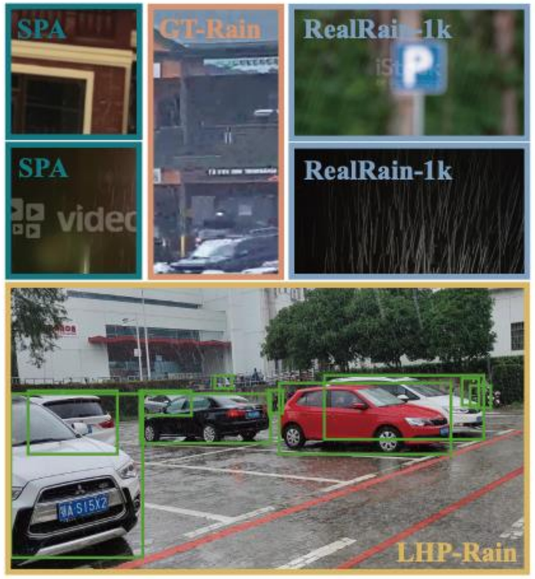
Datasets	Year	Source	Sequence	Frame	Resolution	Rain Categories	Annotation	Paired
RID/RIS[19]	2019	Cam/Internet	None	4.5K	640*368	streak, raindrop	Object detection	-
NR-IQA[43]	2020	Internet	None	0.2K	1000*680	streak, veiling	None	-
Real3000[25]	2021	Internet	None	3.0K	942*654	streak, veiling	None	-
FCRealRain[51]	2022	Camera	None	4.0K	4240*2400	streak, veiling	Object detection	-
SPA-Data[37]	2019	Cam/Internet	170	29.5K	256*256	streak	None	✓
RainDS[32]	2021	Cam	None	1.0K	1296*728	streak, raindrop	None	✓
GT-Rain[1]	2022	Internet	202	31.5K	666*339	streak, veiling	None	✓
RealRain-1K[20]	2022	Cam/Internet	1120	1.1K	1512*973	streak, veiling, occlusion	None	✓
LHP-Rain	2023	Camera	**3000**	**1.0M**	1920*1080	streak, veiling, occlusion, **splashing**	Object detection/**Lane**	✓



(a)

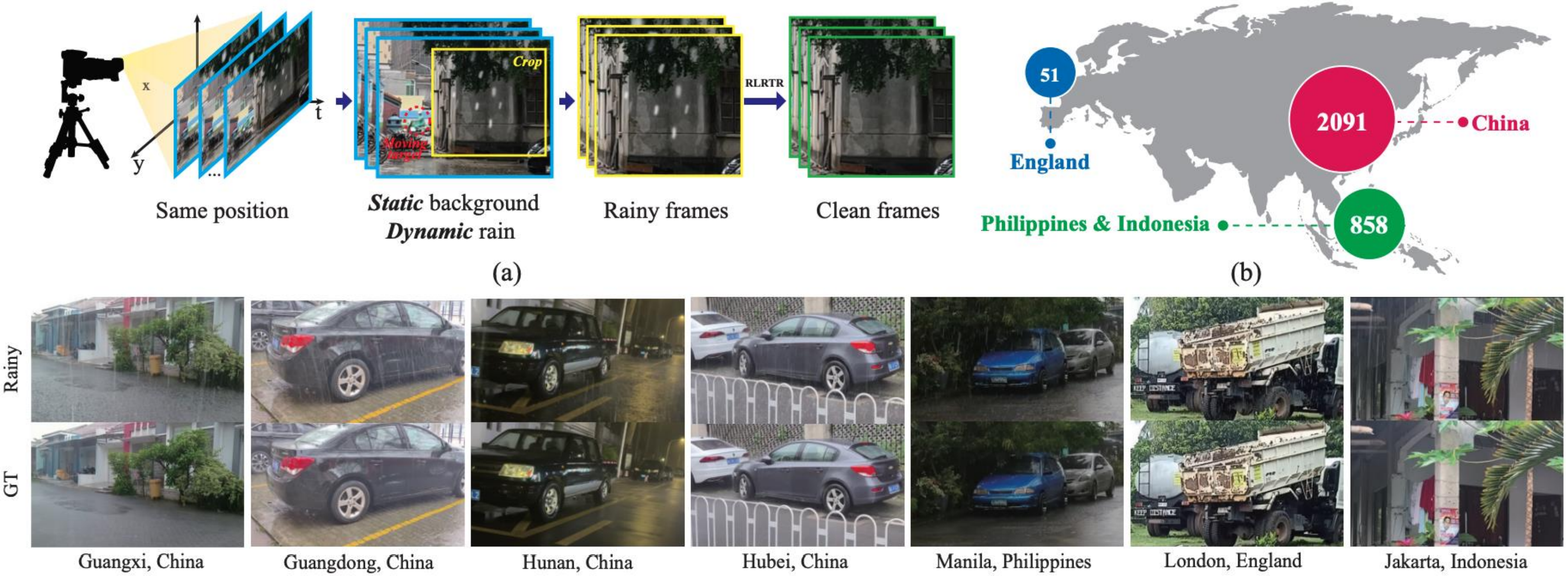


(b)



(c)

Framework


x
y
t
...
Same position
Crop
Moving target
Static background
Dynamic rain
Rainy frames
RLRTR
Clean frames
(a)
51
England
2091
China
Philippines & Indonesia
858
(b)
Rainy
GT
Guangxi, China
Guangdong, China
Hunan, China
Hubei, China
Manila, Philippines
London, England
Jakarta, Indonesia

Robust Low-rank Tensor Recovery Model

RLRTR

$$\mathcal{O} \circ \tau = \mathcal{B} + \mathcal{R} + \mathcal{N}, \quad (1)$$

$$\min_{\mathcal{B},\mathcal{R},\tau} \frac{1}{2}||\mathcal{B} + \mathcal{R} - \mathcal{O} \circ \tau||_F^2 + \boxed{\omega P_b(\mathcal{B})} + \boxed{\mu P_r(\mathcal{R})}, \quad (2)$$

- P_b and P_r are the prior knowledge for the image and rain

Method

$$\min_{\mathcal{B},\mathcal{R},\tau} \frac{1}{2}||\mathcal{B}+\mathcal{R}-\mathcal{O}\circ\tau||_F^2 + \boxed{\omega P_b(\mathcal{B})} + \mu P_r(\mathcal{R}), \quad (2)$$

$$\boxed{P_b(\mathcal{B})} = \omega \sum_i \left(\frac{1}{\lambda_i^2}||\underline{\mathcal{S}_i\mathcal{B}}\times_3 \underline{Q_i} - \mathcal{J}_i||_F^2 + \underline{||\mathcal{J}_i||_{tnn}} \right) + \underline{\gamma||\nabla_t\mathcal{B}||_1}, \quad (3)$$

- tensor nuclear norm
- Total variation

- **non-local low-rank property**
  - along spatial dimension via the non-local clustering of a sub-cubic

$\mathcal{S}_i\mathcal{B} \in \mathbb{R}^{p^2\times k\times t}$ $\quad$ $u_i \in \mathbb{R}^{p\times p\times t}$

- **global low-rank property**
  - along the temporal dimension enforce the orthogonal constraint on

$Q_i \in \mathbb{R}^{d\times t}(d \ll t)$ $\quad$ $Q_i^T Q_i = I$

Method

$$\min_{\mathcal{B},\mathcal{R},\tau} \frac{1}{2}||\mathcal{B} + \mathcal{R} - \mathcal{O} \circ \tau||_F^2 + \omega P_b(\mathcal{B}) + \boxed{\mu P_r(\mathcal{R})}, \quad (2)$$

- As for the rain R, we formulate it as the sparse error via the **L1 sparsity**.

