NEURAL INFORMATION PROCESSING SYSTEMS

UPS: Unified Projection Sharing for Lightweight Single-Image Super-resolution and Beyond

Kun Zhou* , Xinyu Lin* , Zhonghang Liu, Xiaoguang Han✉, and Jiangbo Lu✉

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

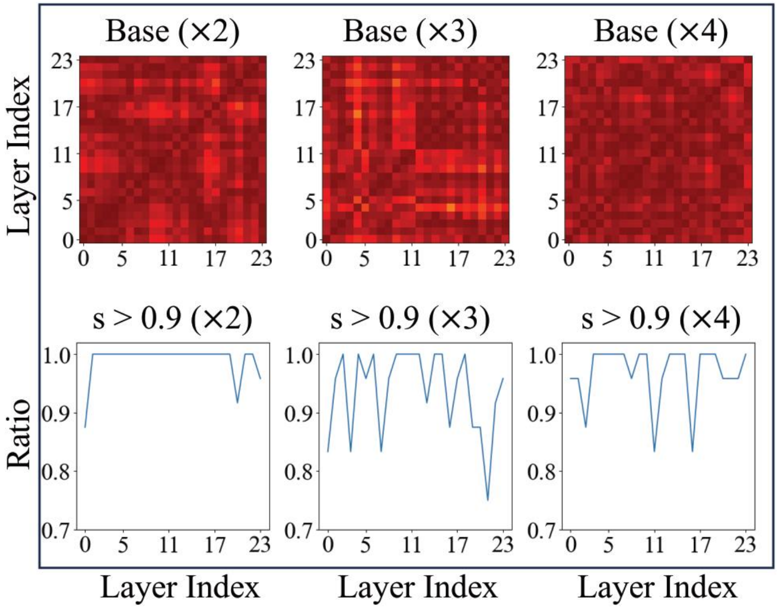
SmartMore 思谋

SMU Singapore Management University

Background

- Under practical lightweight scenarios, the complex interaction of deep image feature extraction and similarity modeling limits the performance of these methods, since they require simultaneous layer-specific optimization of both two tasks.

- We observe that the SwinIR-light (termed as Base) models exhibit **significant similarities (CKA)** in projection layers.



