中国科学技术大学
University of Science and Technology of China

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

PAPER TAG: WED-AM-085

Zero-Shot Dual-Lens Super-Resolution

Ruikang Xu* Mingde Yao* Zhiwei Xiong†
University of Science and Technology of China

VIDAR
Visual Information Discovery And Recovery

NEL-BITA
National Engineering Laboratory for Brain-Inspired Intelligence Technology and Application

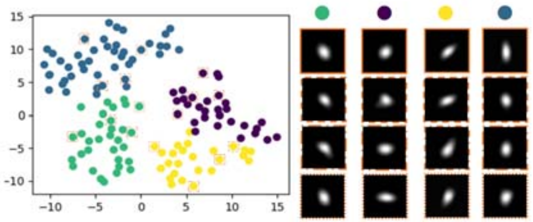
Zero-Shot Dual-Lens Super-Resolution

——OVERVIEW——

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Motivation

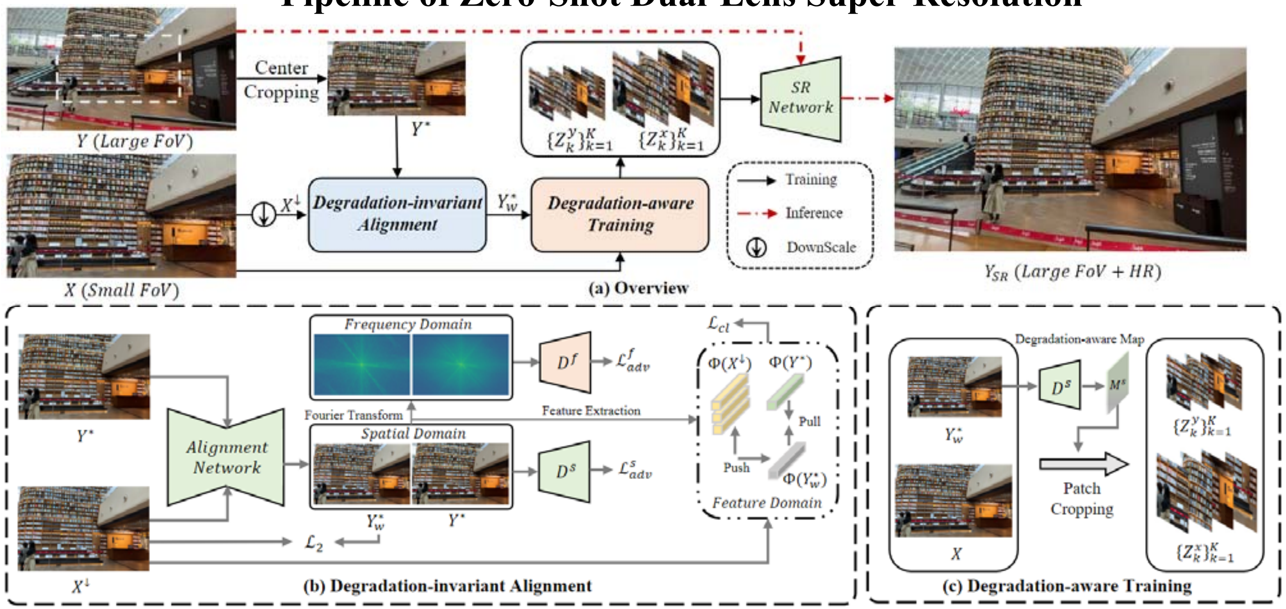
Dual-lens camera & Image-specific Degradation



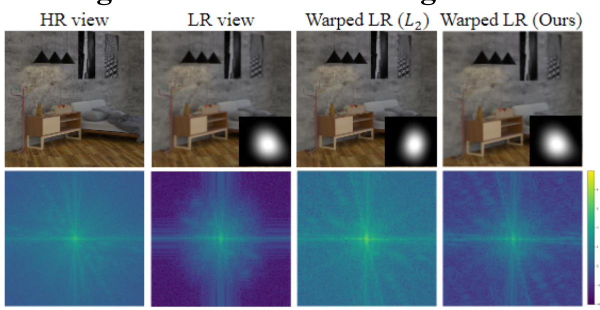
Comparison to existing Dual-lens SR

Method

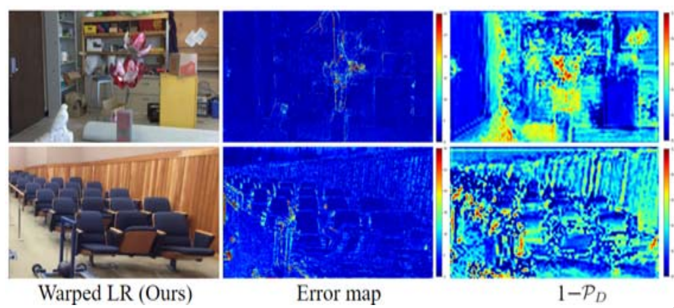
Pipeline of Zero-Shot Dual-Lens Super-Resolution



