Deep Arbitrary-Scale Image Super-Resolution via Scale-Equivariance Pursuit

Xiaohang Wang*, Xuanhong Chen*, BingbingNi**, Hang Wang, Zhengyan Tong, Yutian Liu

{xygz2014010003, chen19910528, nibingbing, 418004}@sjtu.edu.cn, francis970625@gmail.com

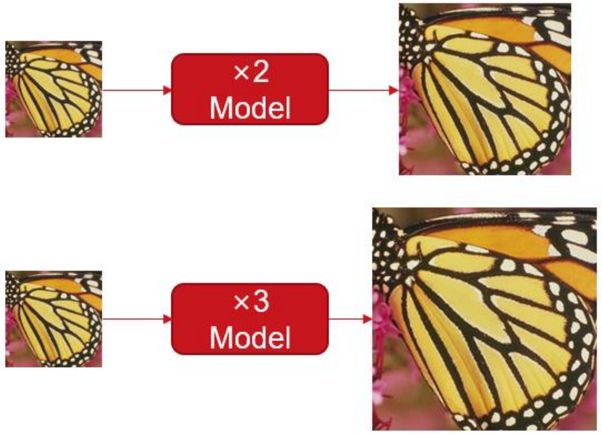
https://github.com/neuralchen/EQSR

TUE-AM-170

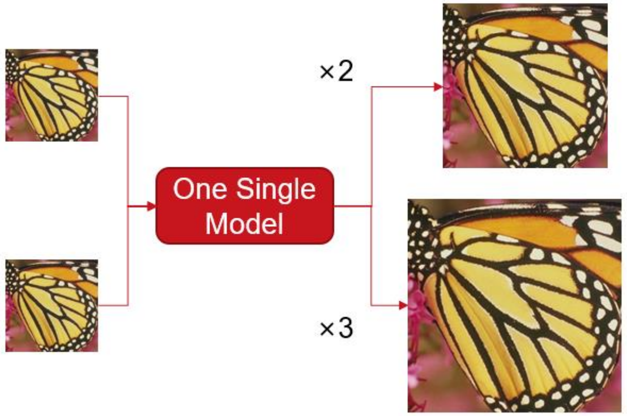
*These authors contributed equally.
**Corresponding author: Bingbing Ni.