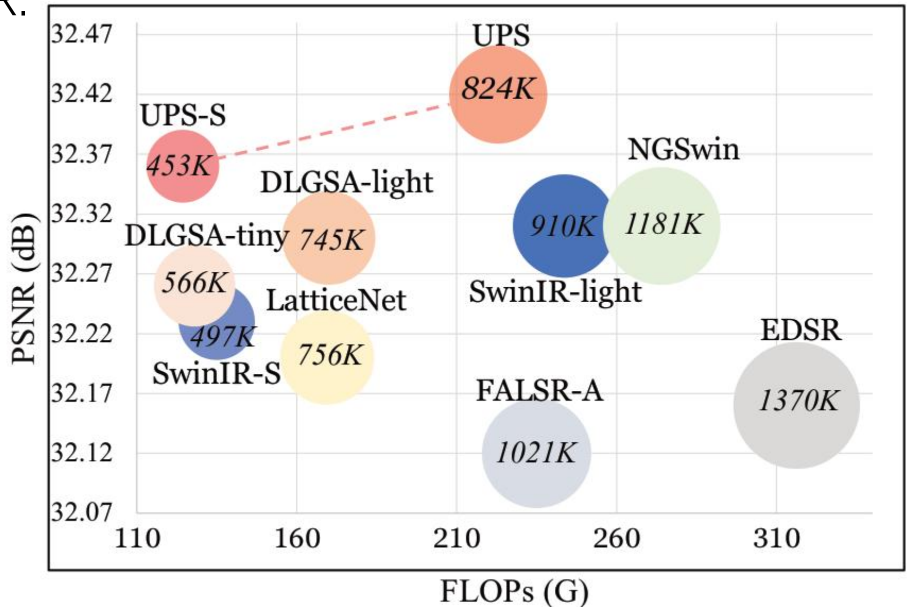
Contribution

NEURAL INFORMATION PROCESSING SYSTEMS

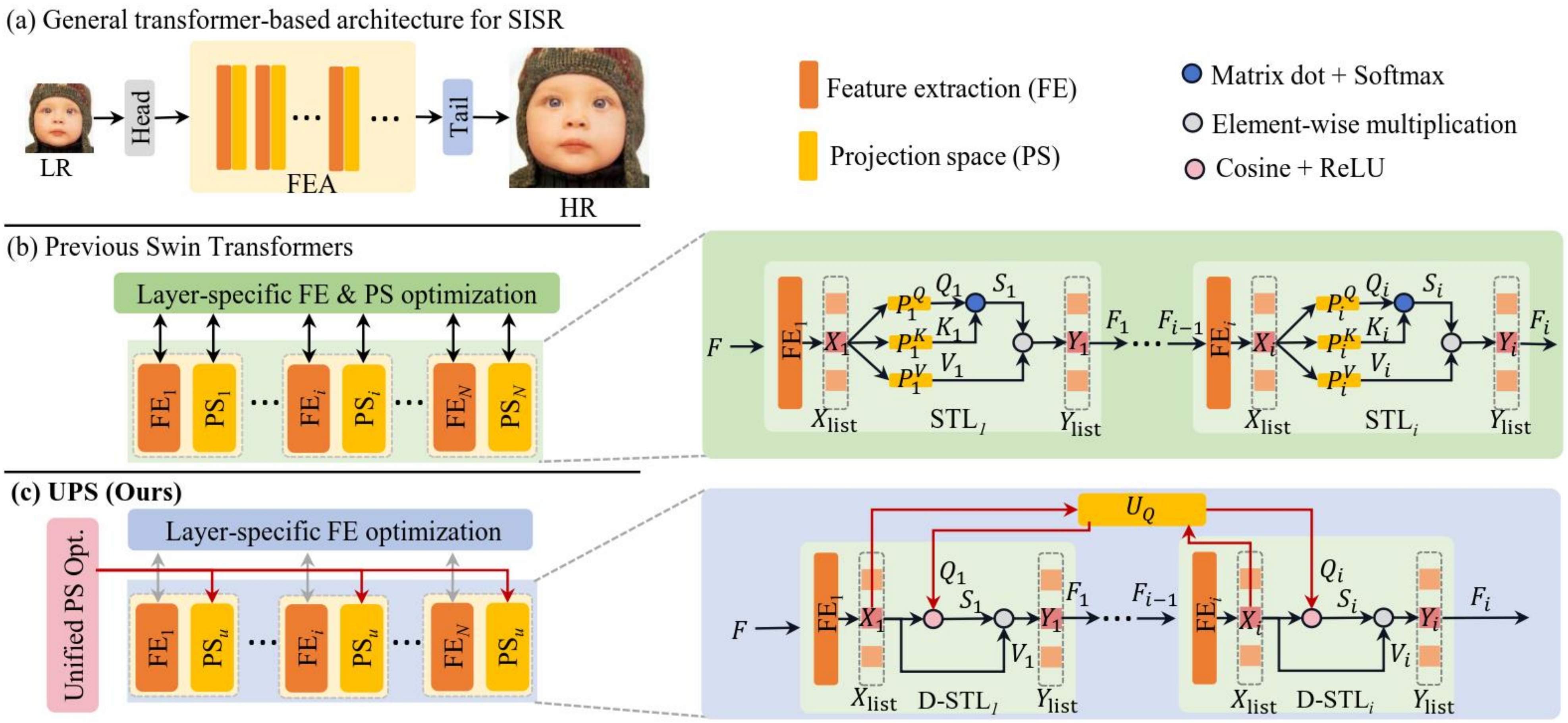
- We propose UPS, an effective decoupled SISR optimization framework, to address the challenge of simultaneous layer-specific feature extraction and similarity modeling for lightweight SISR.

- UPS simplifies the similarity optimization process by learning a layer-invariant projection space.It leads to effective aggregation and improved performance, even with reduced model capacity and less training samples

- UPS demonstrates the good generalization ability for unseen data, such as noisy image and depth map SR.



Method



- Previous Swin Transformers perform layer-specific feature extraction and projection space optimization .
- Instead, UPS adopt a global and unified projection space optimization while keeping the layer-specific FE optimization.