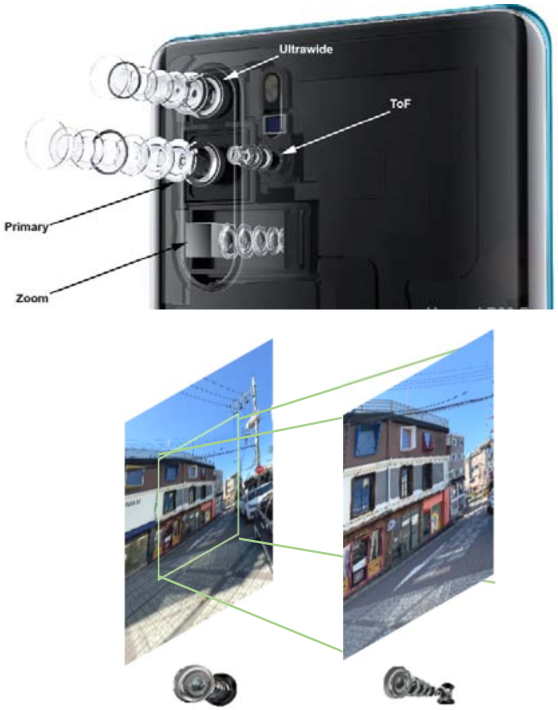
Degradation-invariant Alignment



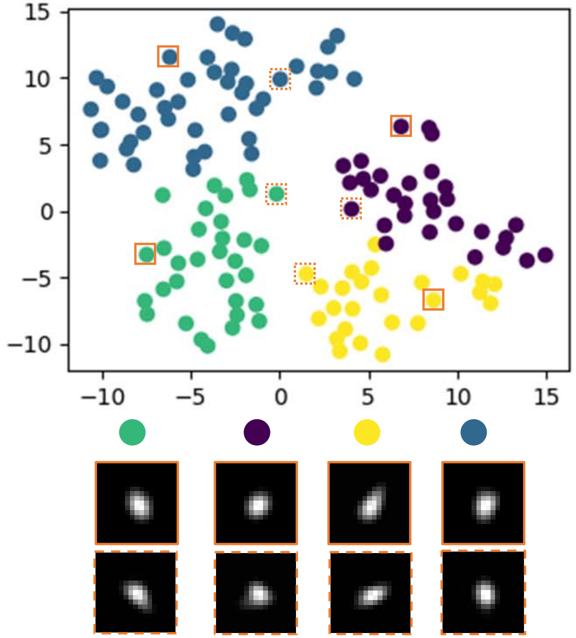
Degradation-aware training



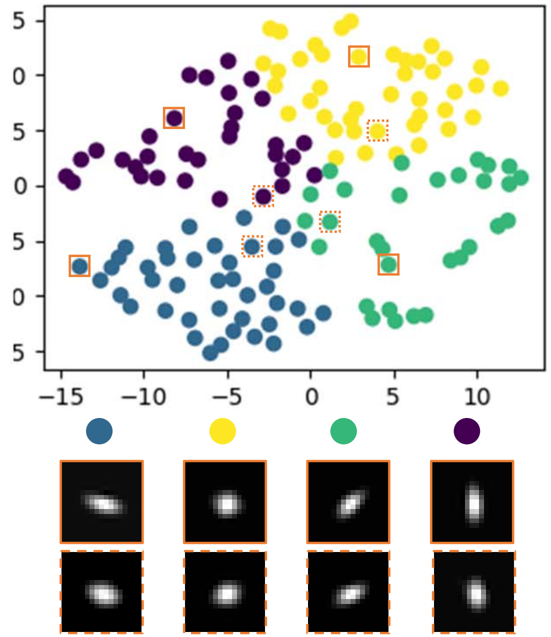
Background



Dual-lens camera



IPhone11 dataset



IPhone12 dataset

Image-Specific Degradation

Compare to Existing Dual-lens SR

Wide-angle Telephoto

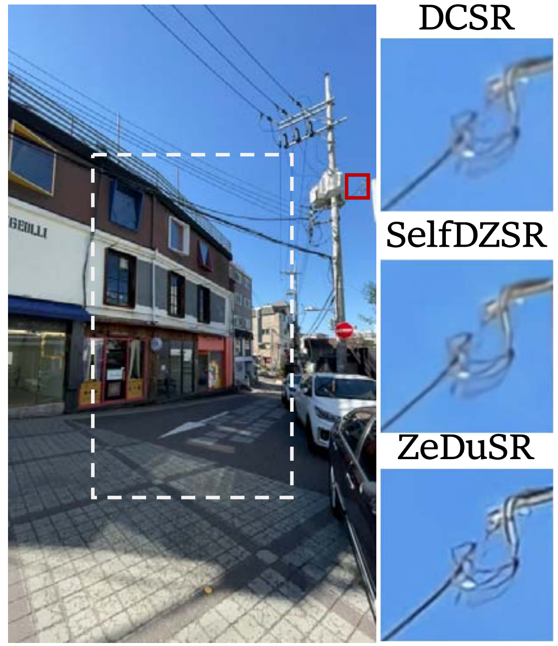
Input

DCSR[1]
- Predefined degradation
- Extra dual-lens dataset
- Supervised learning

SelfDZSR[2]
- Device-consistent degradation
- Extra dual-lens dataset
- Self-supervised learning

ZeDuSR
- Image-specific degradation
- No extra dual-lens dataset
- Zero-shot learning

