

# UCAN: Unified Convolutional Attention Network for Expansive Receptive Fields in Lightweight Super-Resolution

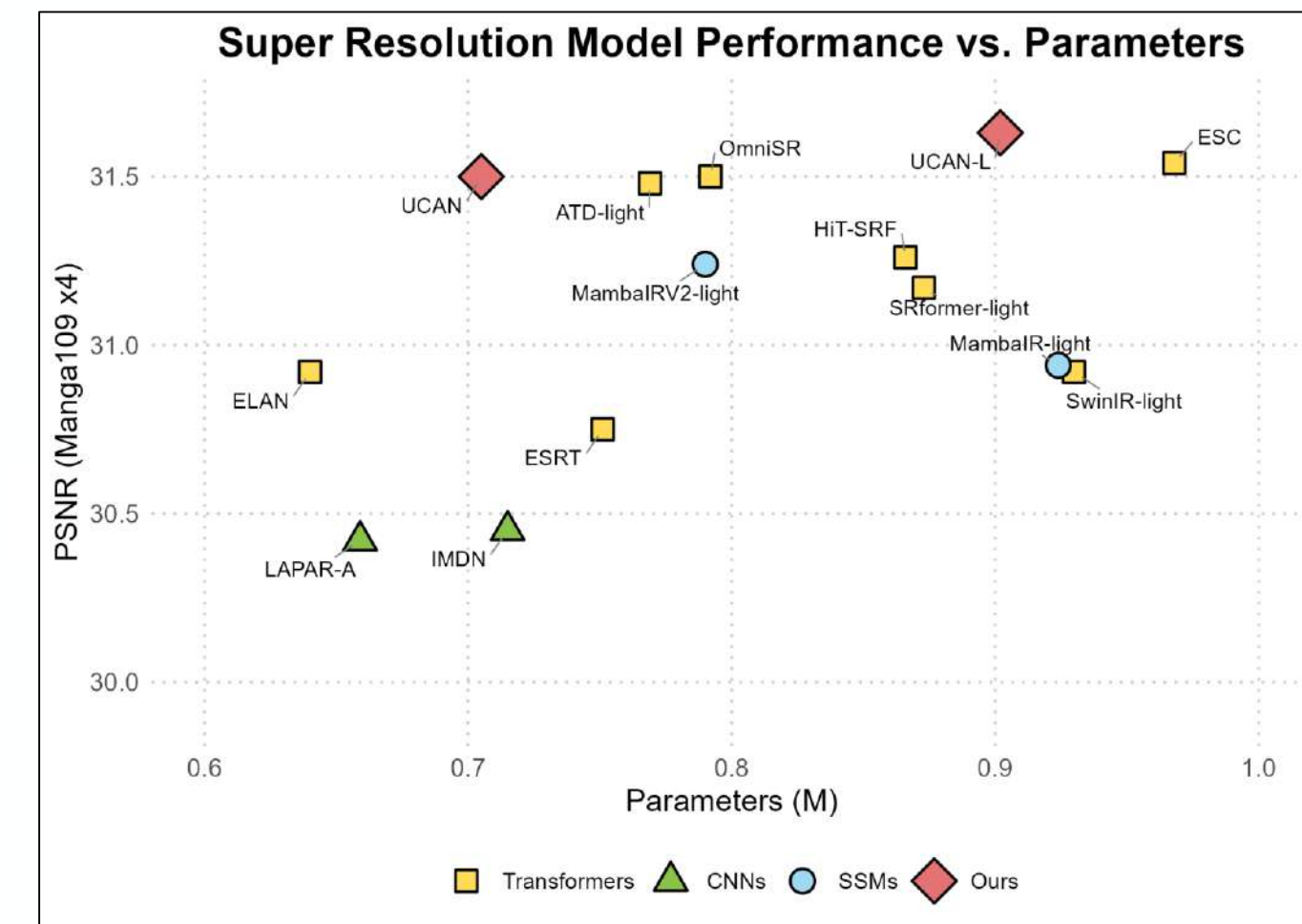
Cao Thien Tan, Phan Thi Thu Trang, Duc Nghiem Do, Ho Ngoc Anh, Hanyang Zhuang\*, Duc Dung Nguyen\*