Method

Algorithm Illustration

Algorithm 1 Pseudo Code of the i-th STL

Require: Input F_{i-1}, window size M
Feature extraction: $\hat{F}_i = \text{Conv}(F_{i-1})$
Partitioning: $X_i^{\text{list}} = \text{Partitioning}(\hat{F}_i, M)$
Define aggregated patch list: Y_i^{list}
for X_i **in** X_i^{list} **do**
    Projection: $Q_i, K_i, V_i = X_i P_i^Q, X_i P_i^K, X_i P_i^V$
    Similarity cal.: $S_i = \text{SoftMax}\left(\frac{Q_i K_i^T}{\sqrt{d}} + B_i\right)$
    Aggregation: $Y_i = S_i V_i$
    $Y_i^{\text{list}}.append(Y_i)$
end for
return $\text{Reshape}(Y_i^{\text{list}})$

Algorithm 2 Pseudo Code of the i-th Decoupled STL

Require: F_{i-1}, M, unified projection matrix U^Q
Feature extraction: $\hat{F}_i = \text{Conv}(F_{i-1})$
Partitioning: $X_i^{\text{list}} = \text{Partitioning}(\hat{F}_i, M)$
Define aggregated patch list: Y_i^{list}
for X_i **in** X_i^{list} **do**
    Projection: $Q_i, V_i = X_i U^Q, X_i$
    Similarity cal.: $S_i = \text{ReLU}(\text{Cosine}(Q_i, Q_i^D) + B_i)$
    Aggregation: $Y_i = S_i V_i$
    $Y_i^{\text{list}}.append(Y_i)$
end for
return $\text{Reshape}(Y_i^{\text{list}})$

- We highlighted the key difference of existing Swin-transformers and our proposed UPS.