Dual-lens SR methods



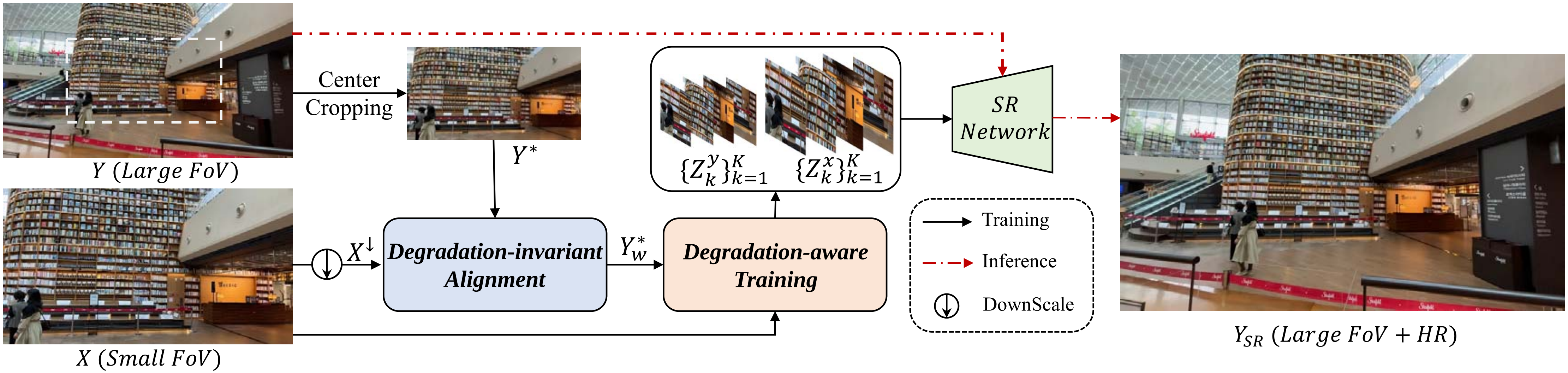
Result

[1] Tengfei Wang et.al. *Dual-camera super-resolution with aligned attention modules.* In ICCV2021.

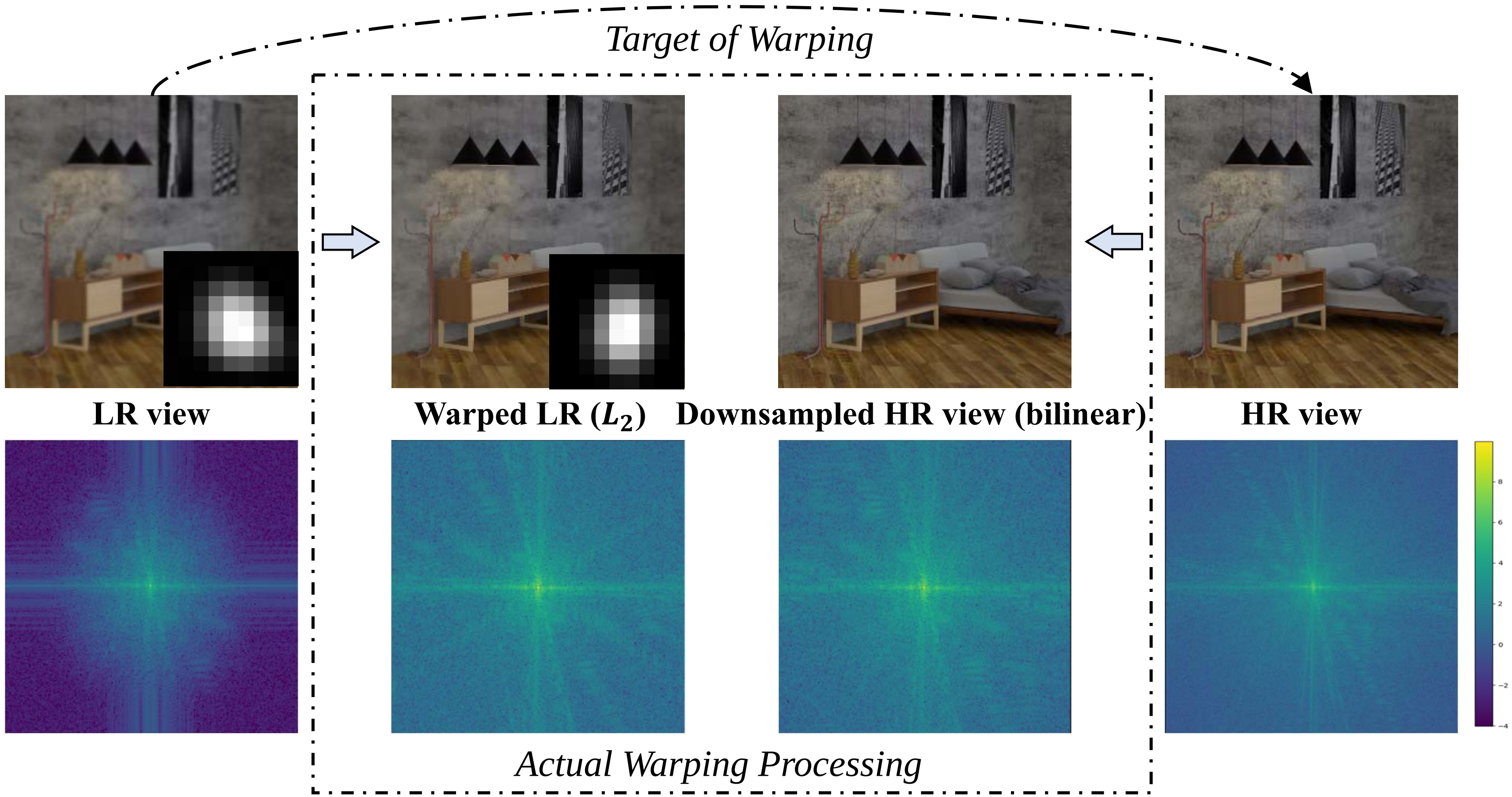
[2] Zhilu Zhang et.al. *Self-supervised learning for real-world super-resolution from dual zoomed observations.* In ECCV2022.