Paper and Code



## Background & Motivation



**Background:** Current SR research seeks to **expand receptive fields** while maintaining efficiency.

**Motivation:**

**CNNs** — restricted to local receptive fields; miss long-range self-similar patterns.

**Transformers** — model long-range dependencies but suffer quadratic complexity in memory and computation, limiting scalability to high-resolution inputs.

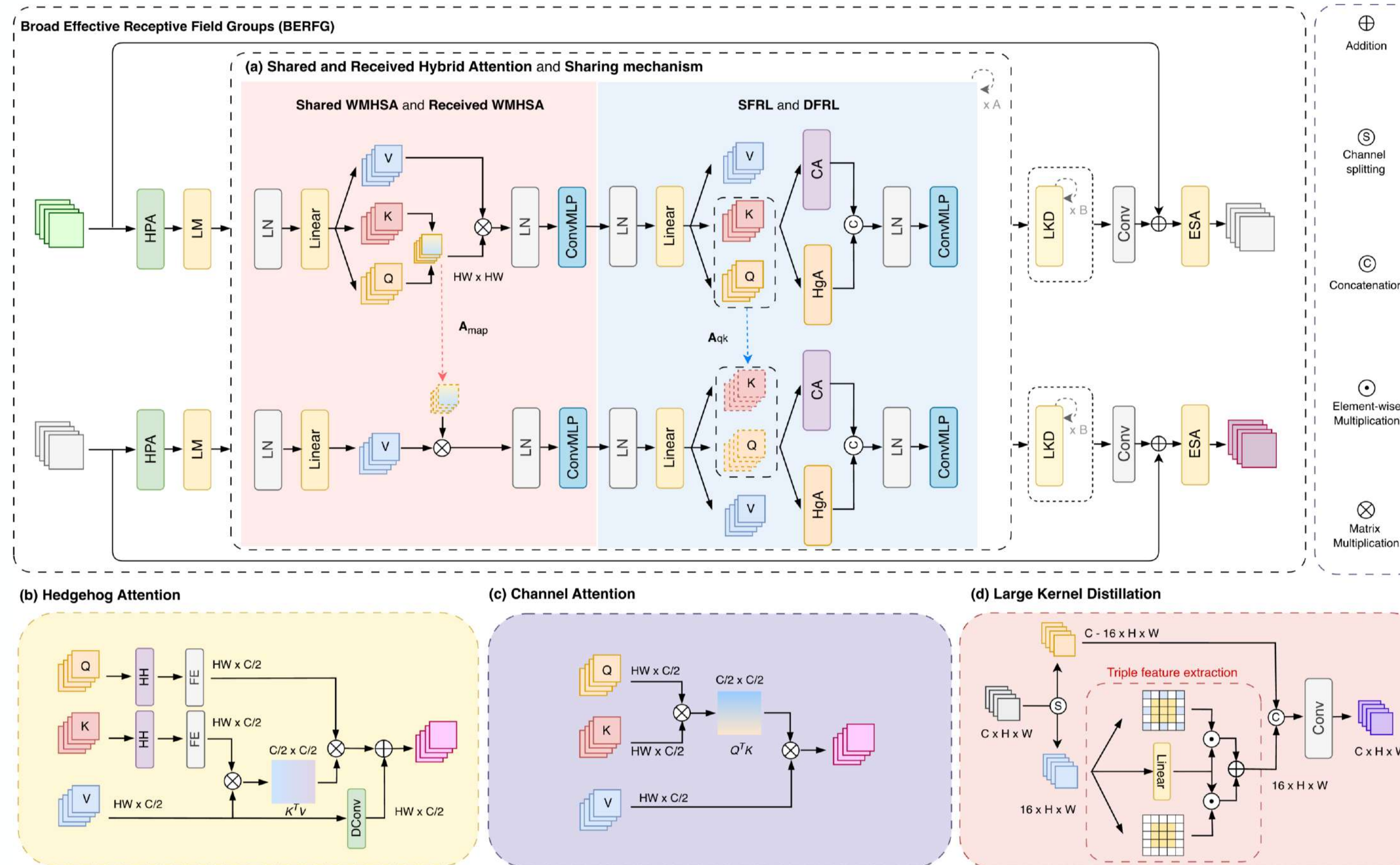
**Linear Attention** — reduces computational cost but suffers from rank collapse, degrading representation quality.