Method

$$\min_{\mathcal{B},\mathcal{R},\tau} \underline{\frac{1}{2}||\mathcal{B}+\mathcal{R}-\mathcal{O}\circ\tau||_F^2} + \boxed{\omega P_b(\mathcal{B})} + \boxed{\mu P_r(\mathcal{R})}, \quad (2)$$

$$\Rightarrow \quad \left\{\hat{\mathcal{B}}, \hat{\mathcal{R}}, \hat{\mathcal{J}}_i, \hat{\tau}, \hat{Q}_i\right\} = \arg \min_{\mathcal{B},\mathcal{R},\mathcal{J}_i,\tau,Q_i} \underline{\tfrac{1}{2}||\mathcal{B}+\mathcal{R}-\mathcal{O}\circ\tau||_F^2} + \boxed{\mu||\mathcal{R}||_1} + \boxed{\omega \sum_i \left(\tfrac{1}{\lambda_i^2}||\mathcal{S}_i\mathcal{B}\times_3 Q_i - \mathcal{J}_i||_F^2 + ||\mathcal{J}_i||_{tnn}\right) + \gamma||\nabla_t\mathcal{B}||_1}. \quad (4)$$

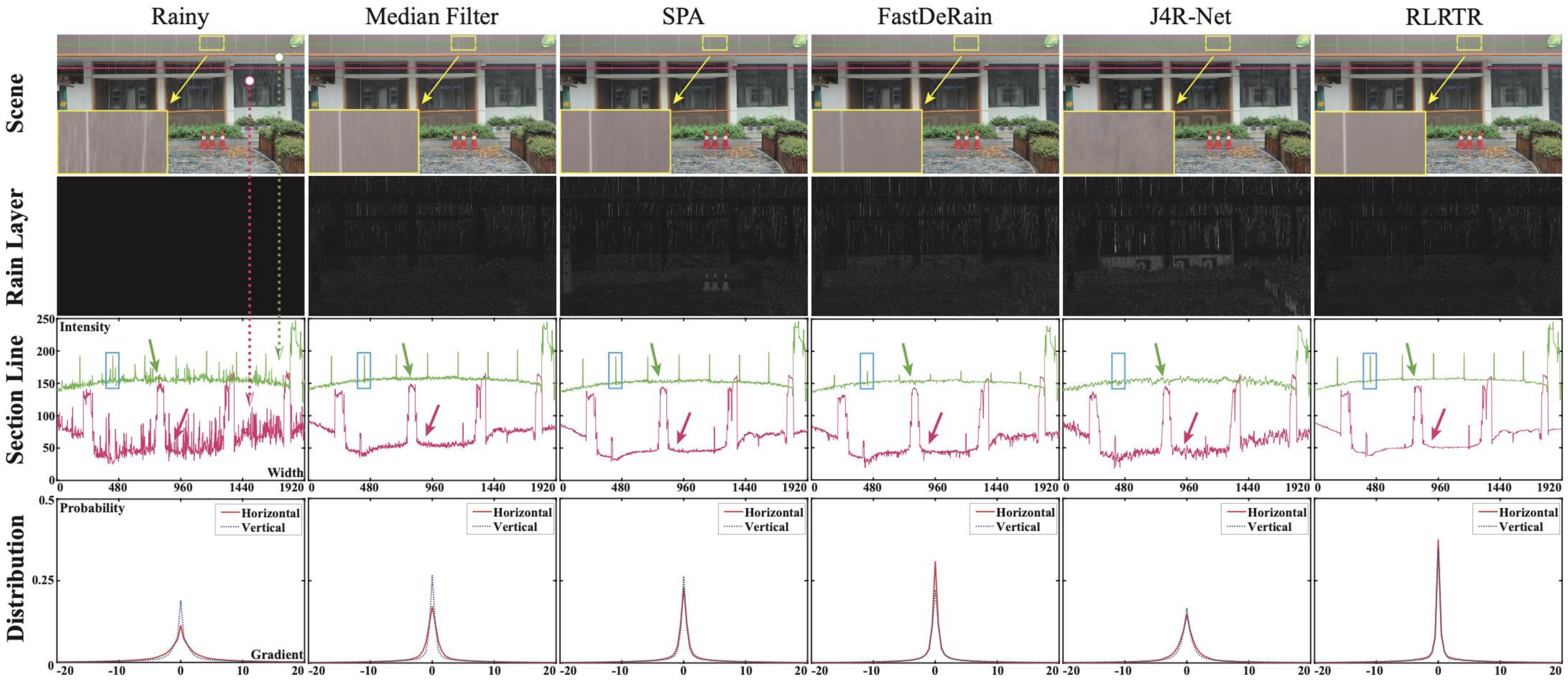
- Optimization.

$\Rightarrow$

  - Due to the difficulty of estimating multiple variables directly, we adopt the **alternating minimization scheme** to solve the Eq. (4) with respect to each variable.