中国科学技术大学
University of Science and Technology of China

Zero-Shot Dual-Lens Super-Resolution

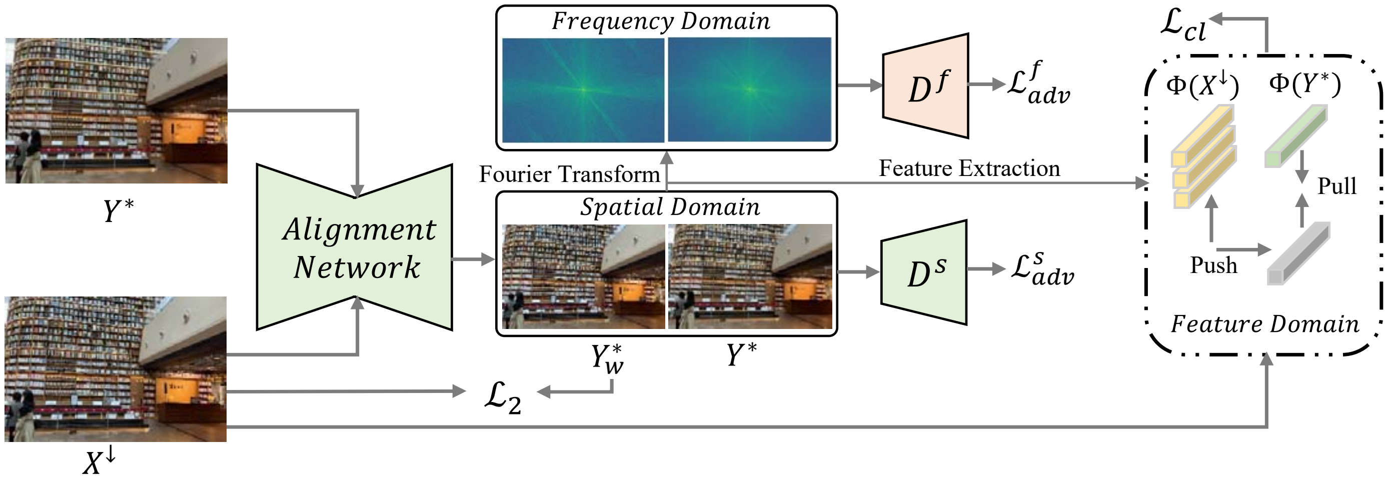


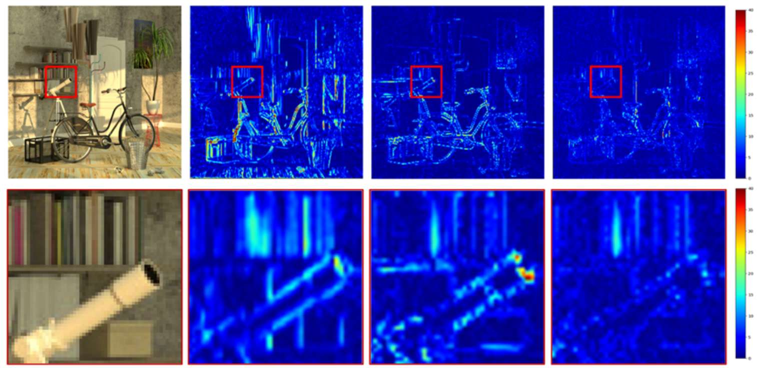
Observation



中国科学技术大学 University of Science and Technology of China

Degradation-invariant Alignment





$$\mathcal{L}_{align} = \mathcal{L}_2 + \lambda_1 \mathcal{L}_{adv}^s + \lambda_2 \mathcal{L}_{adv}^f + \lambda_3 \mathcal{L}_{cl}$$

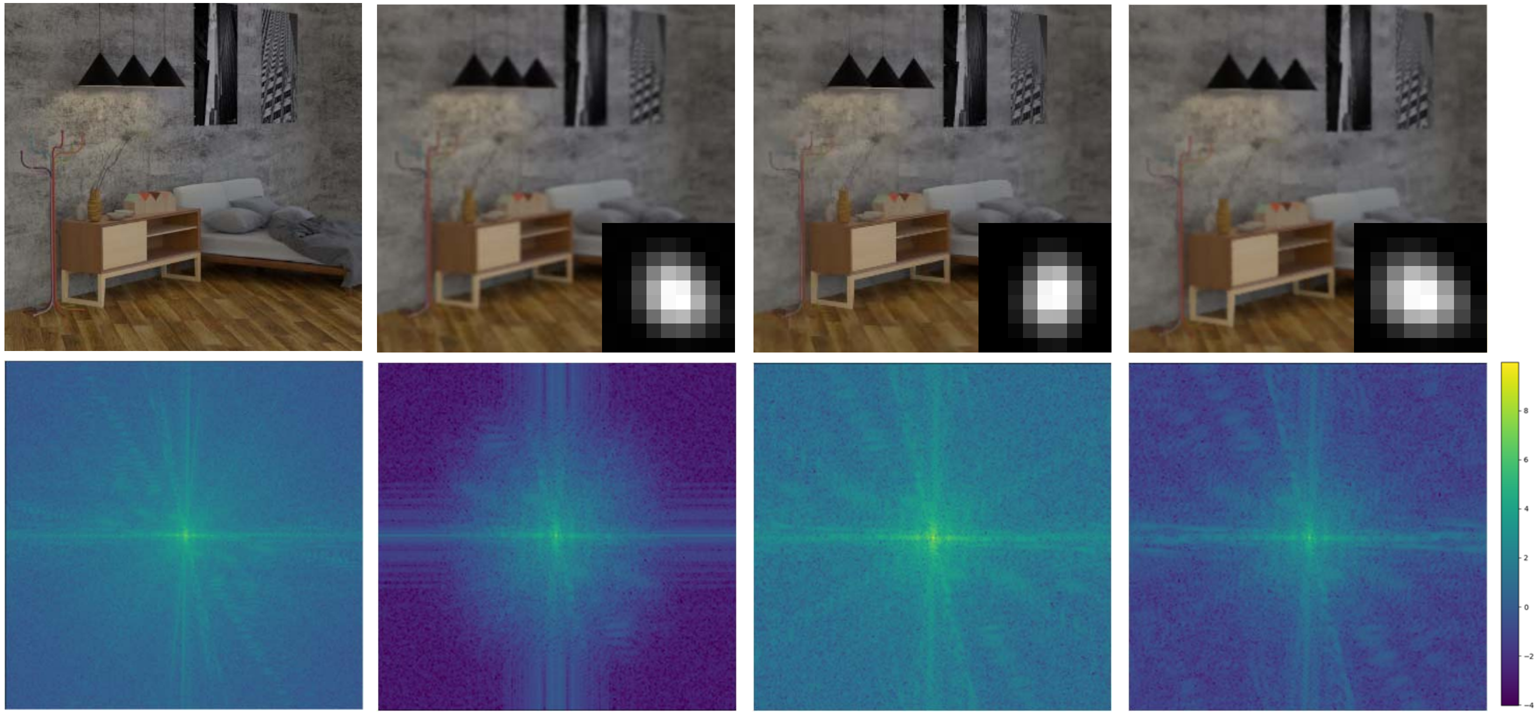
$$\mathcal{L}_2 = \|X^{\downarrow} - Y_w^*\|_2$$