饮 水 思 源 • 爱 国 荣 校

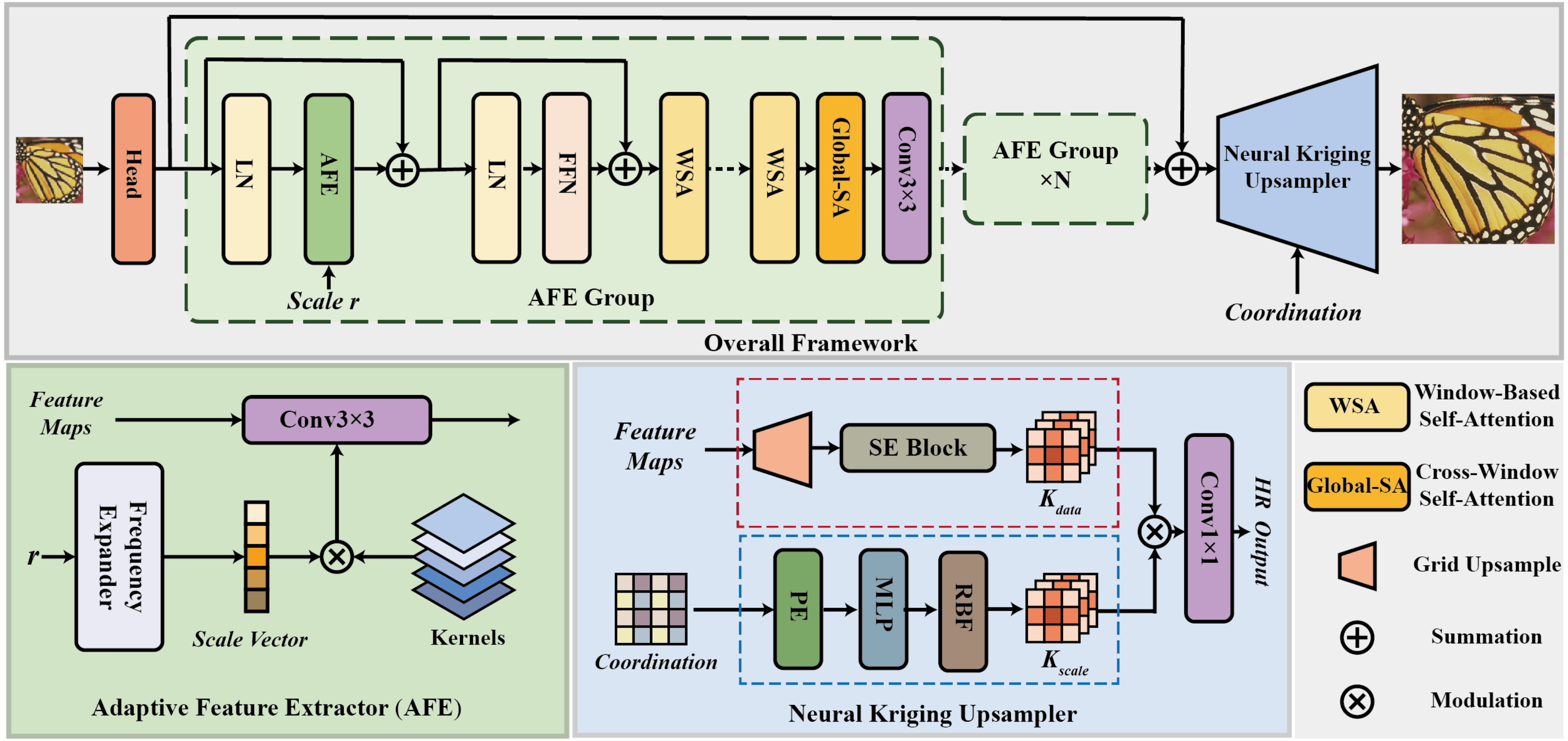
Preview



Fixed-Scale SR



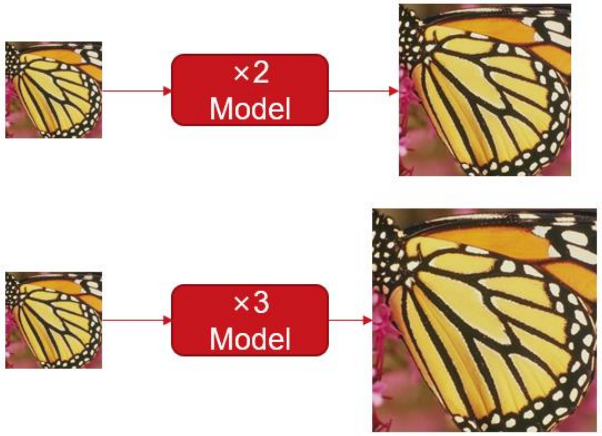
Arbitrary-Scale SR



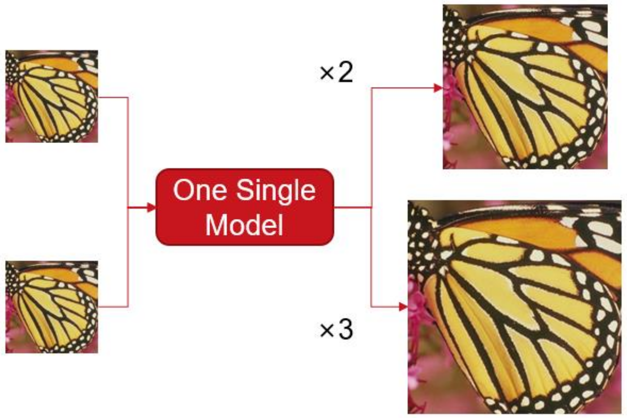
Key idea:

1. **Backbone**: inject scale information into the feature extraction process explicitly.
2. **Upsampler**: neuralize Kriging interpolation and construct an up-sampling operator that possesses a certain level of scale equivariance.
3. **Training**: novel data processing approach for ASISR, which enables more straightforward large-scale pretraining and offers greater flexibility in selecting network layers.

