

OSRT: Omnidirectional Image Super-Resolution with Distortion-aware Transformer

Fanghua Yu^{1*}, Xintao Wang^{2*}, Mingdeng Cao^{2,3}, Gen Li⁴, Ying Shan², Chao Dong^{1,5†}

¹SIAT, Chinese Academy of Sciences

²ARC, Tencent PCG

³The University of Tokyo

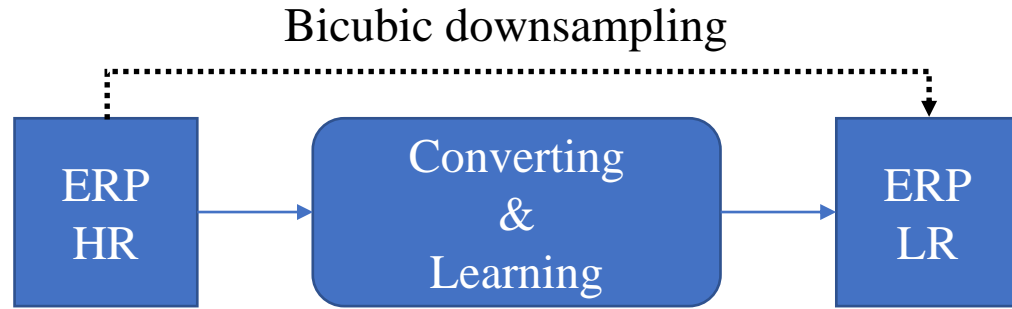
⁴Platform Technologies, Tencent Online Video

⁵Shanghai AI Lab



Codes & Models

Omnidirectional Image Super-Resolution (ODISR)



Previous ODISR methods treated ODIs only as images that followed a specific distribution, ignoring their geometric properties.

Manual generated ERP LR



- Uniform pixel density
- Sharp polar areas

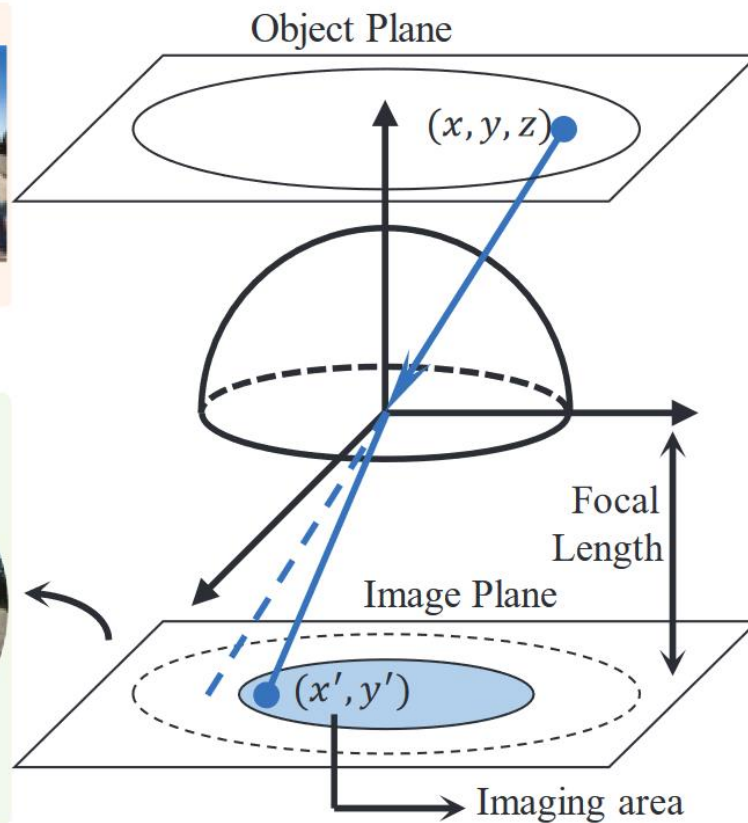
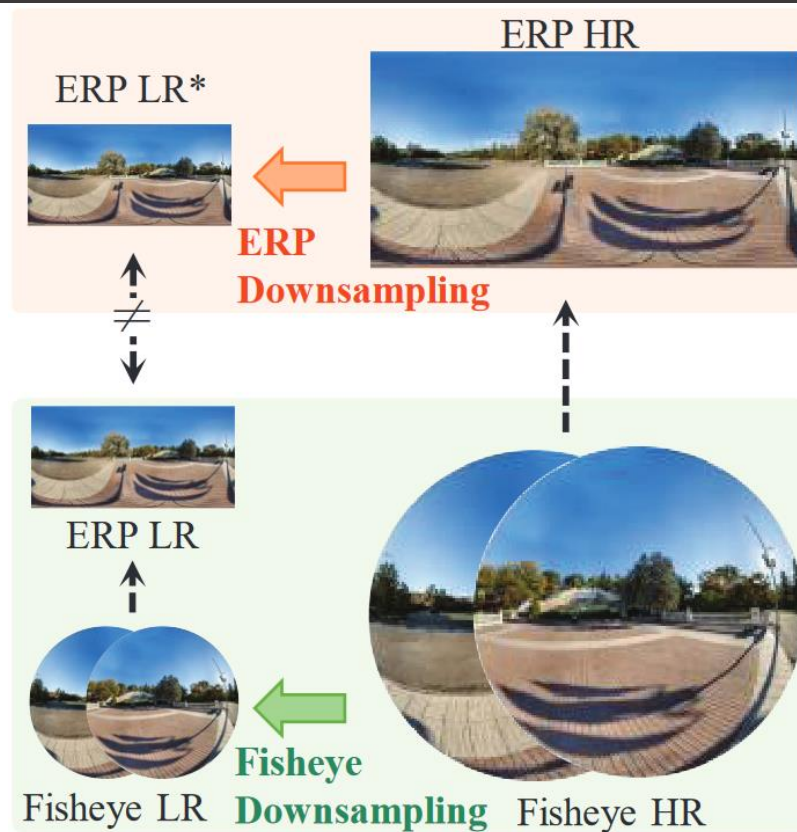
Poor generalization