$$\mathcal{L}_{adv}^f = -\mathbb{E}_{Y_w^*}[\log(D^f(\mathcal{A}[Y_w^*]))]$$

$$\mathcal{L}_{adv}^s = -\mathbb{E}_{Y_w^*}[\log(D^s(Y_w^*))]$$

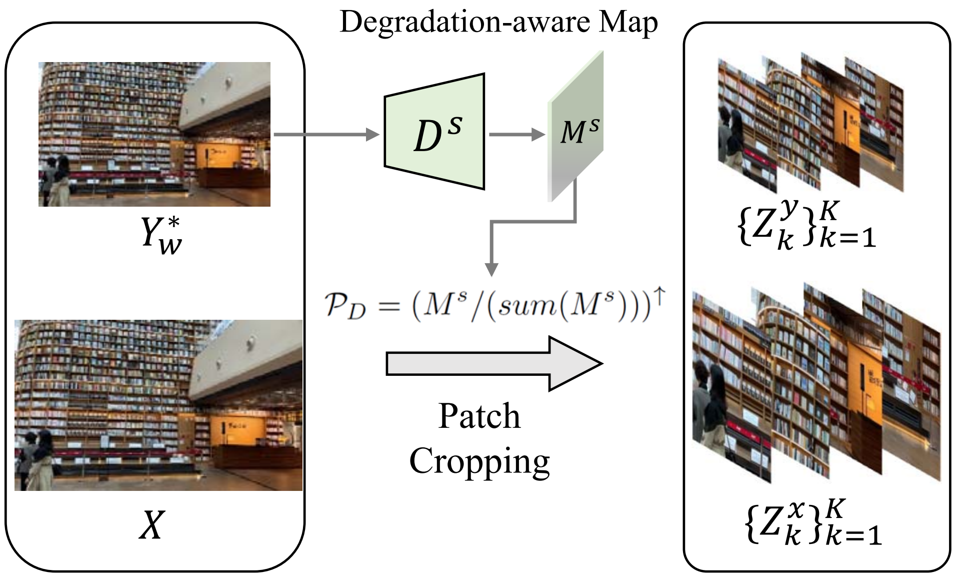
$$\mathcal{L}_{cl} = \frac{\sum_{i=1}^{n} \|\Phi_i(Y_w^*), \Phi_i(Y^*)\|_2}{\sum_{j=1}^{m} \|\Phi_j(Y_w^*), \Phi_j(X^{\downarrow})\|_2}$$

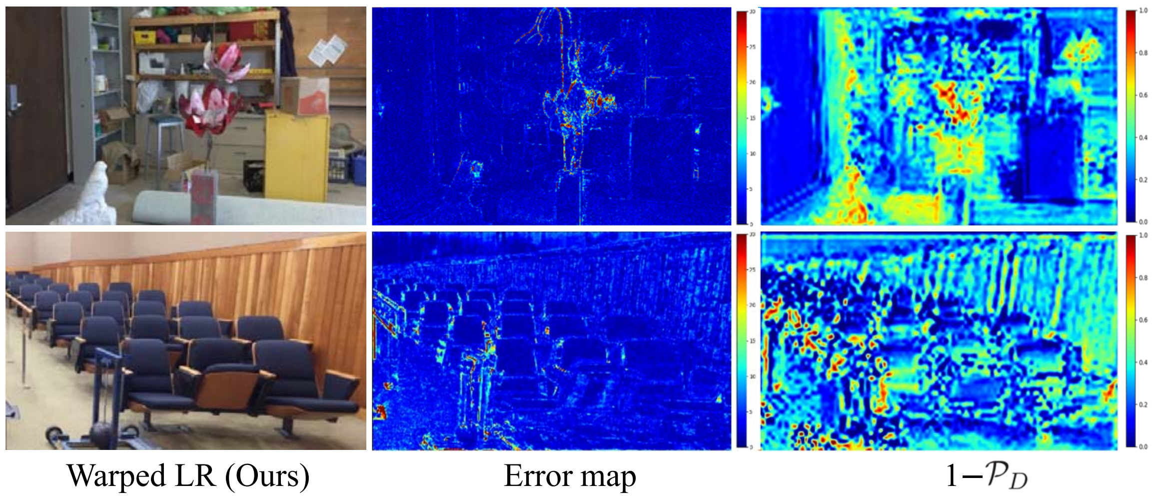
Degradation-invariant Alignment



HR view | LR view | Warped LR (L_2) | Warped LR (Ours)