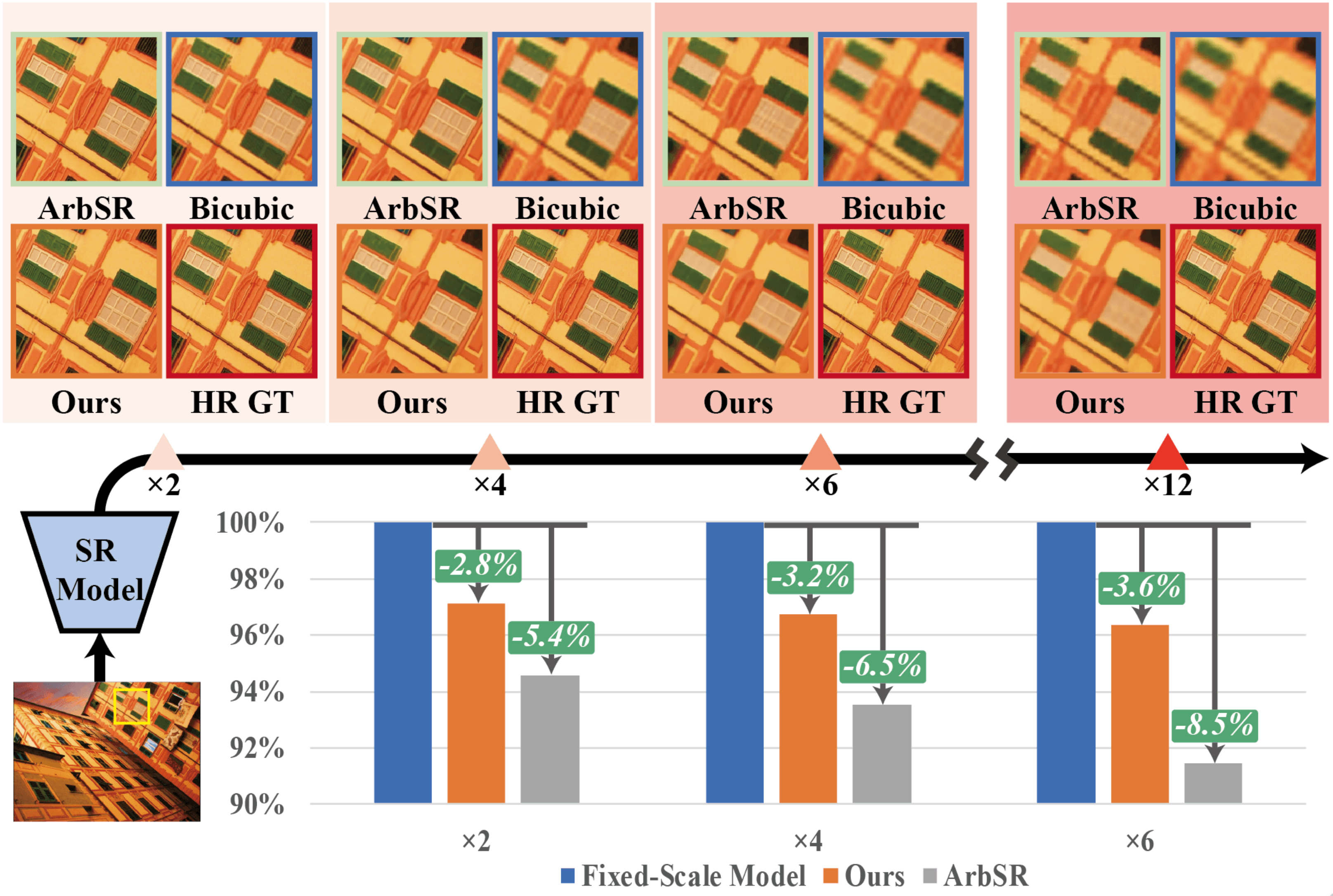
Background & Motivation



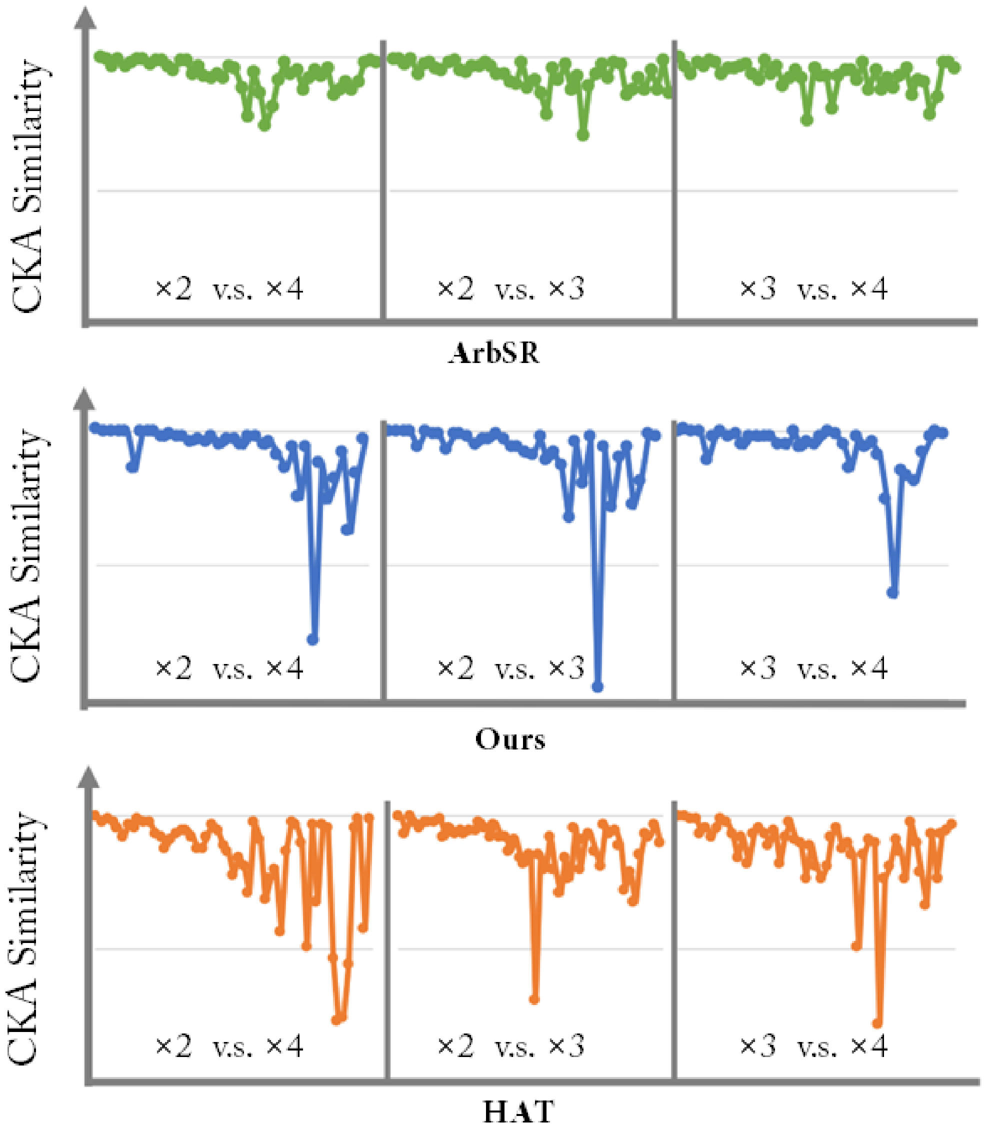
Fixed-Scale SR



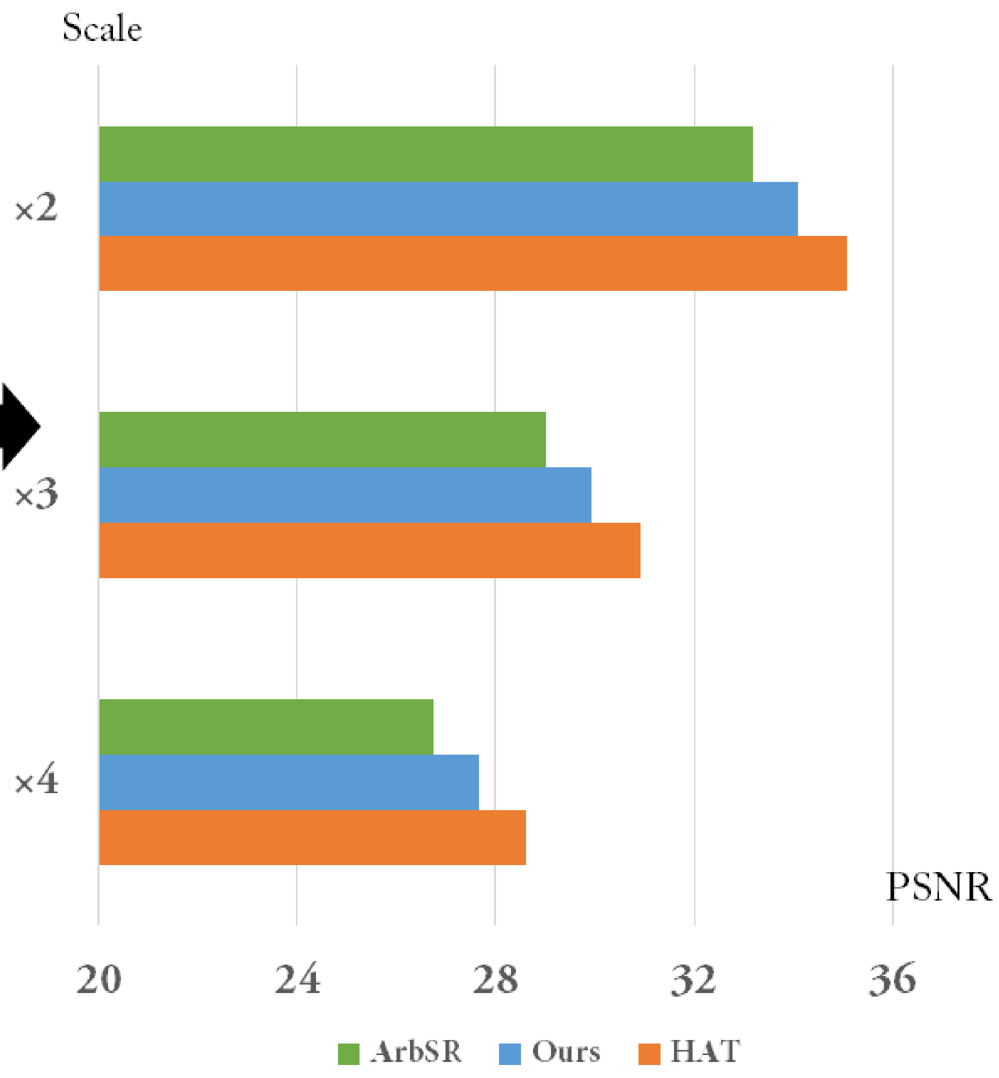
Arbitrary-Scale SR



Background & Motivation



(a)