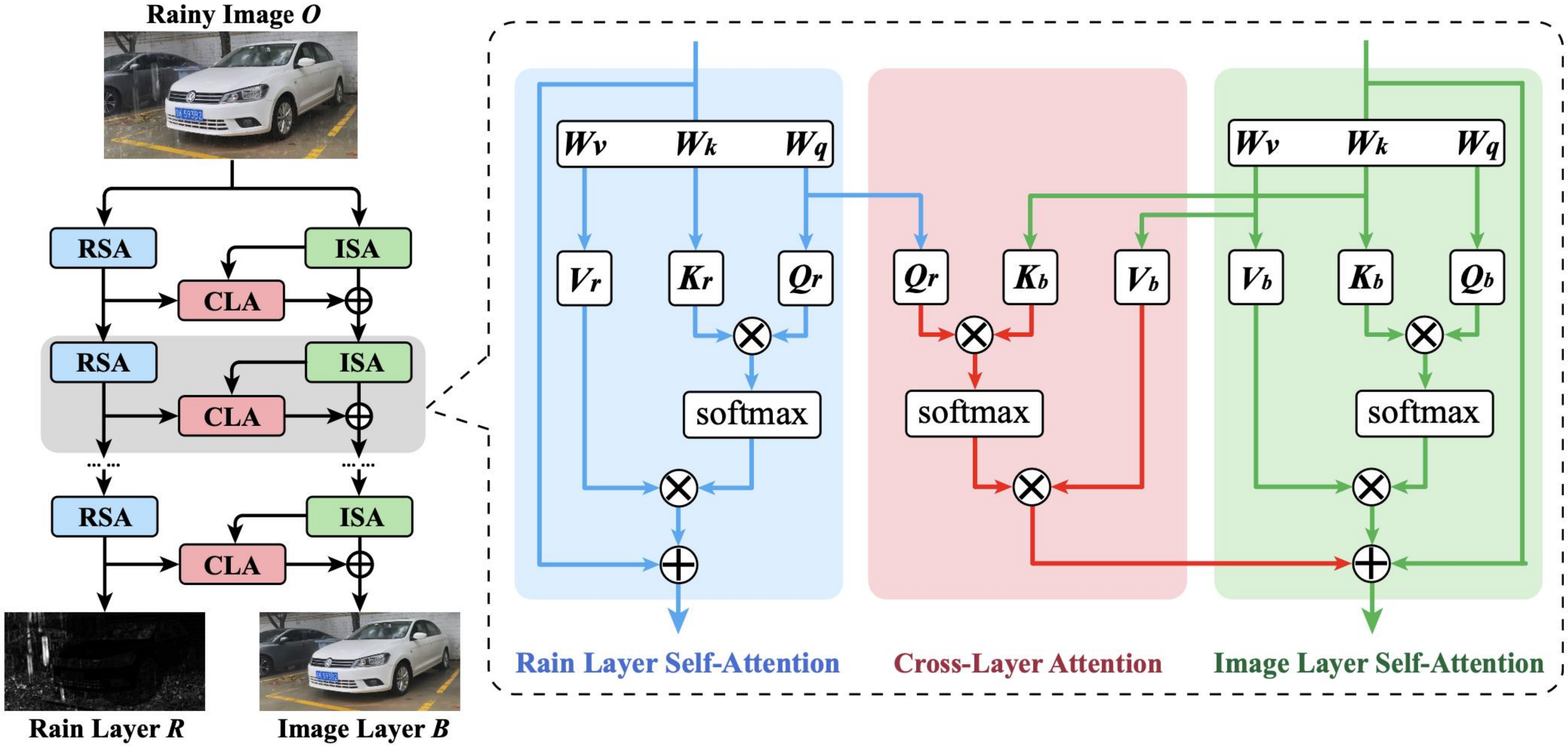
Experiment

different video deraining results



Experiment

framework of the SCD-Former



$$\mathcal{L} = ||O - B - R||_F^2 + \lambda_r ||R - \hat{R}||_1 + \lambda_b ||B - \hat{B}||_1. \quad (12)$$

Experiment

Method	A→A		B→A		C→A		A→B		B→B		C→B		A→C		B→C		C→C	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Rainy Image			32.60 / 0.9173						19.48 / 0.5849						29.97 / 0.8497			
SPANet	38.53	0.9875	22.93	0.8207	31.46	0.9612	20.01	0.6148	21.51	0.7145	19.20	0.5706	28.00	0.8905	20.10	0.8061	31.19	**0.9346**
PReNet	37.05	0.9696	22.44	0.7713	**32.46**	0.9387	20.29	0.5860	20.65	0.6005	19.34	0.5530	27.57	0.8595	20.91	0.7222	32.13	0.9177
RCDNet	39.74	0.9661	22.51	0.8392	32.30	0.9378	20.09	0.5785	21.04	0.6106	19.09	0.5264	25.46	0.7959	21.38	0.8047	32.34	0.9152
JORDER-E	40.63	0.9794	23.47	0.7426	31.23	0.9234	19.98	0.5799	21.24	0.6854	18.76	0.4861	27.13	0.8531	22.14	0.8433	31.24	0.8847
MPRNet	46.06	0.9894	24.27	0.8428	32.37	0.9379	19.87	0.6286	22.00	0.6515	19.47	0.5889	28.41	0.8807	**23.82**	0.8052	33.34	0.9309
GT-Rain	37.21	0.9827	**25.30**	**0.9243**	26.46	0.9145	20.07	**0.6941**	**22.51**	**0.7300**	**21.14**	0.5698	28.62	0.8675	23.19	0.8098	32.18	0.9132
Uformer-B	**46.42**	**0.9917**	24.08	0.8979	32.21	**0.9667**	19.70	0.6875	21.60	0.7124	19.10	**0.6622**	**28.74**	**0.9262**	22.91	**0.8734**	**33.56**	0.9317
IDT	45.74	0.9889	23.80	0.8334	32.38	0.9422	**20.34**	0.6306	21.98	0.6536	19.44	0.5977	26.90	0.8742	23.34	0.7897	33.02	0.9310
SCD-Former	**46.89**	**0.9941**	**26.13**	**0.9122**	**34.38**	**0.9798**	**20.98**	**0.6985**	**22.79**	**0.7684**	**21.71**	**0.6893**	**29.41**	**0.9127**	**23.56**	**0.8626**	**34.33**	**0.9468**

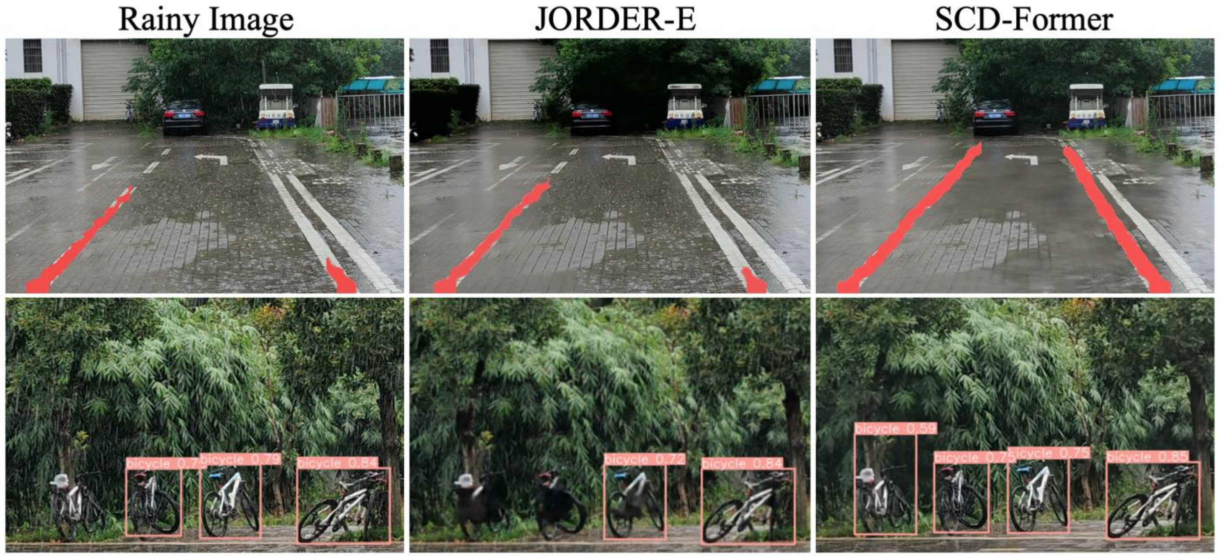
Experiment

train SCD-Former on different datasets

Rainy Image　Rain100L　SPA-data　GT-Rain　LHP-Rain

Experiment

Evaluation of high-level tasks



Method	Det. (mAP)	Gain (Det.)	Seg. (Acc)	Gain (Seg.)
Rainy	0.543	-	0.237	-
SPANet	0.563	+0.020	0.268	+0.031
PReNet	0.560	+0.022	0.255	+0.018
RCDNet	0.556	+0.018	0.361	+0.124
JORDER-E	0.568	+0.025	0.385	+0.148
MPRNet	0.560	+0.017	0.350	+0.113
Uformer-B	0.568	+0.025	0.306	+0.069
IDT	0.570	+0.027	0.365	+0.128
SCD-Former	**0.575**	**+0.031**	**0.449**	**+0.212**