Degradation-aware Training





Quantitative comparisons on Synthesized Data

Method	HCI_new				Middlebury2021			
	IG		AG		IG		AG	
	2× scale	4× scale	2× scale	4× scale	2× scale	4× scale	2× scale	4× scale
Bicubic	29.45/0.8119	27.62/0.7168	28.27/0.7505	27.33/0.7003	33.43/0.9409	30.87/0.9001	32.44/0.9267	30.33/0.8911
RCAN [52]	30.15/0.8371	28.23/0.7415	28.99/0.7670	28.00/0.7191	34.11/0.9465	31.84/0.9084	33.17/0.9326	31.30/0.9053
CSNLN [31]	30.24/0.8376	28.25/0.7379	28.94/0.7682	27.97/0.7172	34.23/0.9487	31.89/0.9136	33.10/0.9366	31.39/0.9040
SwinIR [22]	30.19/0.8329	28.27/0.7405	28.84/0.7652	27.90/0.7140	34.28/0.9492	31.96/0.9144	33.05/0.9361	31.27/0.9022
ZSSR [36]	30.29/0.8236	27.75/0.7186	28.63/0.7580	27.42/0.7041	33.93/0.9453	30.92/0.9021	32.91/0.9331	30.49/0.8925
KernelGAN [1]	-	-	29.78/0.8042	28.36/0.7358	-	-	33.65/0.9399	31.83/0.9032
TTSR [46]	-	28.03/0.7367	-	27.67/0.7082	-	31.39/0.9045	-	30.97/0.8991
MASA [25]	-	28.32/0.7498	-	28.05/0.7212	-	32.01/0.9094	-	31.42/0.9011
DCSR [43]	30.40/0.8306	28.38/0.7440	29.22/0.7627	27.99/0.7176	34.29/0.9434	32.02/0.9073	33.18/0.9382	31.36/0.8992
SelfDZSR [53]	29.86/0.8382	27.91/0.7297	28.97/0.7943	27.60/0.7244	33.91/0.9442	31.32/0.9132	32.96/0.9405	31.03/0.9033
DCSR+SRA [43]	30.61/0.8424	28.56/0.7486	29.44/0.7910	28.18/0.7317	34.37/0.9468	32.17/0.9124	33.38/0.9404	31.59/0.9037
ZeDuSR	31.01/0.8529	28.87/0.7536	30.02/0.8146	28.66/0.7420	34.78/0.9553	32.41/0.9117	33.79/0.9421	31.76/0.9019
ZeDuSR*	**31.17/0.8594**	**29.25/0.7601**	**30.23/0.8183**	**29.09/0.7483**	**34.89/0.9571**	**32.77/0.9212**	**33.94/0.9450**	**32.42/0.9141**

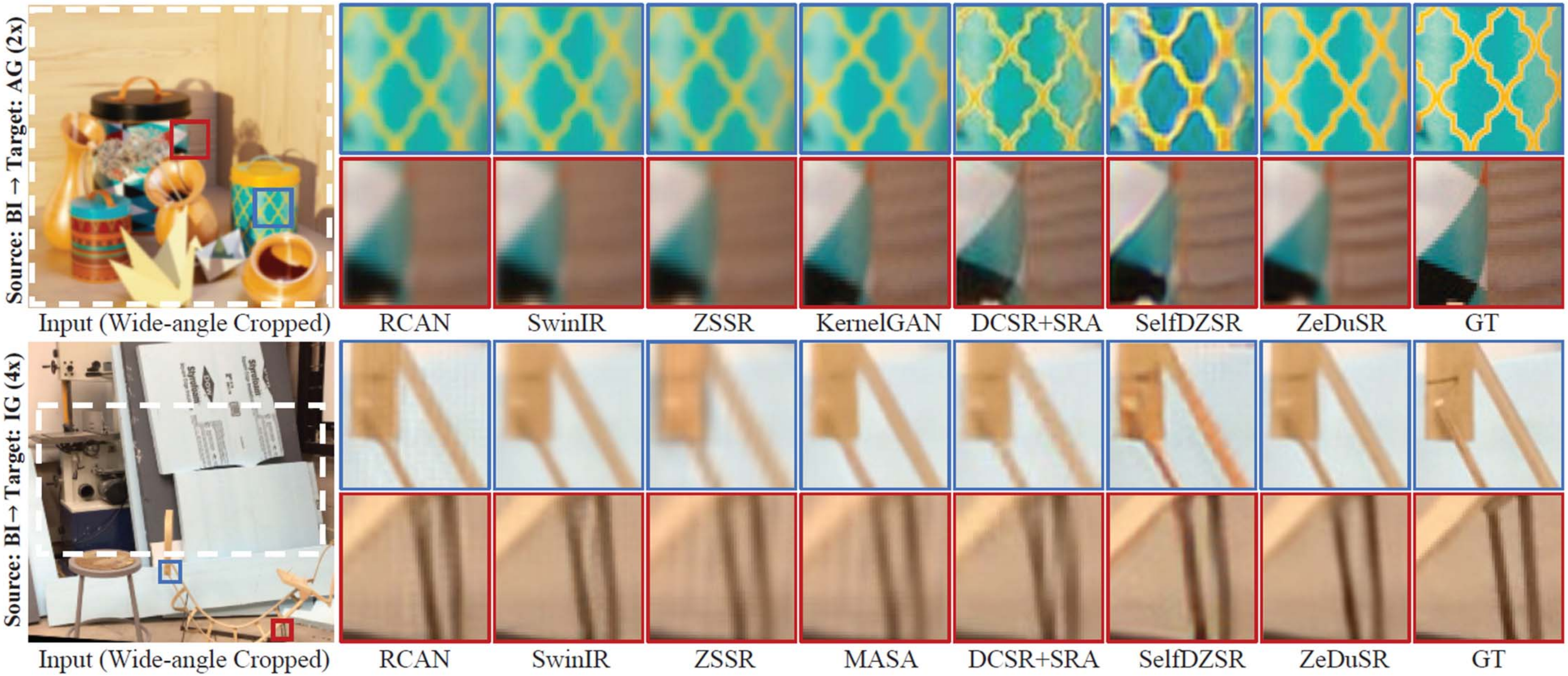
- Comparisons with SISR, RefSR, and Dual-lens SR for 2× and 4× SR on Synthesized Dual-lens Data with image-specific degradation (isotropic and anisotropic Gaussian downsampling (IG and AG)).
- ZeDuSR shows superior performance over the previous methods in most cases, thanks to the image-specific degradation assumption.