LR in real scenarios



- Nonuniform pixel density
- Blur polar areas

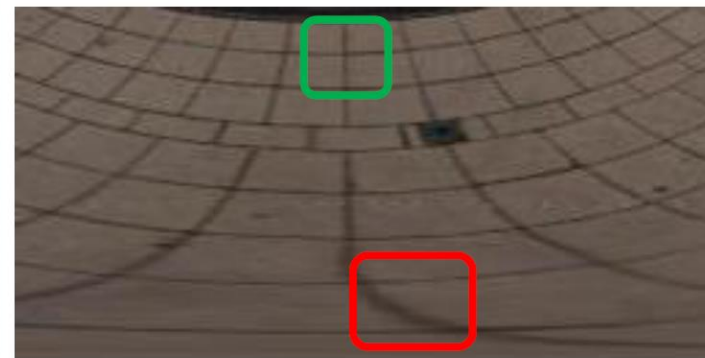
Fisheye downsampling



ERP DS

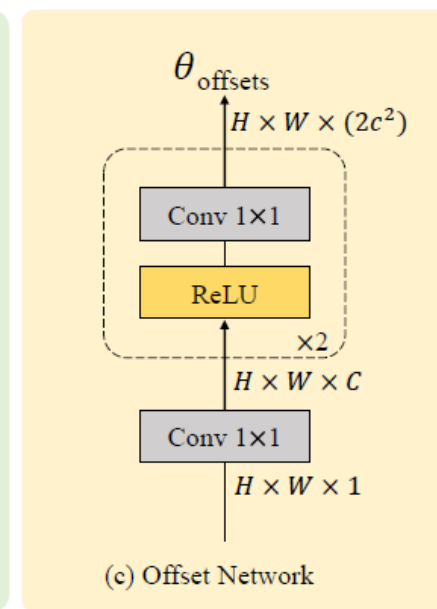
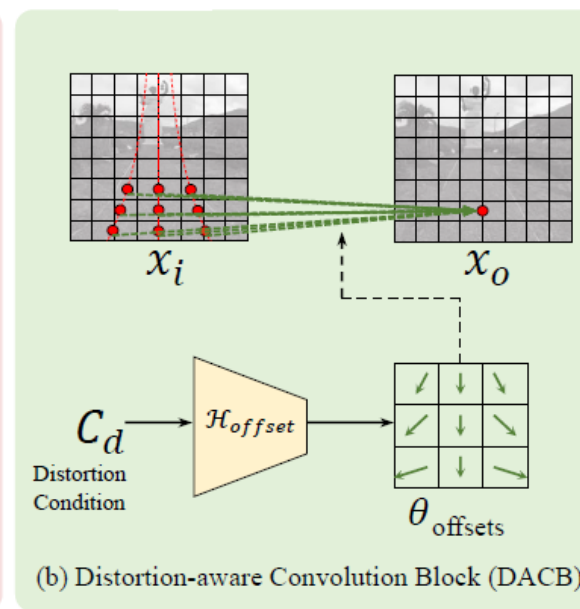
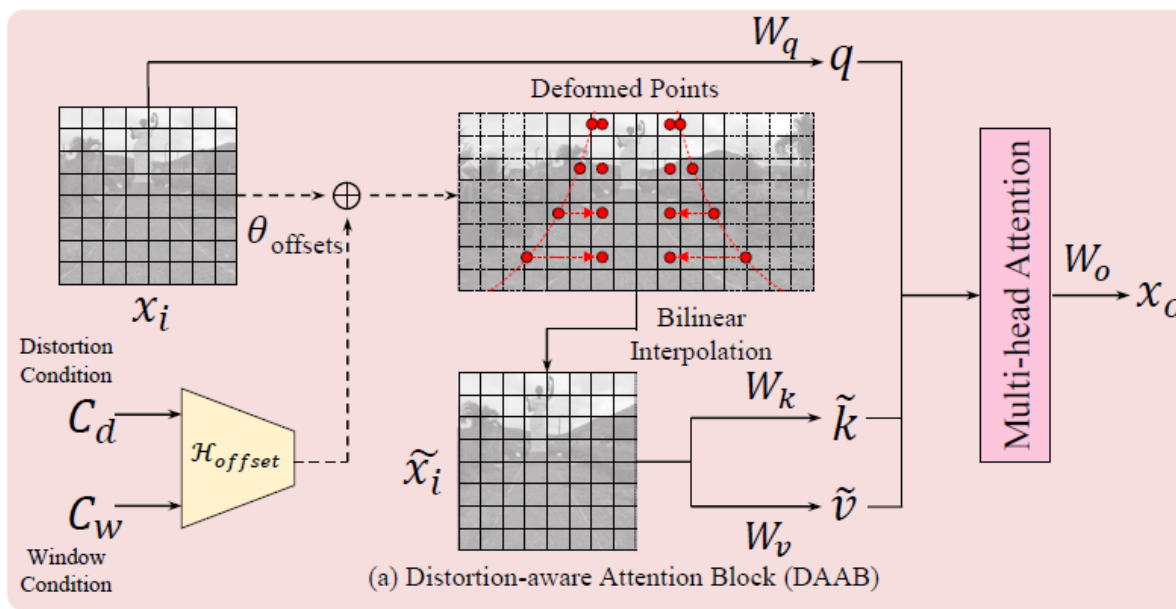
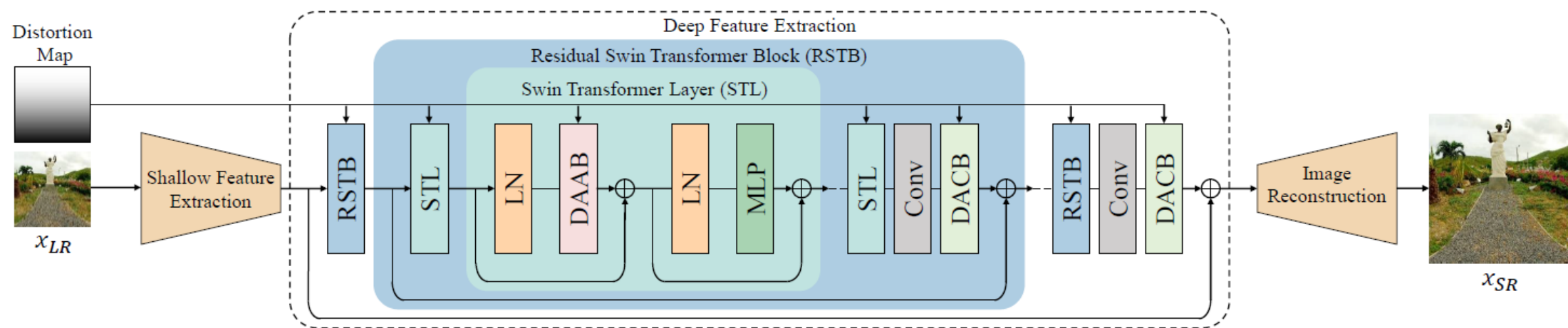