(b)

Methodology

1. How to achieve arbitrary-scale upsampling?
2. How to incorporate scale information during the feature extraction process?
3. How to enhance the equivariance of the upsampling operator?
4. How to train the model (e.g., different scales in one batch)?
5. How to measure equivariance?

1. Implicit Based Upsampling

We follow LIIF [4] and LTE [18], constructing an implicit field for the input image and then reconstructing high-resolution results by querying the RGB values at target coordinates.

$$\boldsymbol{z_0} = \boldsymbol{I^{SR}}(\boldsymbol{x_0}) = \boldsymbol{\Phi}(\boldsymbol{I^{LR}}, \boldsymbol{x_0}; \boldsymbol{\theta})$$

RGB of target point

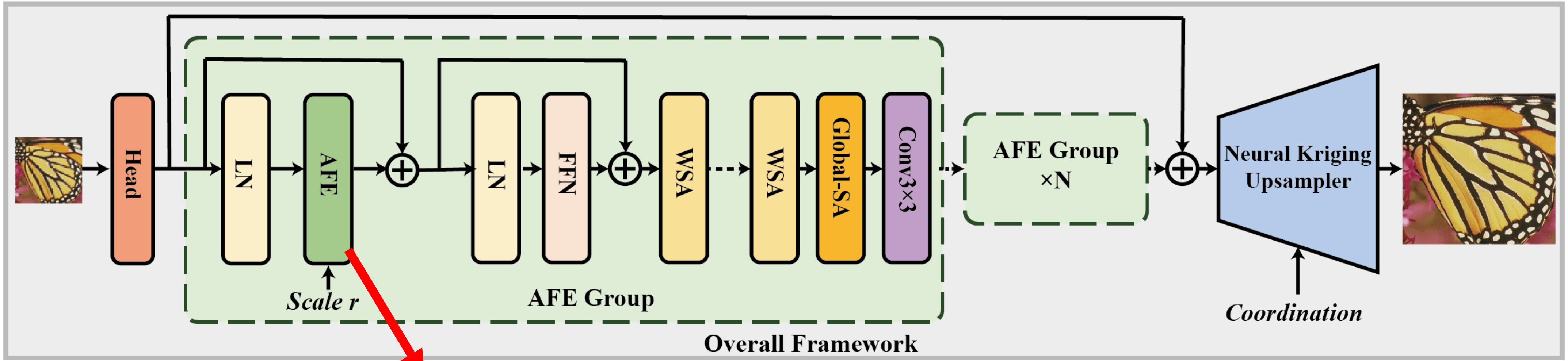
Model Weights

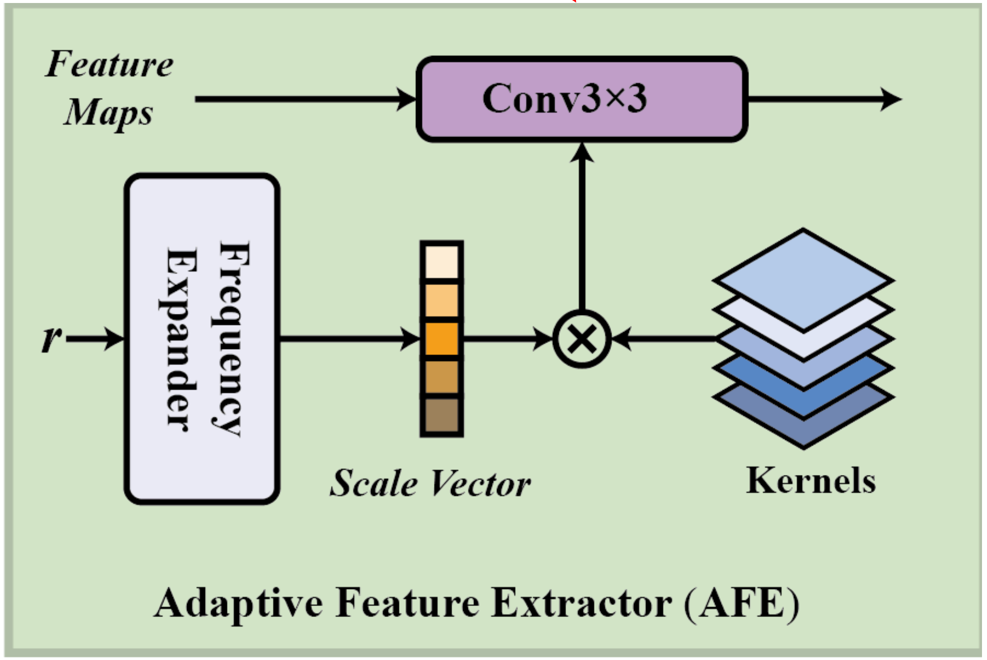
Coordinate of target point

However, our implicit field **differs from** previous works. See Problem 4 for details.

[4] Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function.
[18] Jaewon Lee and Kyong Hwan Jin. Local texture estimator for implicit representation function.