Quantitative comparisons on Synthesized Data

Method	HCI_new				Middlebury2021			
	IG		IG_JPEG		IG		IG_JPEG	
	2× scale	4× scale	2× scale	4× scale	2× scale	4× scale	2× scale	4× scale
Bicubic	29.45/0.8119	27.62/0.7168	29.42/0.7839	27.21/0.6881	33.43/0.9409	30.87/0.9001	33.02/0.9277	30.41/0.8822
DANv1 [27]	30.29/0.8569	28.43/0.7534	29.07/0.7747	27.19/0.6917	34.44/0.9519	32.22/0.9159	33.10/0.9182	30.45/0.8720
DANv2 [28]	30.16/0.8539	28.48/0.7586	28.82/0.7721	27.28/0.6911	34.36/0.9516	32.29/0.9150	32.97/0.9169	30.56/0.8695
DCLS [26]	30.63/0.8651	28.68/0.7524	29.21/0.7841	27.45/0.6928	34.72/0.9528	32.46/0.9173	33.38/0.9250	30.71/0.8725
ZeDuSR	31.01/0.8529	28.87/0.7536	29.98/0.8006	27.82/0.7035	34.78/0.9553	32.41/0.9171	33.63/0.9316	30.97/0.8805
ZeDuSR*	**31.17/0.8594**	**29.25/0.7601**	**30.06/0.8032**	**28.17/0.7179**	**34.89/0.9571**	**32.77/0.9212**	**33.79/0.9332**	**31.57/0.8929**

- Quantitative comparisons with blind SR for 2× and 4× SR on synthesized dual-lens data with image-specific degradation (isotropic Gaussian downsampling with slight JPEG compression (IG JPEG)).
- ZeDuSR shows superior performance over the previous methods in most cases, thanks to the image-specific degradation assumption.