Fisheye DS



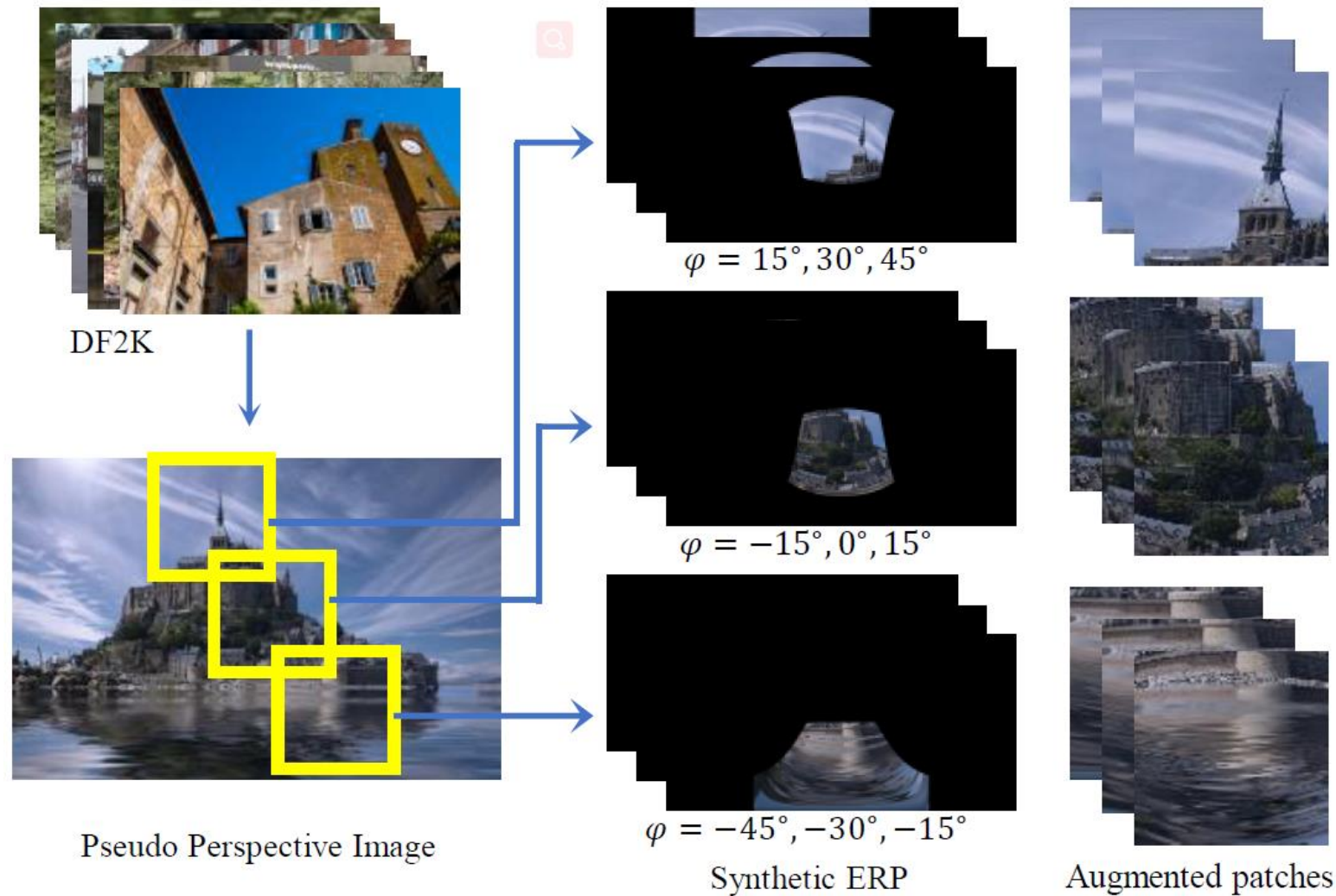
Real-world ERP

OSRT: rectifying ERP distortions by deformable operations

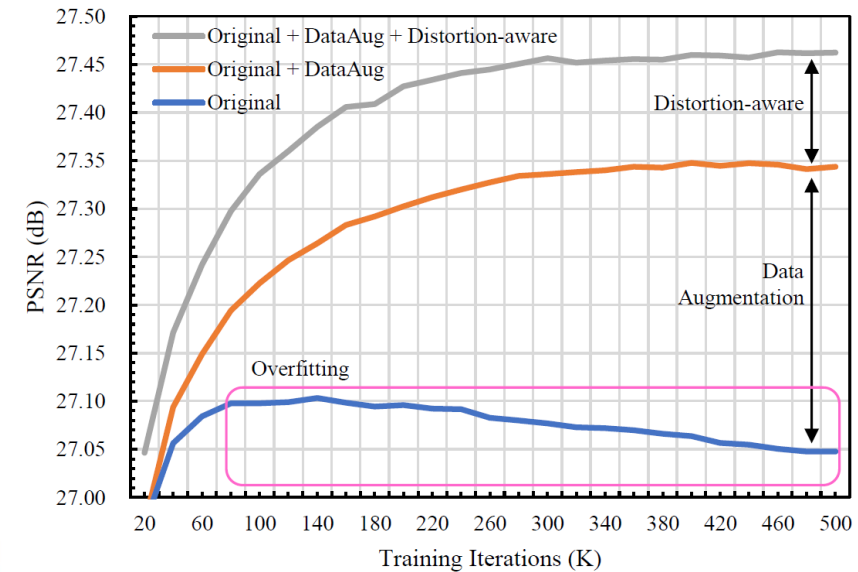


- ✓ Self-adaptively modulating distortions
- ✓ Single end-to-end model

Generating pseudo-ERP patches from plane 2D images



- Plane patches are treated as sampled Perspectives
- 146K additional training patches, 6 times of the previous dataset
- Addressing overfitting problems



Qualitative Comparisons



SUN360 ($\times 4$): 034



HR
PSNR/SSIM



Bicubic
24.38dB/0.6872



RCAN [44]
26.40dB/0.8137



SRResNet [37]
26.21dB/0.7999



EDSR [25]
26.38dB/0.8072



SwinIR [24]
26.77dB/0.8234



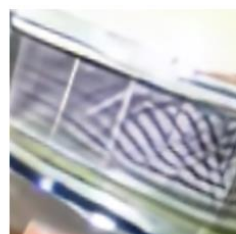
SwinIR† [24]
27.34dB/0.8462



OSRT†
27.68dB/0.8561



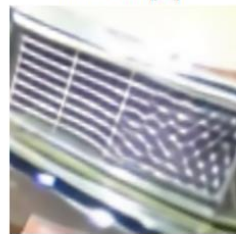
(c) ODI-SR ($\times 4$): 003
Fisheye (Vertical, Right)



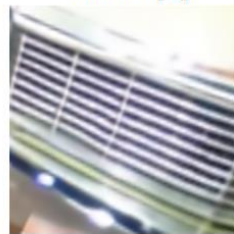
EDSR [4]