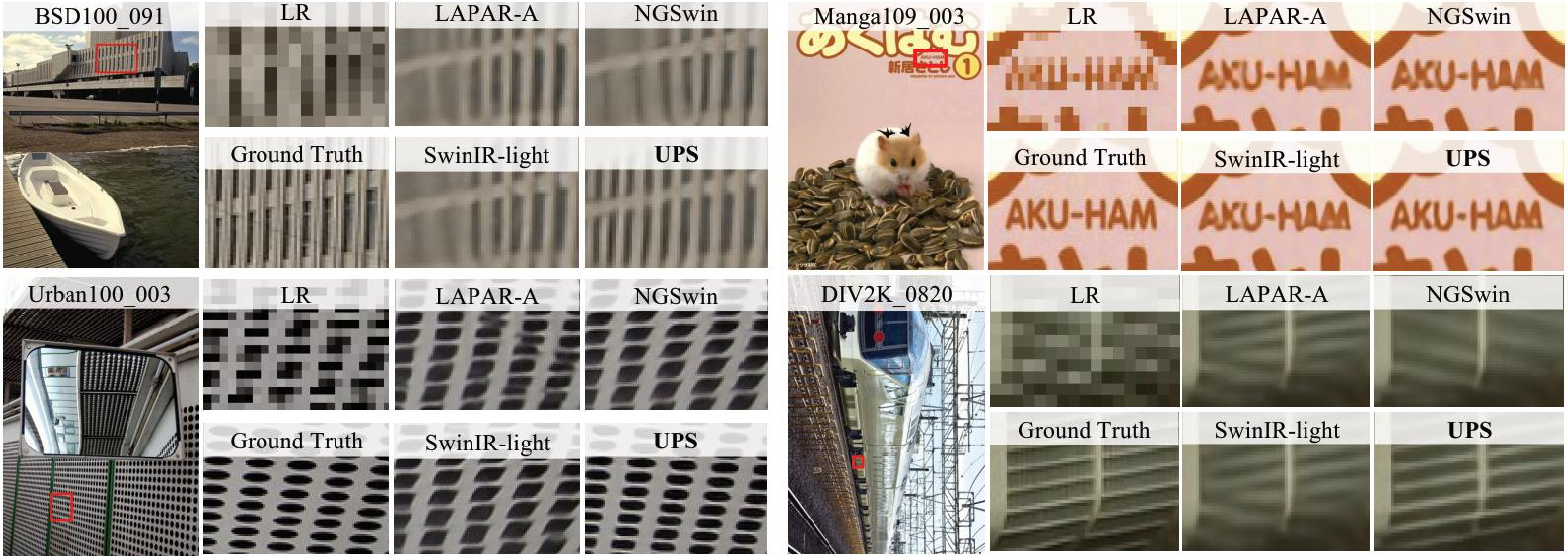
Experiments

Quantitative Comparison

Method	Scale	Parameters (K)	FLOPs (G)	Set5 PSNR / SSIM	Set14 PSNR / SSIM	BSD100 PSNR / SSIM	Urban100 PSNR / SSIM	Manga109 PSNR / SSIM
IMDN	×2	694	158.8	38.00 / 0.9605	33.63 / 0.9177	32.19 / 0.8996	32.17 / 0.9283	38.88 / 0.9774
RFDN-L	×2	626	145.8	38.08 / 0.9606	33.67 / 0.9190	32.18 / 0.8996	32.24 / 0.9290	38.95 / 0.9773
SwinIR-light	×2	910	244.4	38.14 / 0.9611	33.86 / 0.9206	32.31 / 0.9012	32.76 / 0.9340	39.12 / 0.9783
DLGSA-light	×2	745	170.0	38.20 / 0.9612	33.89 / 0.9203	32.30 / 0.9012	32.94 / 0.9355	**39.29** / 0.9780
Omni-SR	×2	772	194.5	**38.22** / 0.9613	33.98 / 0.9210	**32.36** / 0.9020	**33.05** / **0.9363**	39.28 / 0.9784
UPS	×2	824	**162.5**	**38.26** / **0.9642**	**34.16** / **0.9232**	**32.42** / **0.9031**	**33.08** / **0.9373**	**39.62** / **0.9800**
SwinIR-S	×2	**497**	**107.3**	38.06 / 0.9603	33.80 / 0.9186	32.23 / 0.9006	32.24 / 0.9301	38.76 / 0.9778
UPS-S	×2	**453**	**90.6**	38.16 / **0.9638**	**34.00** / **0.9220**	**32.36** / **0.9023**	32.79 / 0.9346	39.26 / **0.9790**
Omni-SR+	×2	772	194.5	38.29 / 0.9617	34.27 / 0.9238	32.41 / 0.9026	33.30 / 0.9386	39.53 / 0.9792
UPS+	×2	824	162.5	38.31 / 0.9643	34.37 / 0.9247	32.43 /0.9032	33.34 / 0.9388	39.80 / 0.9802
IMDN	×3	703	71.5	34.36 / 0.9270	30.32 / 0.8417	29.09 / 0.8046	28.17 / 0.8519	33.61 / 0.9445
RFDN-L	×3	633	65.6	34.47 / 0.9280	30.35 / 0.8421	29.11 / 0.8053	28.32 / 0.8547	33.78 / 0.9458
SwinIR-light	×3	918	110.8	34.62 / 0.9289	30.54 / 0.8463	29.20 / 0.8082	28.66 / 0.8624	33.98 / 0.9478
DLGSA-light	×3	752	75.4	**34.70** / 0.9295	**30.58** / **0.8465**	**29.24** / 0.8089	28.83 / 0.8653	34.16 / 0.9483
Omni-SR	×3	780	88.4	**34.70** / 0.9294	30.57 / 0.8469	29.28 / **0.8094**	**28.84** / **0.8656**	**34.22** / **0.9487**
UPS	×3	832	**72.4**	**34.66** / **0.9322**	**30.72** / **0.8489**	**29.31** / **0.8114**	**28.98** / **0.8685**	**34.53** / **0.9505**
SwinIR-S	×3	**503**	**47.9**	34.38 / 0.9281	30.46 / 0.8448	29.15 / 0.8073	28.37 / 0.8572	33.77 / 0.9464
UPS-S	×3	**459**	**40.4**	34.53 / **0.9312**	30.55 / 0.8463	**29.24** / 0.8093	28.60 / 0.8614	34.12 / 0.9484
Omni-SR+	×3	780	88.4	34.77 / 0.9304	30.70 / 0.8489	29.33 / 0.8111	29.12 / 0.8712	34.64 / 0.9507
UPS+	×3	832	72.4	34.78 / 0.9325	30.78 / 0.8492	29.36 / 0.8122	29.28 / 0.8728	34.84 / 0.9517
IMDN	×4	715	40.9	32.21 / 0.8948	28.58 / 0.7811	27.56 / 0.7353	26.04 / 0.7838	30.45 / 0.9075
RFDN-L	×4	643	37.4	32.28 / 0.8957	28.61 / 0.7818	27.58 / 0.7363	26.20 / 0.7883	30.61 / 0.9096
SwinIR-light	×4	930	63.6	32.44 / 0.8976	28.77 / 0.7858	27.69 / 0.7406	26.47 / 0.7980	30.92 / 0.9151
DLGSA-light	×4	761	42.5	**32.54** / 0.8993	**28.84** / **0.7871**	**27.73** / **0.7415**	**26.66** / **0.8033**	**31.13** / 0.9161
Omni-SR	×4	792	50.9	32.49 / 0.8988	28.78 / 0.7859	27.71 / **0.7415**	26.64 / 0.8018	31.02 / 0.9151
UPS	×4	843	**41.3**	**32.50** / **0.9024**	**28.90** / **0.7892**	**27.79** / **0.7435**	**26.83** / **0.8073**	**31.39** / **0.9194**
SwinIR-S	×4	**512**	**27.3**	32.14 / 0.8955	28.67 / 0.7832	27.63 / 0.7382	26.22 / 0.7906	30.68 / 0.9111
UPS-S	×4	**468**	**23.0**	32.41 / **0.9008**	28.80 / 0.7863	**27.73** / 0.7414	26.58 / 0.7995	**31.13** / **0.9163**
Omni-SR+	×4	792	50.9	32.57 / 0.8993	28.95 / 0.7898	27.81 / 0.7439	26.95 / 0.8105	31.50 / 0.9192
UPS+	×4	843	41.3	32.60 / 0.9029	28.97 / 0.7896	27.83 / 0.7446	27.10 / 0.8136	31.79 / 0.9223