⇒ Our model combines all of this to expand the receptive field **efficiently** and **without sacrificing feature diversity**.

## Our Contribution

- ✓ Hedgehog Attention boosts feature rank in linear attention for richer, more expressive representations.
- ✓ Semi-sharing distillation balances shared and task-specific weights, achieving strong SR performance with fewer parameters and lower compute cost.
- ✓ UCAN unifies Flash Attention, Hedgehog Attention, and multi-kernel convolution to capture both local and global dependencies efficiently.

## UCAN



### High performance attention

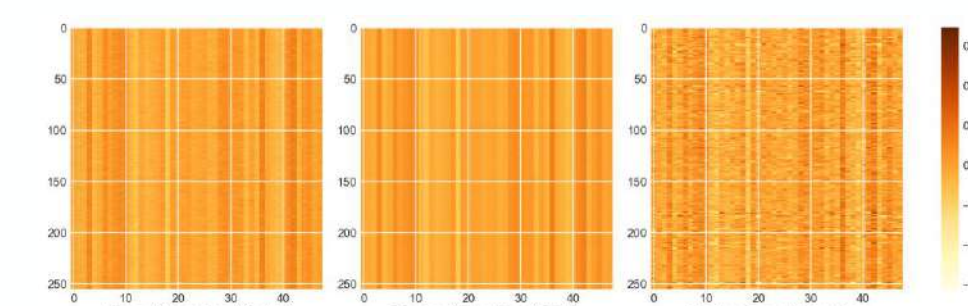
**Flash Attention** enables exact large-window attention ( $32 \times 32$ ) with lower memory and faster execution, replacing standard  $O(N^2)$  window attention

### Semi-sharing Mechanism (BERFG)

SHA computes full attention and shares the attention maps  $A_{qk}, A_{map}$  with RHA, eliminating redundant recomputation. This semi-sharing strategy maintains strong feature representation while substantially reducing computational cost.

### Hedgehog Attention (HgA)

- Fixes rank collapse in linear attention using the Hedgehog Feature Map.
- Symmetric exponential pairs preserve both positive and negative projection directions.
- Achieves feature rank up to 46, exceeding compared methods



### Large Kernel Distillation (LKD)

- Splits channels: heavy computation only on fine-grained subset  $C_{fg} = \max(C/4, 16)$
- Three parallel branches: Channel Branch + Local Context Branch + Hierarchical Large Kernel Branch.

### Architecture Overview & Complexity

Core encoder: HPA (local) → SHA/RHA (global) → LKD (spatial refinement).

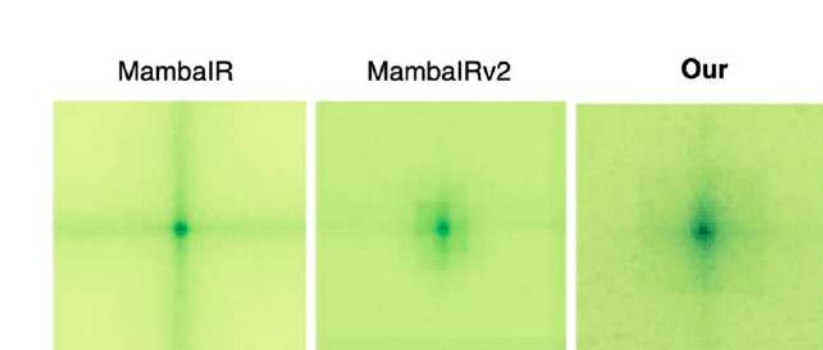
Linear complexity with  $O(N)$  and Flash Window attention.