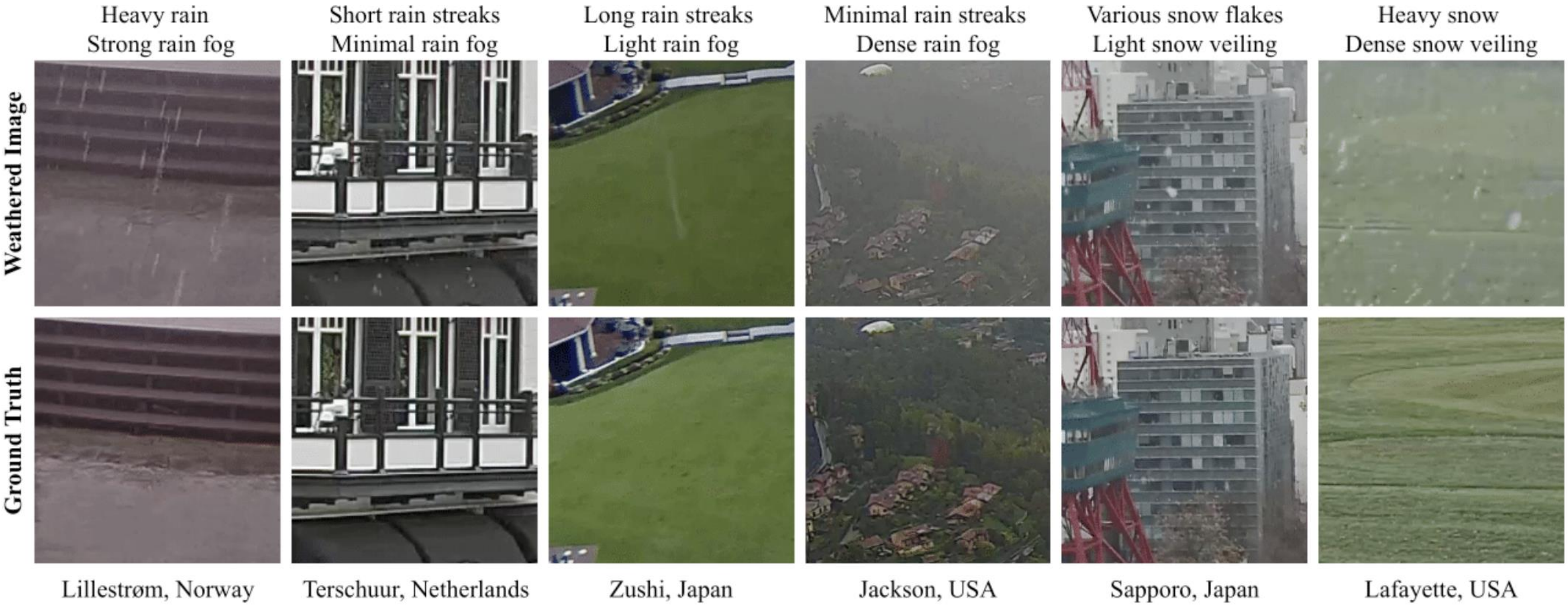
Outline

- Introduction
- LHP-Rain
  - Method
  - Experiment
- WeatherStream
  - Method
  - Experiment
- Conclusion

WeatherStream: Light Transport Automation of Single Image Deweathering

Howard Zhang[1*] Yunhao Ba[1*] Ethan Yang[1] Varan Mehra[1] Blake Gella[1]
Akira Suzuki[1] Arnold Pfahnl[1] Chethan Chinder Chandrappa[1]
Alex Wong[2] Achuta Kadambi[1]
[1]University of California, Los Angeles [2]Yale University

Example


Heavy rain
Strong rain fog
Short rain streaks
Minimal rain fog
Long rain streaks
Light rain fog
Minimal rain streaks
Dense rain fog
Various snow flakes
Light snow veiling
Heavy snow
Dense snow veiling
Weathered Image
Ground Truth
Lillestrøm, Norway
Terschuur, Netherlands
Zushi, Japan
Jackson, USA
Sapporo, Japan
Lafayette, USA

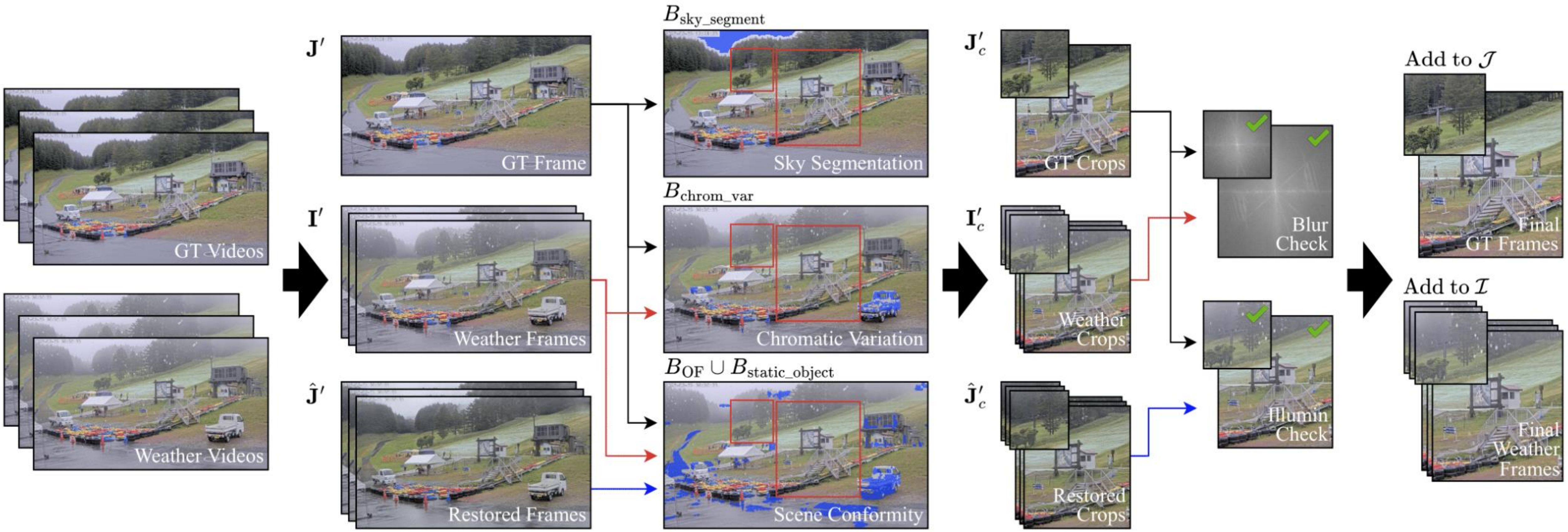
Example

(a) Snowy Input Image (b) Model [57] trained on Previous Dataset (c) **Model [57] trained on Our Dataset**

Existing dataset

Dataset	Multi-weather?	# Pairs	Time Multiplexed?	Scalable?
All-in-One [32]	✓	111.6K	✗	✓
Snow100k [40]	✗	100K	✗	✓
SRRS [8]	✗	15K	✗	✓
CSD [9]	✗	10K	✗	✓
Rain100L [69]	✗	300	✗	✓
Rain100H [69]	✗	1.9K	✗	✓
Outdoor-Rain [31]	✗	10.5K	✗	✓
RainCityscapes [24]	✗	10.62K	✗	✓
Rain12000 [73]	✗	13.2K	✗	✓
Rain14000 [17]	✗	14K	✗	✓
GT-RAIN [2]	✗	31.5K	✓	✗
WeatherStream (Ours)	✓	202K	✓	✓

Framework



$$\mathcal{D}_{\mathbf{c}} = \{\widetilde{\mathbf{c}}^n\}_{n=1}^{N}, \mathcal{D}_{\mathbf{d}} = \{\widetilde{\mathbf{d}}^m\}_{m=1}^{M}$$

$$f_{\text{WeatherStream}} : \mathcal{D} \to (\mathbf{c}, \mathbf{d}).$$

Principle 1 (Background Conformity)
Theory

Objects in a suitable paired scene exhibit no motion and color constancy.