Experiments

Qualitative Comparison

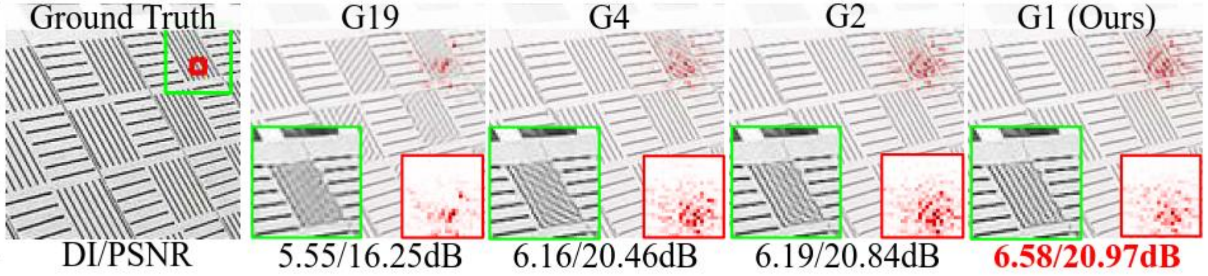


Experiments

Ablation Studies

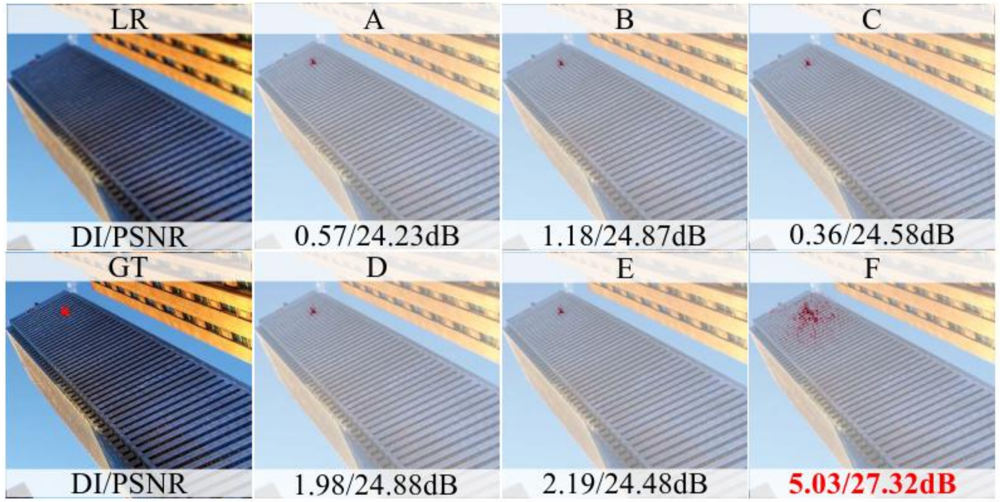
- Analysis of several UPS-S models with different projection groups

Proj. Group	G19	G4	G2	G1 (Ours)
PSNR (dB)	33.63	33.91	33.97	**34.00**
SSIM	0.9186	0.9205	0.9216	**0.9220**