## Result and Analysis

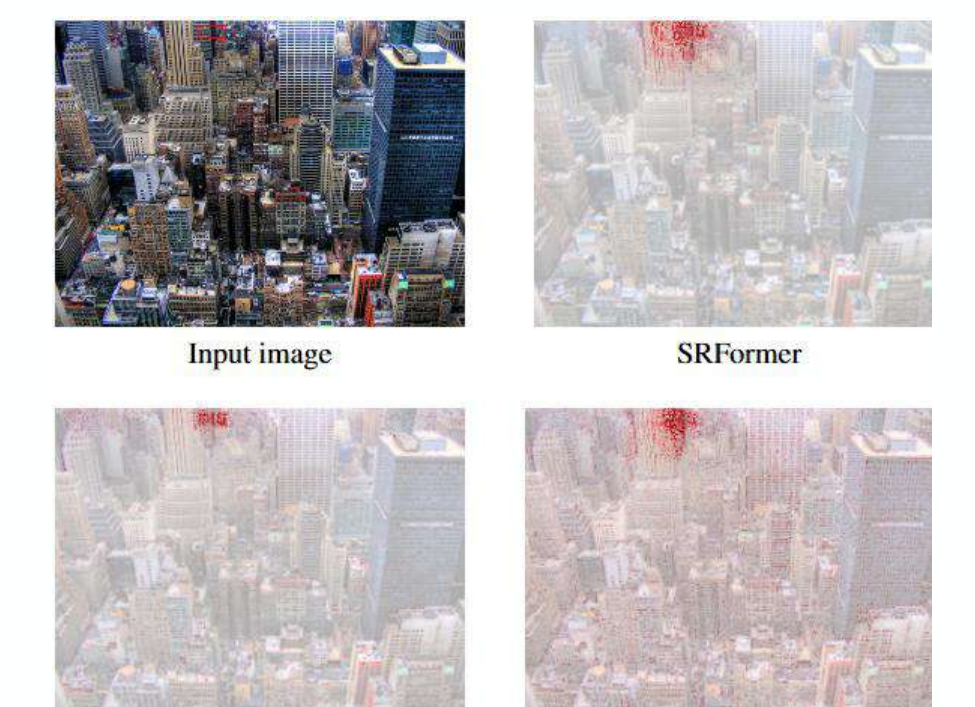
Table 1. Quantitative comparison on *lightweight image super-resolution* with state-of-the-art methods. The best and second-best results are shown in **bold** and underlined, respectively.