2. Adaptive Feature Extractor





$$X_{out} = \mathcal{F}\left(\begin{bmatrix} sin(2^i\pi r/10) \\ cos(2^i\pi r/10) \\ r \end{bmatrix}\right) \otimes \mathcal{W} * X_{in} + \mathcal{B},$$

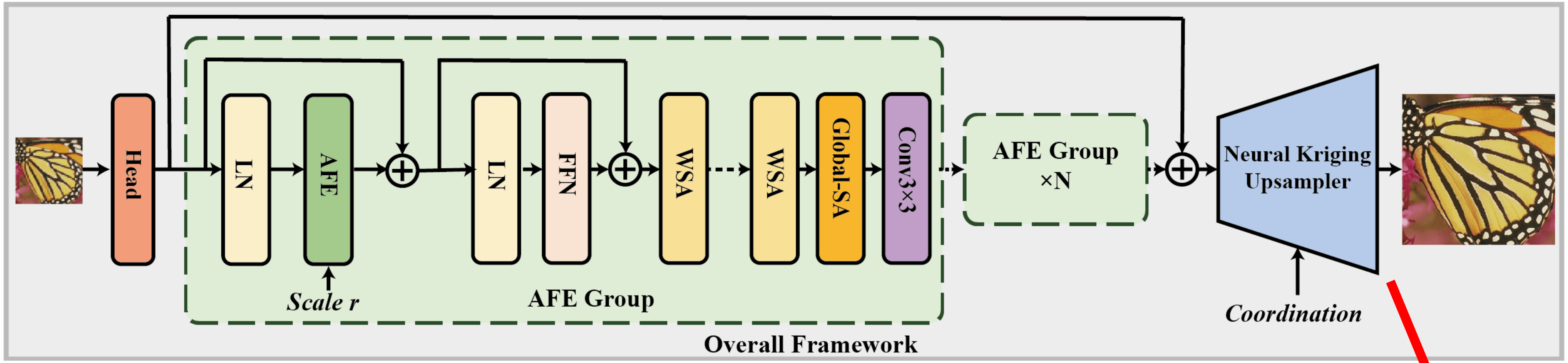
Linear layer

Sine-cosine encoding

Kernel basis

Input Features

3. Neural Kriging Upsampler

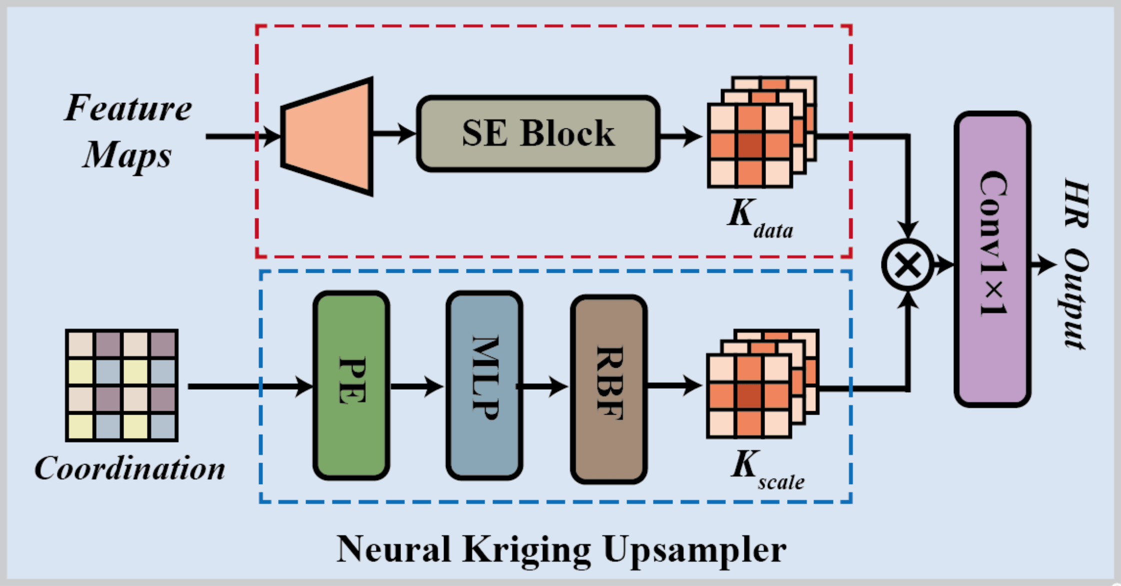


$$\hat{z}_0 = \sum_i^N \lambda_i z_i \qquad \langle z(x_0), z(x_j) \rangle = \sum_{i=1}^{N} \lambda \langle z(x_i), z(x_j) \rangle, \forall j = 0, 1, \ldots, n,$$

$$\underbrace{\begin{bmatrix} c(0,1) \\ c(0,2) \\ \vdots \\ c(0,N) \end{bmatrix}}_{K_{scale}} = \underbrace{\begin{bmatrix} c(1,1) & c(1,2) & \cdots & c(1,N) \\ \vdots & \ddots & & \vdots \\ c(N,1) & \cdots & \ddots & c(N,N) \end{bmatrix}}_{K_{data}} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix}$$