- Impact of different similarity calculation methods

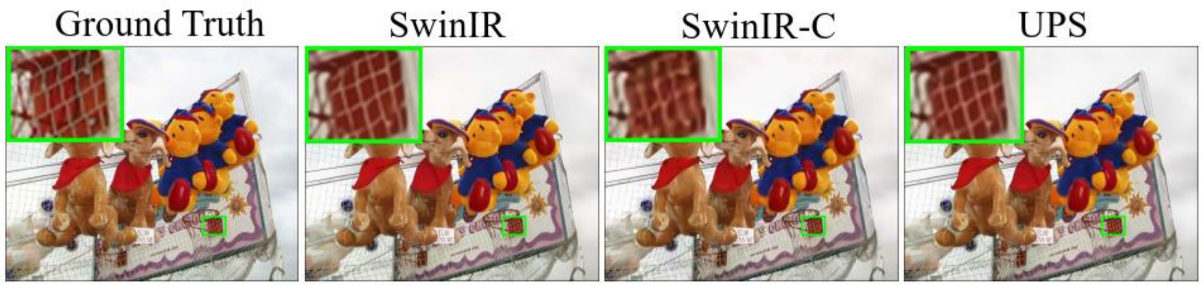
	Matrix dot	Cosine	SoftMax	ReLU	PSNR/SSIM
A	✓				33.60/0.9192
B	✓		✓		33.84/0.9202
C	✓			✓	33.41/0.9169
D		✓			33.61/0.9208
E		✓	✓		33.73/0.9194
F		✓		✓	**34.00/0.9220**



Extension

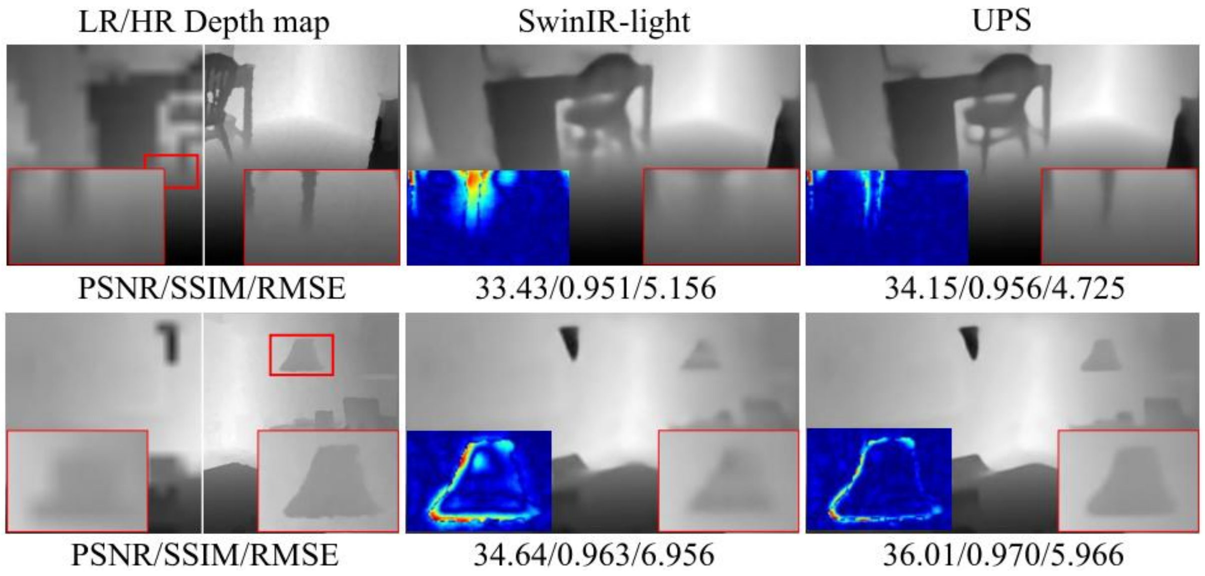
- Extension on other image restoration problems (Denoising)

Tasks	Metrics	SwinIR	SwinIR-C	UPS
Deblocking	PSNR	29.86	29.63	**29.98**
$q = 40$	Param.	11.50M	3.89M	**3.49M**
Denoising	PSNR	**28.56**	28.20	28.37
$\sigma = 50$	Param.	11.50M	0.959M	**0.873M**



- Generalization comparison between baseline model and UPS (Depth Map SR)

Settings	Metrics	SwinIR-light	UPS
$\times 4$	PSNR	47.25	**47.79**
	SSIM	0.994	**0.995**
	RMSE	2.339	**2.198**
$\times 16$	PSNR	37.25	**37.98**
	SSIM	0.969	**0.972**
	RMSE	7.832	**7.236**



Thank You

UPS: Unified Projection Sharing for Lightweight Single-Image Super-resolution and Beyond

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

SmartMore 思谋

SMU SINGAPORE MANAGEMENT UNIVERSITY