Method	scale	#param	MACs	Set5		Set14		BSDS100		Urban100		Manga109	
				PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MambalR-light [10]	2x	905K	334.2G	38.13	0.9610	33.95	0.9208	32.31	0.9013	32.85	0.9349	39.20	0.9782
OmniSR [35]	2x	772K	194.5G	38.22	0.9613	33.98	0.9210	32.36	0.9020	33.05	0.9363	39.28	0.9784
SRFormer-light [46]	2x	853K	236.3G	38.23	0.9613	33.94	0.9209	32.36	0.9019	32.91	0.9353	39.28	0.9785
ATD-light [40]	2x	753K	380.0G	38.29	0.9616	34.10	0.9217	32.39	0.9023	33.27	0.9375	39.52	0.9789
HIT-SRF [43]	2x	847K	226.5G	38.26	0.9615	34.01	0.9214	32.37	0.9023	33.13	0.9372	39.47	0.9787
ASID-D8 [27]	2x	732K	190.5G <sup>†</sup>	38.32	0.9618	34.24	0.9232	32.40	0.9028	33.35	0.9387	-	-
MambalRV2-light [9]	2x	774K	286.3G	38.26	0.9615	34.09	0.9221	32.36	0.9019	33.26	0.9378	39.35	0.9785
RDN [45]	2x	22123K	5096.2G	38.24	0.9614	34.01	0.9212	32.34	0.9017	32.89	0.9353	39.18	0.9780
RCAN [44]	2x	15445K	3529.7G	38.27	0.9614	34.12	0.9216	32.41	0.9027	33.34	0.9384	39.44	0.9786
ESC [27]	2x	947K	592.0G	38.35	0.9619	34.11	0.9223	32.41	0.9027	33.46	0.9395	39.54	0.9790
UCAN (Our)	2x	689K	146.3G	38.34	0.9618	34.27	0.9242	32.39	0.9025	33.22	0.9379	39.54	0.9790
UCAN-L (Our)	2x	886K	182.4G	38.37	0.9619	34.19	0.9224	32.44	0.9031	33.39	0.9393	39.66	0.9789
MambalR-light [10]	3x	913K	148.5G	34.63	0.9288	30.54	0.8459	29.23	0.8084	28.70	0.8631	34.12	0.9479
OmniSR [35]	3x	780K	88.4G	34.70	0.9294	30.57	0.8469	29.28	0.8099	28.84	0.8656	34.22	0.9487
SRFormer-light [46]	3x	861K	105.4G	34.67	0.9296	30.57	0.8469	29.26	0.8099	28.81	0.8655	34.19	0.9489
ATD-light [40]	3x	760K	168.0G	34.74	0.9300	30.68	0.8485	29.32	0.8109	29.17	0.8709	34.60	0.9506
HIT-SRF [43]	3x	855K	101.6G	34.75	0.9300	30.61	0.8475	29.29	0.8106	28.99	0.8687	34.53	0.9502
ASID-D8 [27]	3x	739K	86.4G	34.84	0.9307	30.66	0.8491	29.32	0.8119	29.08	0.8706	-	-
MambalRV2-light [9]	3x	781K	126.7G	34.71	0.9298	30.68	0.8483	29.26	0.8098	29.01	0.8689	34.41	0.9497
RDN [45]	3x	22308K	2281.2G	34.71	0.9296	30.57	0.8468	29.26	0.8093	28.80	0.8653	34.13	0.9484
RCAN [44]	3x	15629K	1586.1G	34.74	0.9299	30.65	0.8482	29.32	0.8111	29.09	0.8702	34.44	0.9499
ESC [27]	3x	955K	267.6G	34.84	0.9308	30.74	0.8493	29.34	0.8118	29.28	0.8739	34.66	0.9512
UCAN (Our)	3x	696K	64.6G	34.83	0.9308	30.72	0.8493	29.32	0.8121	29.15	0.8712	34.62	0.9508
UCAN-L (Our)	3x	893K	81.3G	34.81	0.9311	30.75	0.8500	29.34	0.8127	29.29	0.8738	34.79	0.9516
MambalR-light [10]	4x	924K	84.6G	32.42	0.8977	28.74	0.7847	27.68	0.7400	26.52	0.7983	30.94	0.9135
OmniSR [35]	4x	792K	50.9G	32.49	0.8988	28.78	0.7859	27.71	0.7415	26.64	0.8018	31.02	0.9151
SRFormer-light [46]	4x	873K	62.8G	32.51	0.8988	28.82	0.7872	27.73	0.7422	26.67	0.8032	31.17	0.9165
ATD-light [40]	4x	769K	100.1G	32.63	0.8998	28.89	0.7886	27.79	0.7440	26.97	0.8107	31.48	0.9198
HIT-SRF [43]	4x	866K	58.0G	32.55	0.8999	28.87	0.7880	27.75	0.7432	26.80	0.8069	31.26	0.9171
ASID-D8 [27]	4x	748K	49.6G	32.57	0.8990	28.89	0.7898	27.78	0.7449	26.89	0.8096	-	-
MambalRV2-light [9]	4x	790K	75.6G	32.51	0.8992	28.84	0.7878	27.75	0.7426	26.82	0.8079	31.24	0.9182
RDN [45]	4x	22271K	1309.2G	32.47	0.8990	28.81	0.7871	27.72	0.7419	26.61	0.8028	31.00	0.9151
RCAN [44]	4x	15592K	917.6G	32.63	0.9002	28.87	0.7889	27.77	0.7436	26.82	0.8087	31.22	0.9173
ESC [27]	4x	968K	149.2G	32.68	0.9011	28.93	0.7902	27.80	0.7447	27.07	0.8144	31.54	0.9207
UCAN (Our)	4x	705K	38.1G	32.65	0.9010	28.95	0.7899	27.79	0.7454	26.89	0.8097	31.50	0.9200
UCAN-L (Our)	4x	902K	48.4G	32.68	0.9015	28.99	0.7917	27.80	0.7459	27.06	0.8134	31.63	0.9212