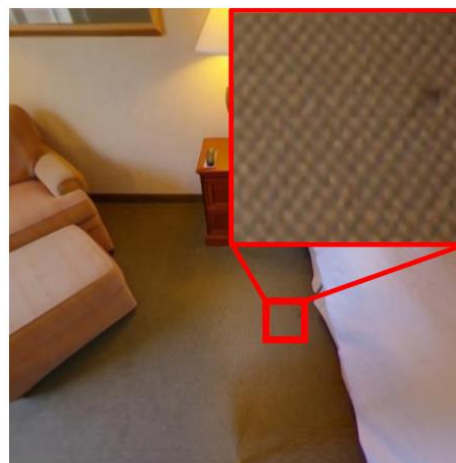
SwinIR [3]



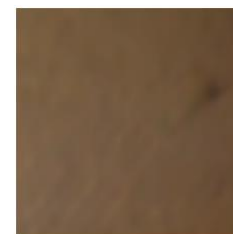
SwinIR† [3]



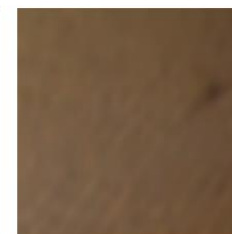
OSRT†



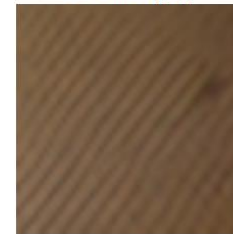
(e) SUN360 ($\times 4$): 096
Perspective (φ : -45° ; FOV: 90°)



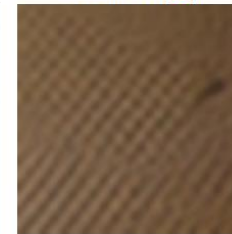
EDSR [4]



SwinIR [3]



SwinIR† [3]



OSRT†

Quantitative Comparisons

Method	Scale	ODI-SR				SUN 360 Panorama			
		PSNR	SSIM	WS-PSNR	WS-SSIM	PSNR	SSIM	WS-PSNR	WS-SSIM
Bicubic	×4	25.59	0.7118	24.95	0.6923	25.29	0.6993	24.90	0.7083
RCAN [43]		26.85	0.7621	26.15	0.7485	27.10	0.7660	26.99	0.7856
SRResNet [36]		26.91	0.7597	26.24	0.7457	27.10	0.7618	26.99	0.7812
EDSR [24]		26.97	0.7589	26.30	0.7458	27.19	0.7633	27.10	0.7827
SwinIR [23]		27.12	0.7663	26.44	0.7523	27.39	0.7707	27.30	0.7901
SwinIR [†] [23]		27.31	0.7735	26.61	0.7589	27.71	0.7804	27.64	0.7996
OSRT [†]		27.41	0.7762	26.70	0.7609	27.84	0.7835	27.77	0.8020

SR results under ERP downsampling

Scale	×8				×16			
Method	ODI-SR		SUN 360 Panorama		ODI-SR		SUN 360 Panorama	
	WS-PSNR	WS-SSIM	WS-PSNR	WS-SSIM	WS-PSNR	WS-SSIM	WS-PSNR	WS-SSIM
360-SS [30]	21.65	0.6417	21.48	0.6352	19.65	0.5431	19.62	0.5308
LAU-Net [12]	24.36	0.6801	24.02	0.6708	22.07	0.5901	21.82	0.5824
SphereSR [40]	24.37	0.6777	24.17	0.6820	22.51	0.6370	21.95	0.6342
OSRT	24.53	0.6780	24.38	0.7072	22.69	0.6261	22.13	0.6388

SR results under ERP downsampling



- Full training and testing codes
- Full script to rebuild augmented ODI dataset from DF2K
- Fisheye Downsampling as an independent downsampling kernel



Codes & Models

Arxiv: <https://arxiv.org/abs/2302.03453>

We provide detailed mathematical derivations of pixel density changes between ERP and Fisheye transforming.