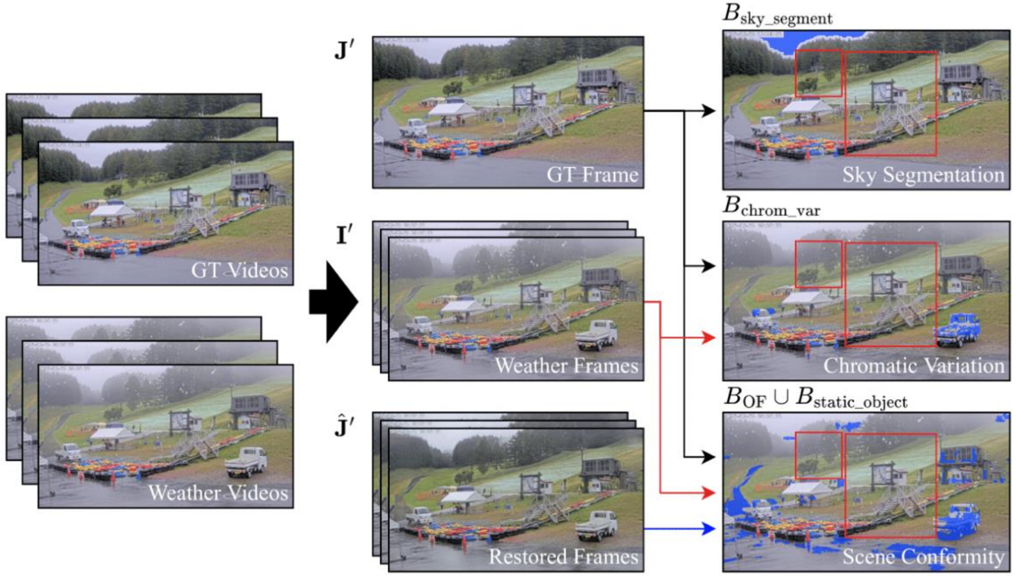
- Rain

$$\mathbf{K}_r(x) = \mathbf{J}(x) + \sum_{i}^{n} \mathbf{S}_i(x), \tag{1}$$

- Snow

$$\mathbf{K}_s(x) = \mathbf{J}(x)(1 - \mathbf{Z}(x)) + \mathbf{C}(x)\mathbf{Z}(x), \tag{2}$$

Principle 1 (Background Conformity)
Implementation



$$\mathbf{1}^{\tilde{\mathbf{c}},\hat{\mathbf{d}}}_{\text{static}}(x) = \begin{cases} 1 \text{ if } & |\tilde{\mathbf{c}}(x) - \hat{\mathbf{d}}(x)| > \gamma_{\text{static}}, \\ 0 \text{ otherwise.} \end{cases} \tag{12}$$

$$\hat{\mathbf{d}} = f_{\text{weather\_removal}}(\tilde{\mathbf{d}}), \tag{13}$$

$$\mathbf{B}_{\text{static}}(x) = \mathbf{1}^{\tilde{\mathbf{c}},\hat{\mathbf{d}}}_{\text{static}}(x) \text{ for } \tilde{\mathbf{c}} \in \mathcal{D}_{\mathbf{c}}, \tilde{\mathbf{d}} \in \mathcal{D}_{\mathbf{d}}. \tag{14}$$

- Filtered by **computing optical flow** between two temporally time frames
- Avoid **sky regions**, as these are not typically regions that contain rain, snow, or fog weather effects

$$\mathbf{B} = \underline{\mathbf{B}_{\text{static}}} \cup \underline{\mathbf{B}_{\text{OF}}} \cup \underline{\mathbf{B}_{\text{sky}}}$$

Principle 2 (Particle Chromatic Variation)

Theory

Rain streaks and snowflakes exhibit isotropic derivatives in RGB pixel intensity

$$I_p = \int_0^{\tau} E_p \, dt + \int_{\tau}^{T} E_{bg} \, dt, \quad (3)$$

$$I_p = \tau \bar{E}_p + (T - \tau) E_{bg}, \quad (4)$$

$$\bar{E}_p = \frac{1}{\tau} \int_0^{\tau} E_p \, dt, \quad (5)$$

$$\Delta I = \tau(\bar{E}_p - E_{bg}). \quad (6)$$

$$\Delta I = -\beta I_{bg} + \alpha, \beta = \frac{\tau}{T}, \alpha = \tau \bar{E}_p. \quad (7)$$

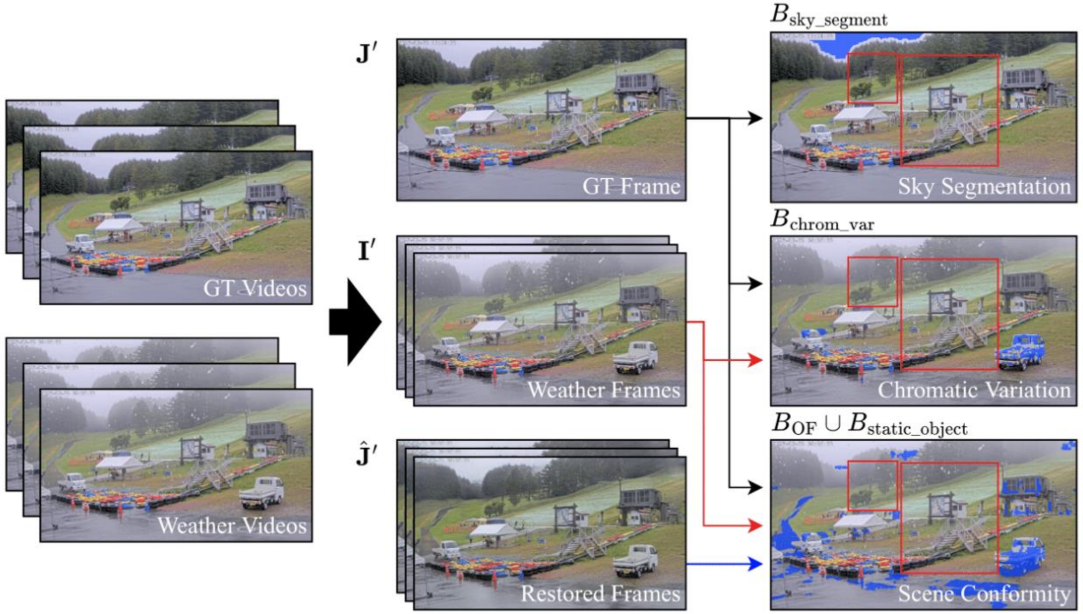
$\boxed{[\Delta R, \Delta G, \Delta B]^T \text{ should be isotropic}}$

Principle 2 (Particle Chromatic Variation)
Implementation

$$\mathbf{C}_{\max}(x) = \max_{c \in \{r,g,b\}} \widetilde{\mathbf{d}}_c(x) - \widetilde{\mathbf{c}}_c(x),$$

$$\mathbf{C}_{\min}(x) = \min_{c \in \{r,g,b\}} \widetilde{\mathbf{d}}_c(x) - \widetilde{\mathbf{c}}_c(x),$$

$$\mathbf{B}_{\text{chrom\_var}} = (\mathbf{C}_{\max} - \mathbf{C}_{\min}) > \gamma_{\text{chrom\_var}}.$$