$$\text{where } c(i,j) = \langle z(x_i), z(x_j) \rangle$$

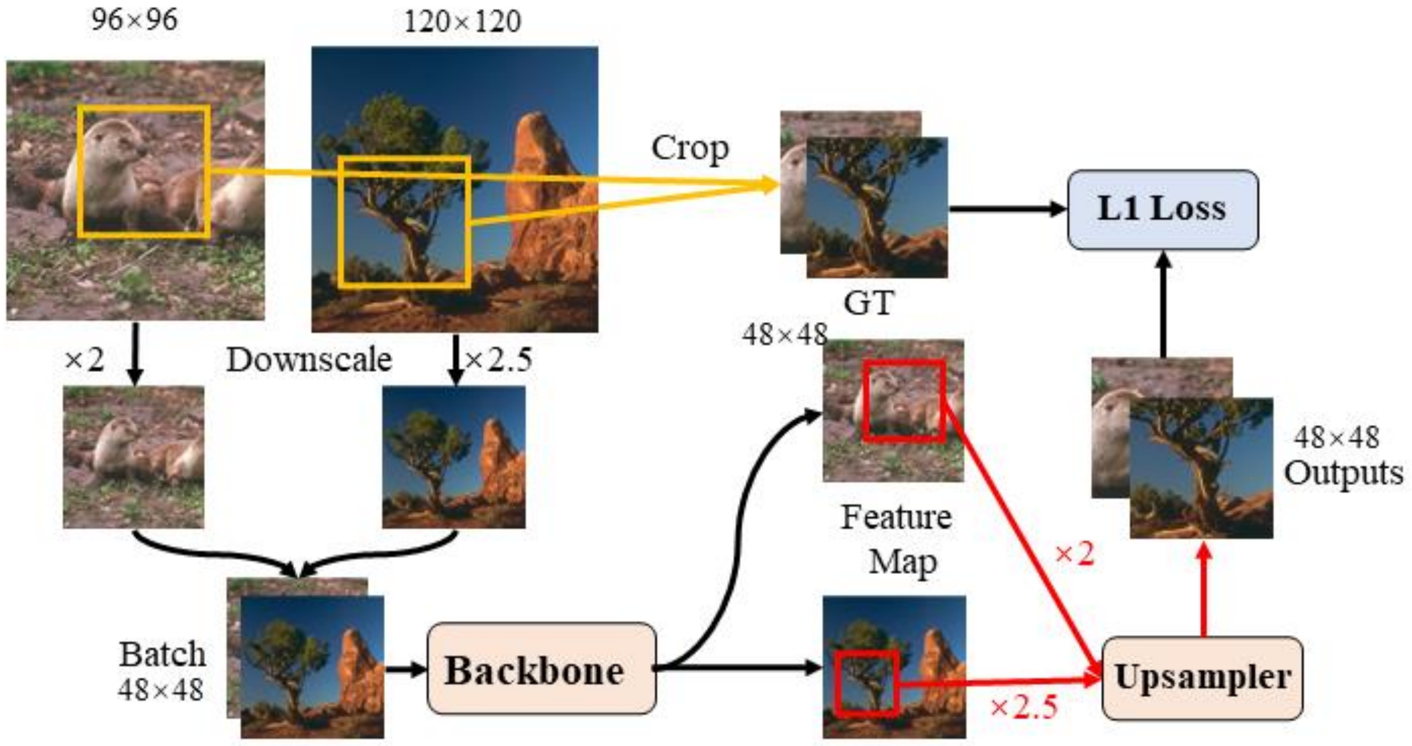
$$\hat{z}_0 = \overbrace{\mathcal{K}_{data}^{-1}(\mathcal{D})}^{scale\ independent} \ \overbrace{\mathcal{K}_{scale}(x_0; X, r)}^{scale\ dependent} \mathbf{z}$$



4. Data Processing

ArbSR [33]:
Downscales the HR images by discrete scales to generate LR data before training.
- bias towards fixed scale
- hard to handle large-scale datasets

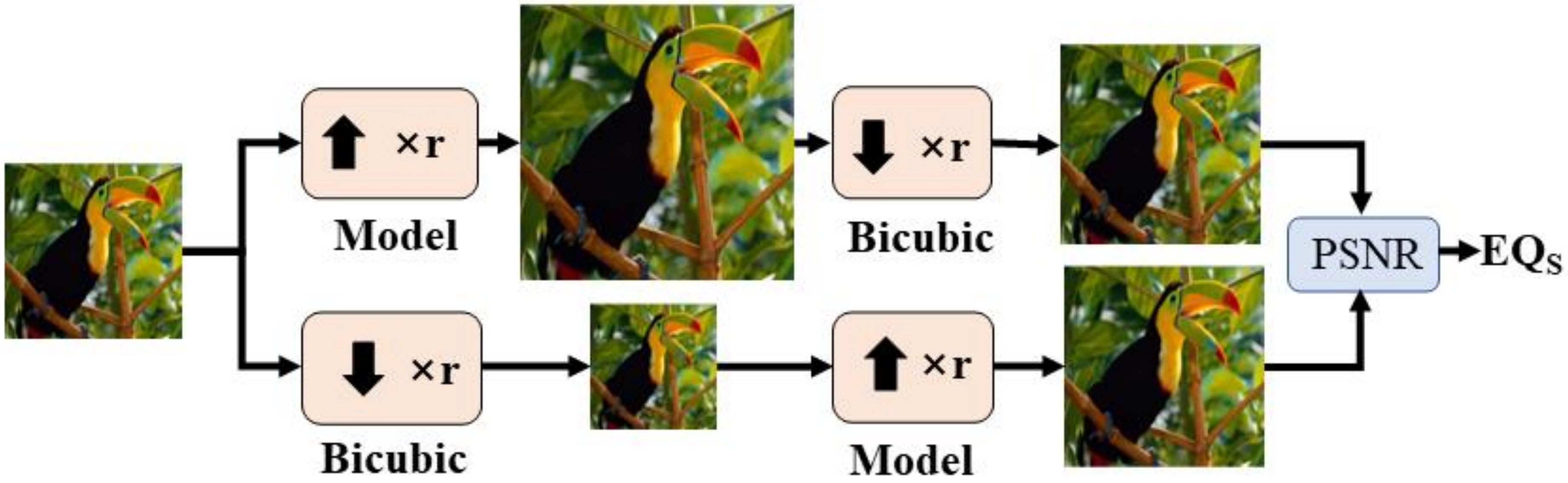
LIIF [4] :
Convert the images into coordinate-RGB pairs and sample 48x48 samples.
- multiple batches of predictions required
- MLP upsampling is necessary



Ours

[4] Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function.
[33] Longguang Wang, Yingqian Wang, Zaiping Lin, Jungang Yang, Wei An, and Yulan Guo. Learning a single network for scale-arbitrary super-resolution.