### Analysis

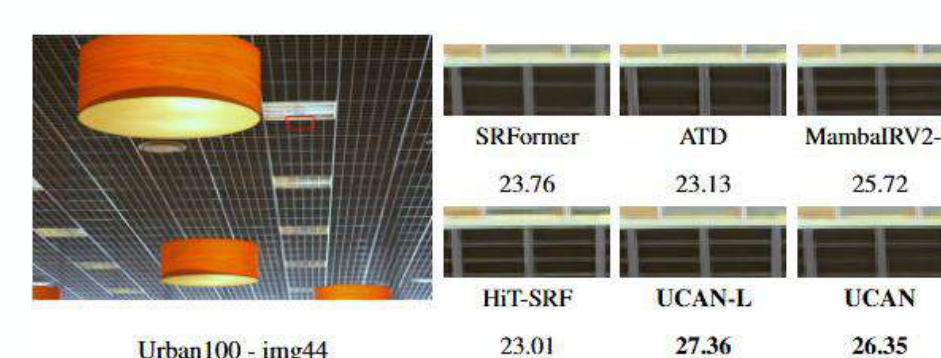
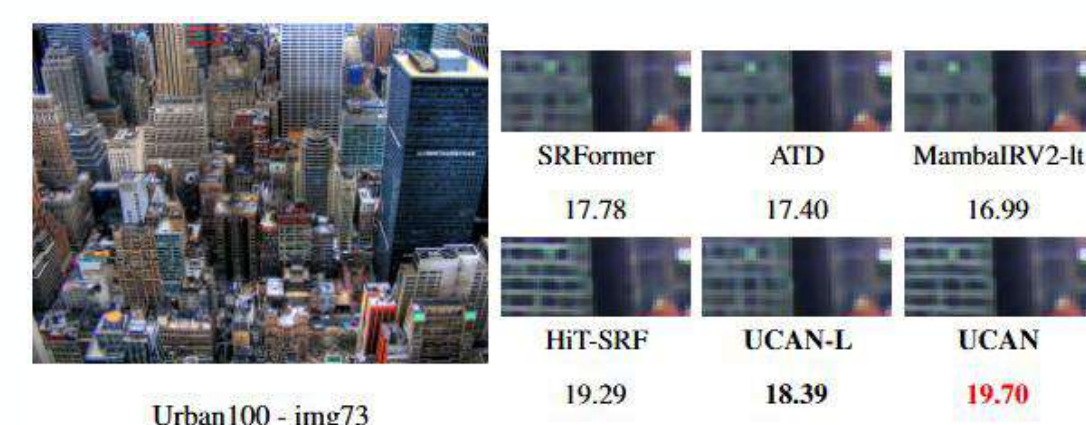


Efficient Receptive Field Comparison



Local Attribution Maps

## Visualization



Manga109 - Gakuen Noise