- catches **static object differences** present in the image pairs while **preserving rain streaks and snowflakes**

Principle 3 (Scatter-dependent Blur)

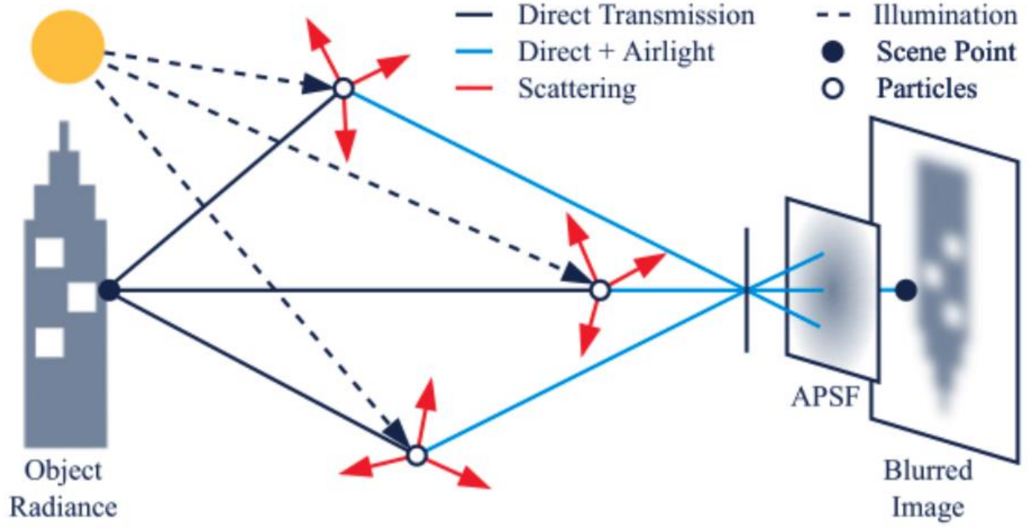
Theory

A degraded image is blurrier than a clean image, due to scattering effects.

$$\begin{cases} L_s(d) = L_o e^{-\int_0^d \beta dl} + \int_0^d L_\infty \beta e^{-\beta l} dl \\ \quad = \underline{L_o e^{-\beta d}} + L_\infty(1 - e^{-\beta d}). \end{cases} \quad (8)$$

$$L_m(d) = L_s(d) * \underline{\text{APSF}(d)}. \quad (9)$$

$$\Longrightarrow \mathbf{I}(x) = [\underline{\mathbf{K}(x) e^{-\beta d(x)}} + \mathbf{L}_\infty(1 - e^{-\beta d(x)})] * \underline{h_{\text{APSF}}}, \quad (10)$$



Principle 3 (Scatter-dependent Blur)
Implementation

- An approximation of the accompanying APSF as a generalized Gaussian distribution leads to consider two metrics

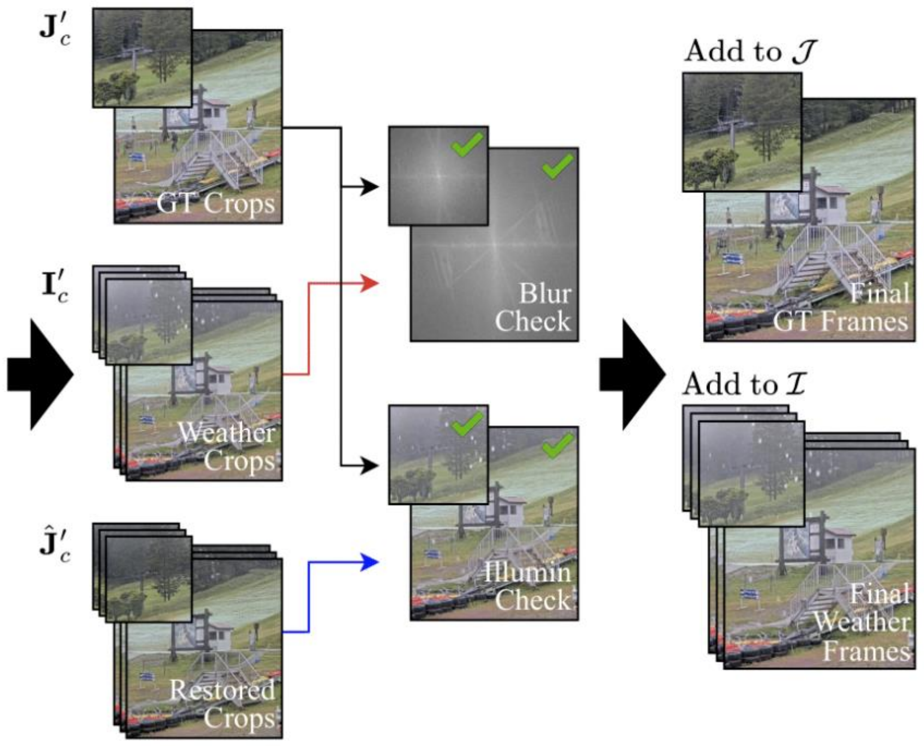
  - **Lower contrast**

$$\bar{\mathbf{d}} = \frac{1}{m}\sum_{i=0}^{m} \widetilde{\mathbf{d}}, \tag{18}$$

$$f_{\text{grad}} = \frac{|\nabla \bar{\mathbf{d}}| - |\nabla \widetilde{\mathbf{c}}|}{|\nabla \widetilde{\mathbf{c}}|}. \tag{19}$$

  - **Blur**

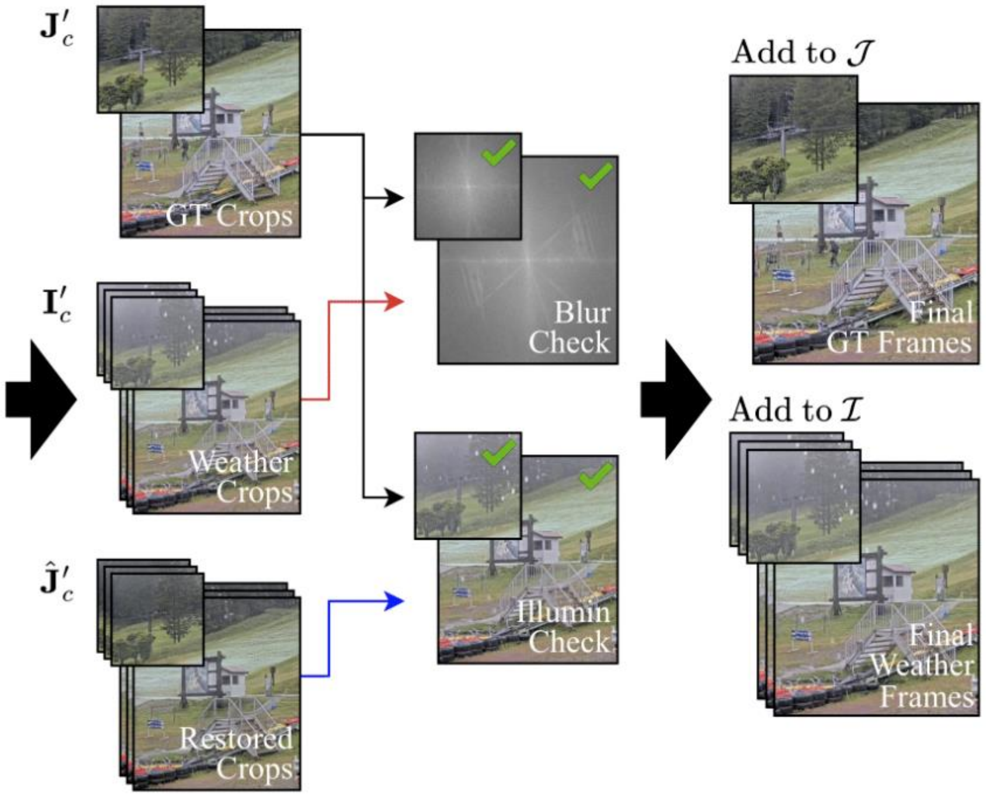
$$f_{\text{fft}} = \frac{|f_{\text{lp}}(\bar{\mathbf{d}}, \gamma_{\text{lp}})| - |f_{\text{lp}}(\widetilde{\mathbf{c}}, \gamma_{\text{lp}})|}{|f_{\text{lp}}(\widetilde{\mathbf{c}}, \gamma_{\text{lp}})|}.$$