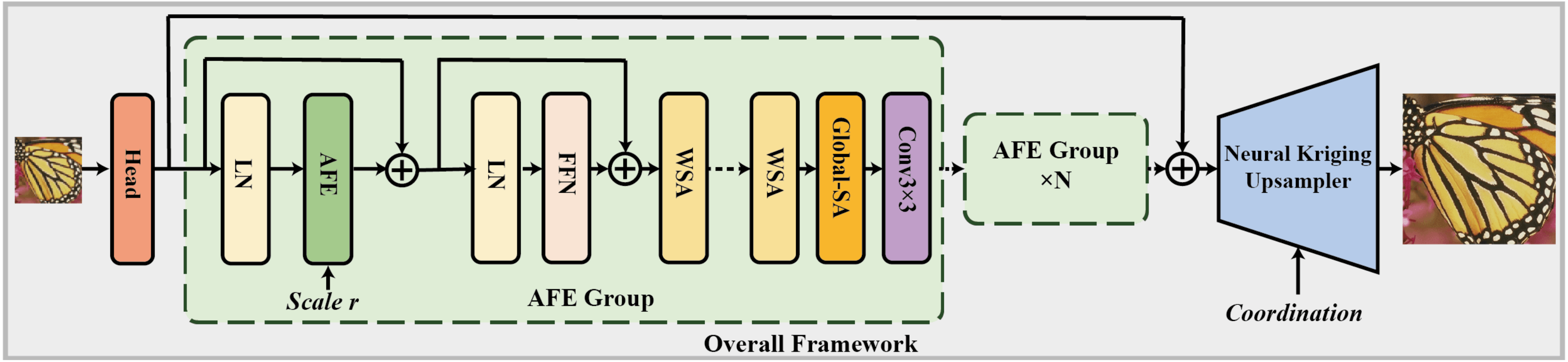
5. Measure of Scale-Equivariance



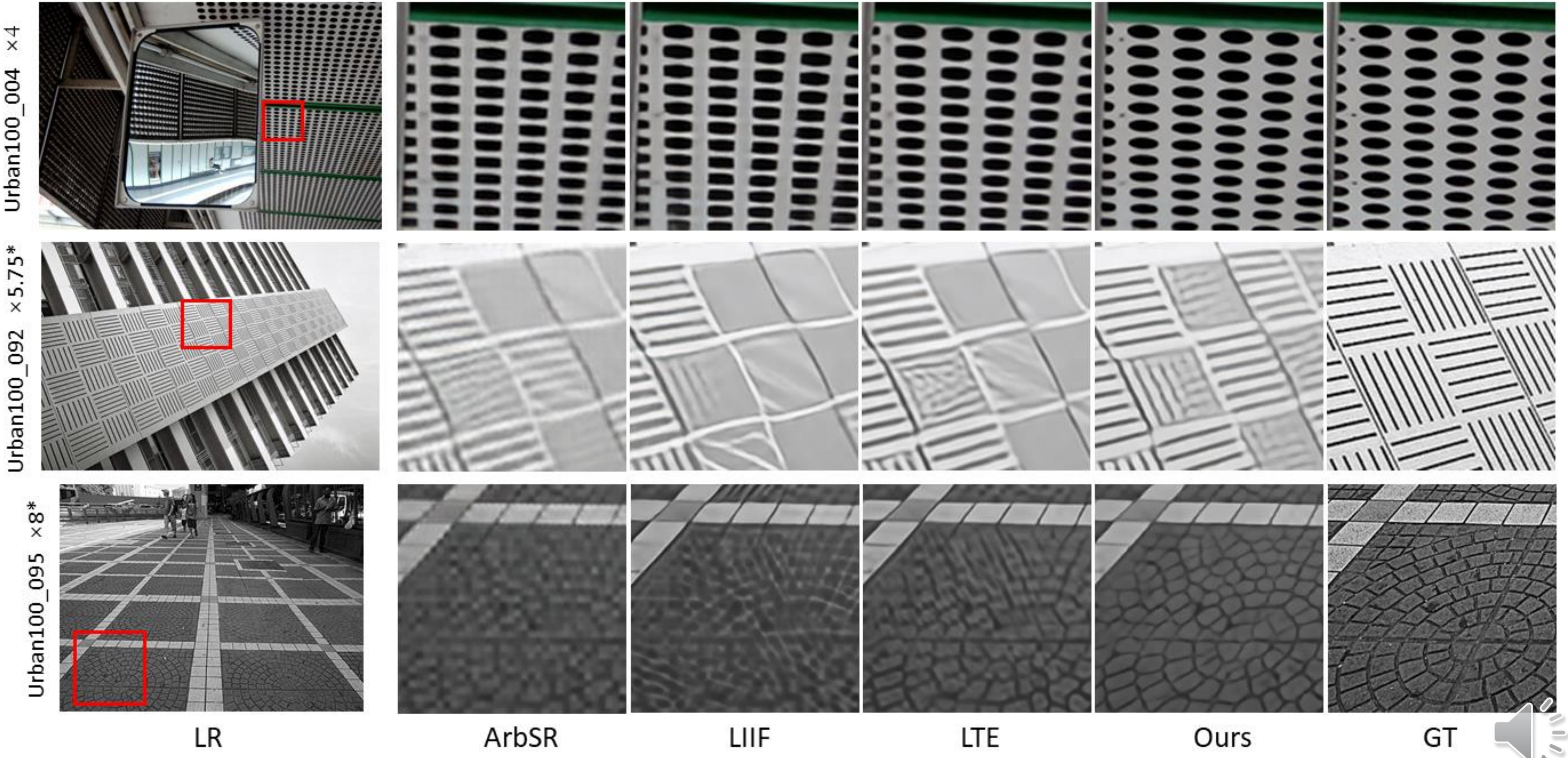
$$EQ_S(\phi, r) = 10 \log \left[\left(\phi(f(x; r), \theta; r) - f(\phi(x, \theta; r); r) \right)^2 \right] \tag{6}$$