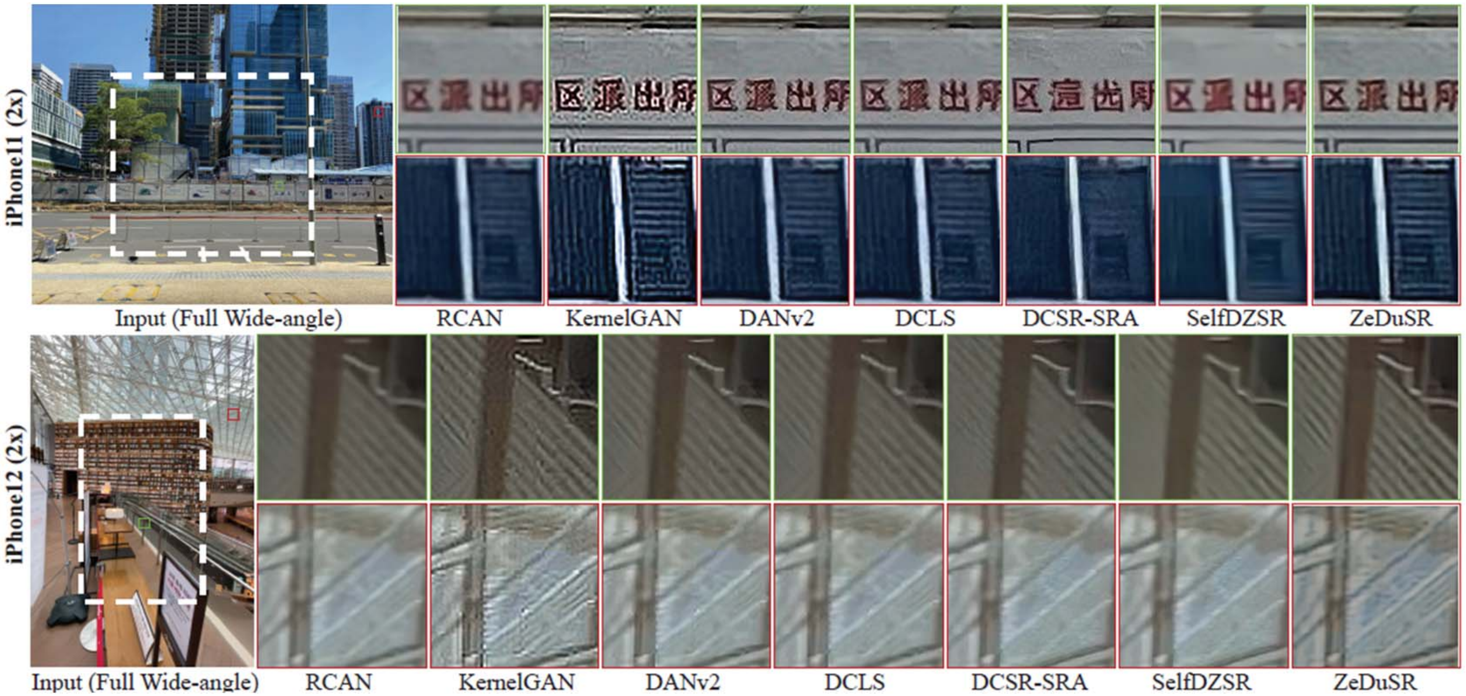
Visual comparisons on Synthesized Data

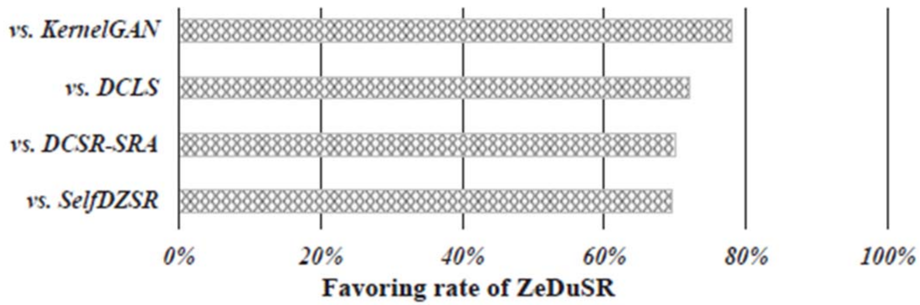


- The white dotted box indicates the overlapped FoV.
- Top: “Origami” from HCI new. Bottom: “Bandsaw1” from Middlebury2021.

University of Science and Technology of China

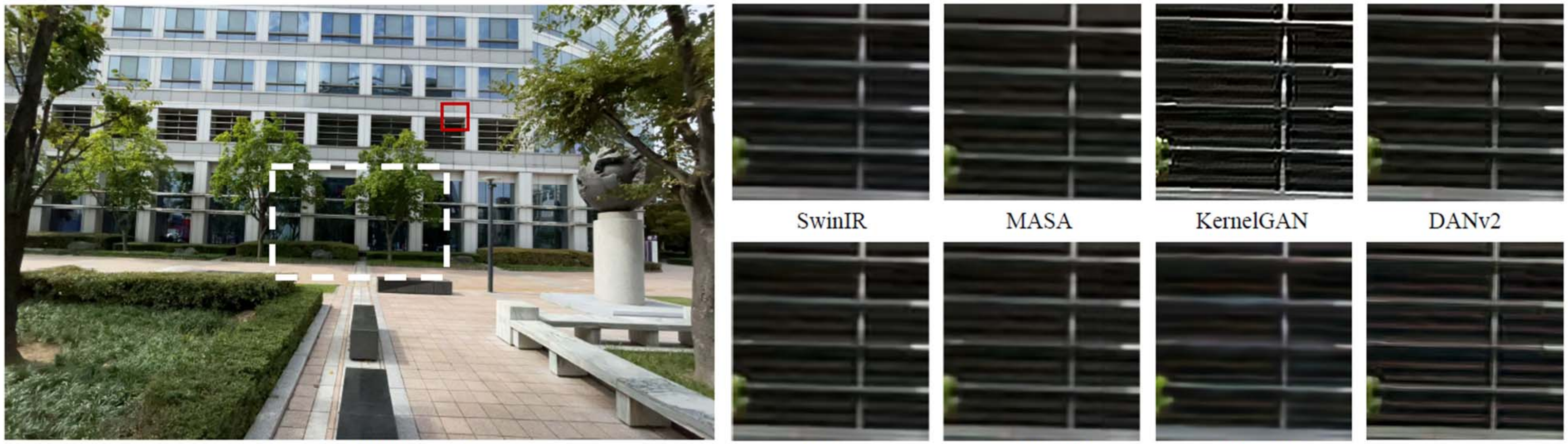
Comparisons on Real-world Data





- Visual comparisons on real-world data.
- User study on the dual-lens SR results on real-world data.

Comparisons on Real-world Data



中国科学技术大学 University of Science and Technology of China

Ablation Studies

Ablation on the Alignment backbone.

Alignment Network	IG	AG
FlowNet-S [8]	31.01/0.8529	30.02/0.8146
SPyNet [34]	31.05/0.8533	29.97/0.8139
PWCNet [39]	29.98/0.8515	30.06/0.8159

Ablation on the SR backbone.

SR Network	IG	AG
RCAN [52]	30.15/0.8371	28.99/0.7670
RCAN [52] + ZeDuSR	31.01/0.8529	30.02/0.8146
CSNLN [31]	30.24/0.8376	28.94/0.7682
CSNLN [31] + ZeDuSR	31.03/0.8516	29.97/0.8135
SwinIR [22]	30.19/0.8329	28.84/0.7652
SwinIR [22] + ZeDuSR	31.04/0.8531	29.91/0.8128

Ablation on the loss and training strategy.

L^s_{adv}	L^f_{adv}	L_{cl}	DaTS	PSNR	SSIM
✗	✗	✗	✗	30.48	0.8362
✓	✗	✗	✗	30.59	0.8429
✓	✓	✗	✗	30.72	0.8452
✓	✓	✓	✗	30.89	0.8512
✓	✓	✓	✓	31.01	0.8529

- For training the alignment network PSNR gradually increases by adding proposed losses.
- For training the image-specific SR network, the PSNR further increases when using DaTS.
- ZeDuSR is robust for different backbones, as performances of embodiments are close.

Conclusion

- *Zero-shot learning solution for realistic SR on dual-lens devices.*
- *Degradation-invariant alignment to constrain the alignment process.*
- *Degradation-aware training to fully exploit the highly limited data.*
- *Effective on synthesized and real-world experiments.*

Thanks!