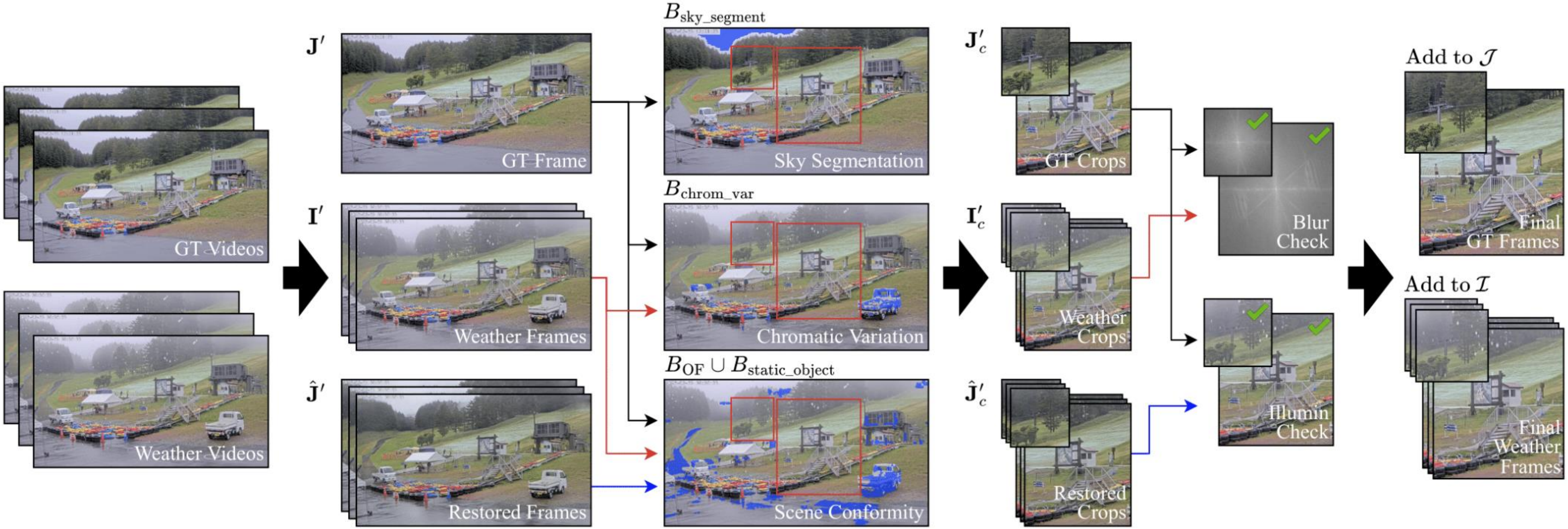
Principle 4 (Illumination Verification)
Theory & Implementation

The ambient illumination should remain consistent despite weather effects. This is difficult to analyze in closed-form since illumination stimuli to uncontrolled scenes are not known a priori.

- The **seed model** can be adapted to distinguish between illumination shifts and veiling effects, if the illumination is consistent between frames.

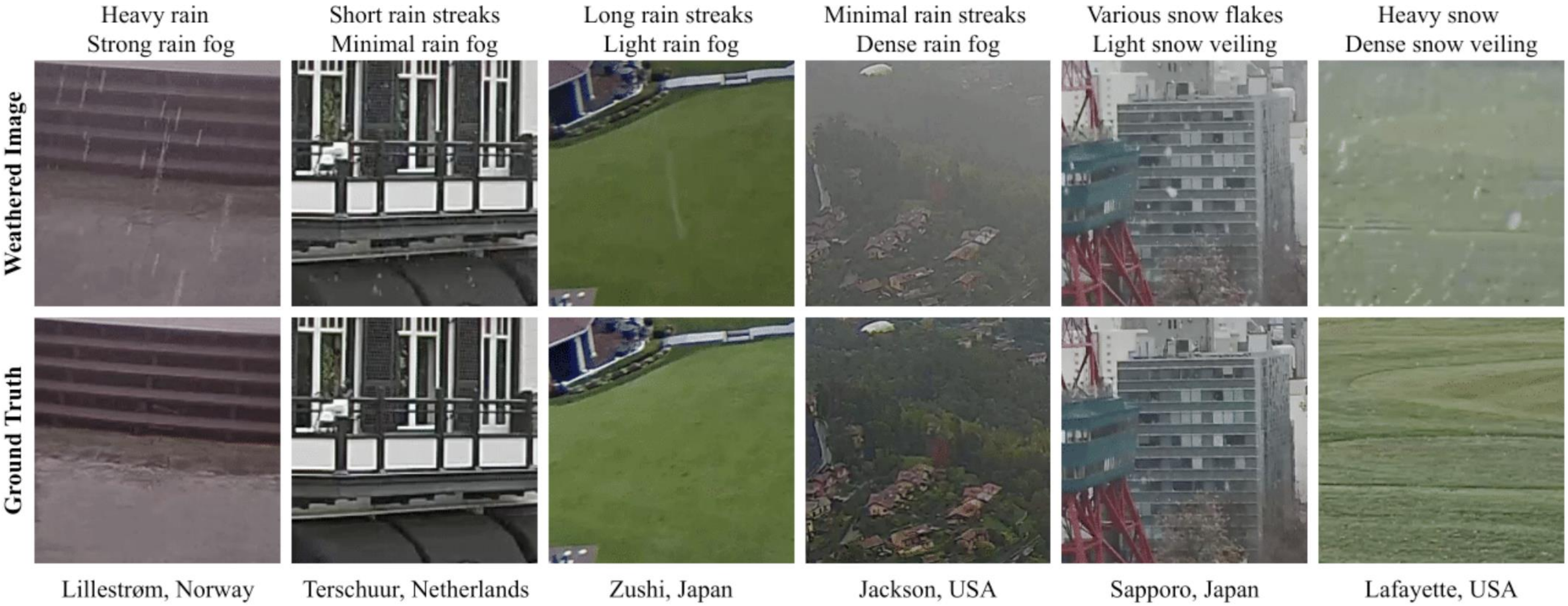


Framework



- Frames that remain in D_c and D_d are assigned as pairs based on temporal closeness within a hysteresis threshold.

Experiment

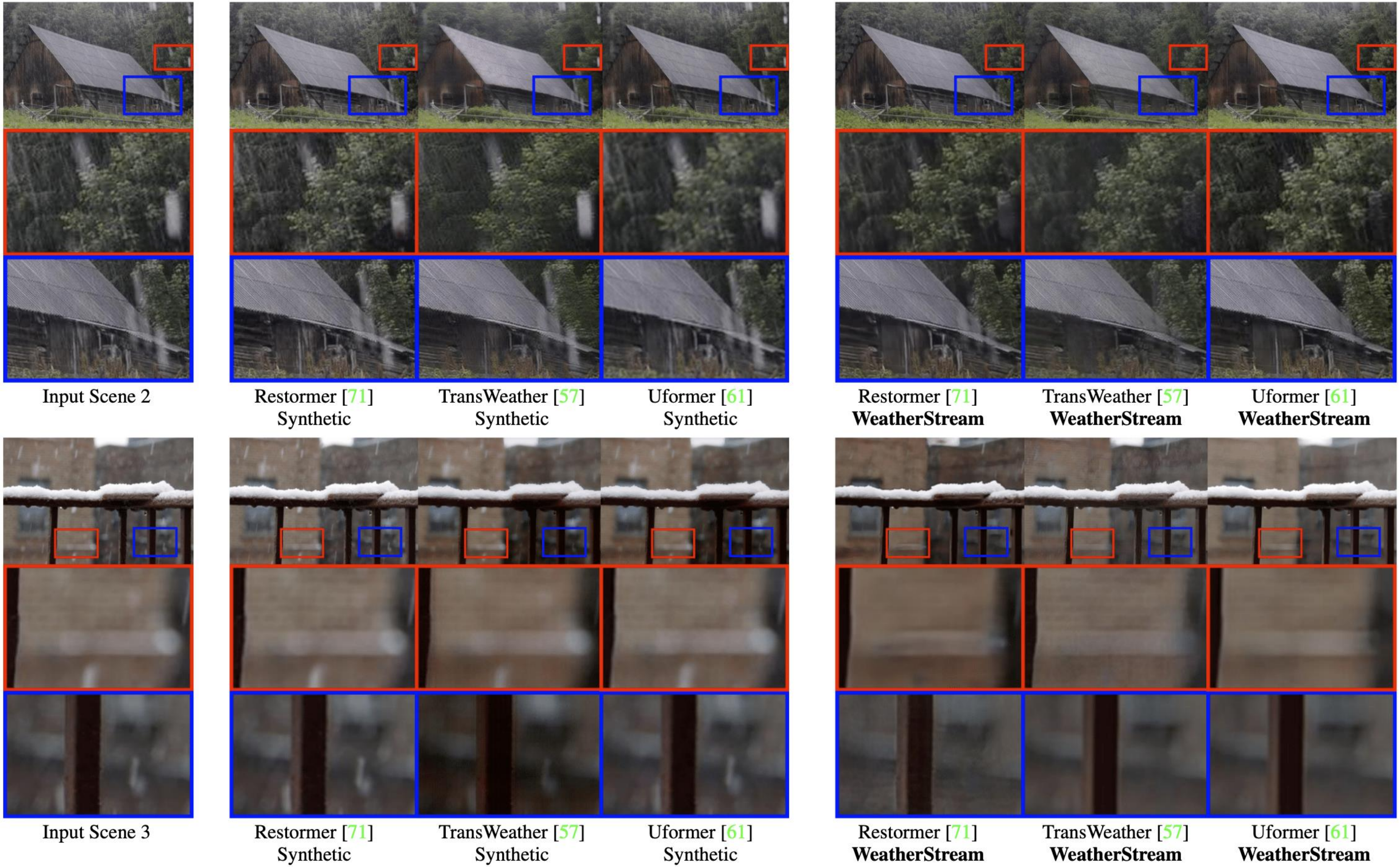


Heavy rain
Strong rain fog
Short rain streaks
Minimal rain fog
Long rain streaks
Light rain fog
Minimal rain streaks
Dense rain fog
Various snow flakes
Light snow veiling
Heavy snow
Dense snow veiling
Weathered Image
Ground Truth
Lillestrøm, Norway
Terschuur, Netherlands
Zushi, Japan
Jackson, USA
Sapporo, Japan
Lafayette, USA

Experiment

Improvements in PSNR and SSIM across many dataset

	Method	Rain		Fog		Snow		Overall	
		PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
Syn.	TransWeather [57] (CVPR'22)	21.36	0.7434	17.49	0.7118	21.04	0.7736	19.98	0.7416
	Restormer [71] (CVPR'22)	21.29	0.7664	18.22	0.7532	20.84	0.7890	20.14	0.7685
	Uformer [61] (CVPR'22)	22.15	0.7723	18.20	0.7519	20.66	0.7920	20.40	0.7712
	Rain-robust [2] (ECCV'22)	17.96	0.7075	17.13	0.6931	17.52	0.7318	17.56	0.7097
Manual	TransWeather [57] (CVPR'22)	21.73	0.7622	21.59	0.7612	21.42	**0.7932**	21.59	0.7708
	Restormer [71] (CVPR'22)	22.63	0.7940	20.14	0.7700	21.62	0.8123	21.51	0.7913
	Uformer [61] (CVPR'22)	22.13	0.7759	18.65	0.7445	20.91	0.7821	20.62	0.7672
	Rain-robust [2] (ECCV'22)	22.83	0.7887	20.95	0.7691	22.17	0.8058	22.01	0.7871
Ours	TransWeather [57] (CVPR'22)	**22.21**	**0.7716**	**22.55**	**0.7735**	**21.79**	0.7919	**22.20**	**0.7781**
	Restormer [71] (CVPR'22)	**23.67**	**0.8027**	**22.90**	**0.8029**	**22.51**	**0.8279**	**23.08**	**0.8100**
	Uformer [61] (CVPR'22)	**22.25**	**0.7911**	**18.81**	**0.7628**	**20.94**	**0.8009**	**20.72**	**0.7845**
	Rain-robust [2] (ECCV'22)	**23.43**	**0.7961**	**22.84**	**0.7901**	**22.29**	**0.8128**	**22.90**	**0.7989**


Input Scene 2
Restormer [71] Synthetic
TransWeather [57] Synthetic
Uformer [61] Synthetic
Restormer [71] **WeatherStream**
TransWeather [57] **WeatherStream**
Uformer [61] **WeatherStream**
Input Scene 3
Restormer [71] Synthetic
TransWeather [57] Synthetic
Uformer [61] Synthetic
Restormer [71] **WeatherStream**
TransWeather [57] **WeatherStream**
Uformer [61] **WeatherStream**

Outline

- Introduction
- LHP-Rain
  - Method
  - Experiment
- WeatherStream
  - Method
  - Experiment
- Conclusion

Comparation Review

- Ways to construct the pairs
  - video-based generation (LHP-Rain)
  - time-interval acquisition (WeatherStream)
- Source
  - Smartphone (LHP-Rain)
  - Youtube stream (WeatherStream)
- Categories
  - Diverse Rain, including streak, veiling, occlusion, splashing (LHP-Rain)
  - Rain and Snow (WeatherStream)
- Amount
  - 1.0M pairs (LHP-Rain)
  - 202K pairs (WeatherStream)

Conclusion

- From Sky to the Ground: A Large-scale Benchmark and Simple Baseline Towards Real Rain Removal
  - provides **diverse rain categories**, especially the ground splashing
  - proposed **low-rank tensor recovery model** could generate high-quality GT and detailed analysis confirms better results than others

- WeatherStream: Light Transport Automation of Single Image Deweathering
  - dataset improves almost all PSNR and SSIM metrics for all weather modalities, with the Restormer model notably improving by 1.49dB.