network scale bicubic degradation

Overall Architecture



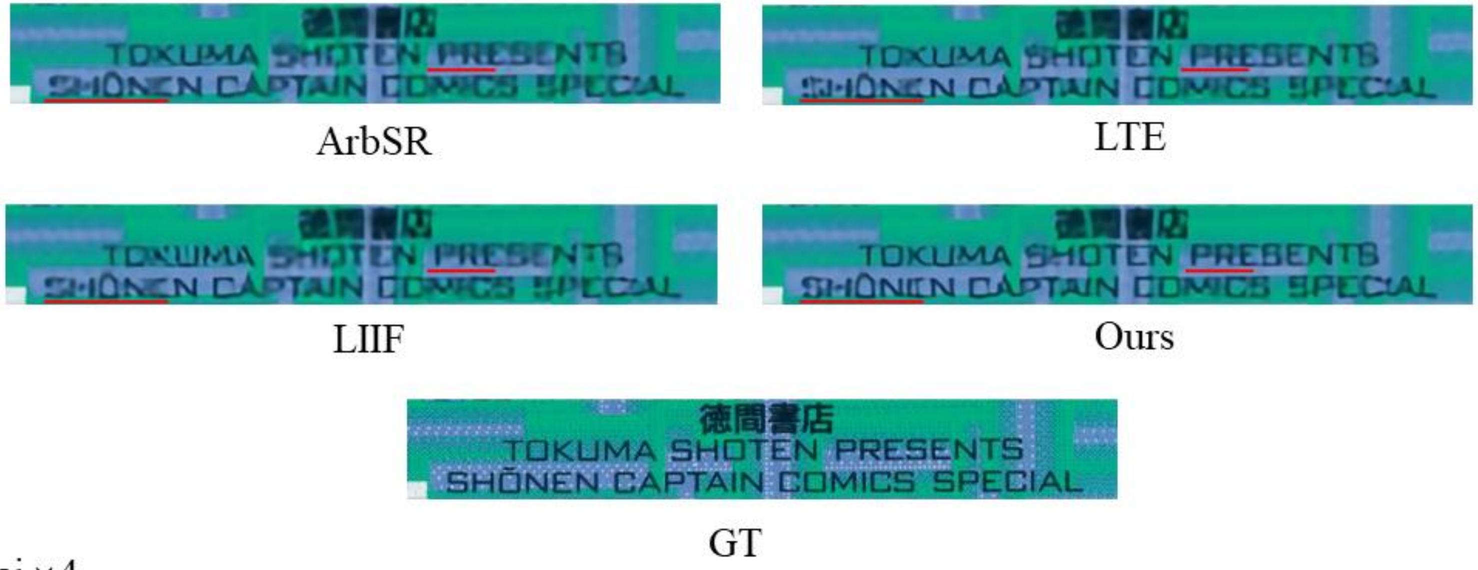
Results


Urban100_004 ×4
Urban100_092 ×5.75*
Urban100_095 ×8*
LR
ArbSR
LIIF
LTE
Ours
GT

Results

Manga109_HealingPlanet ×3.25

Manga109_PrayerHaNemurenai ×4



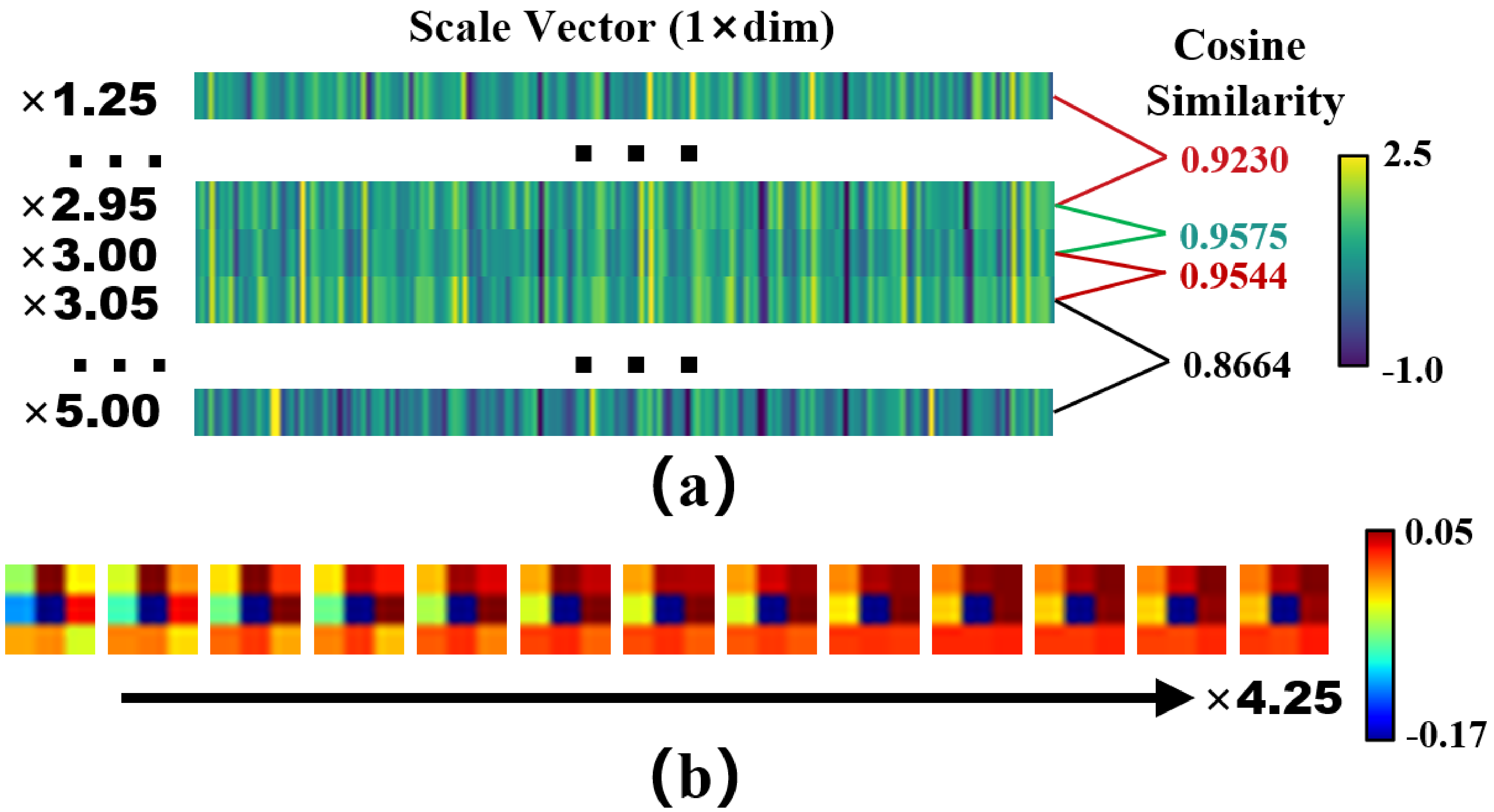
Results

	Set5			Set14			B100			Urban100			Manga109		
	×2	×1.6	×1.55	×2	×1.5	×1.65	×2	×1.4	×1.85	×2	×1.9	×1.95	×2	×1.7	×1.95
Bicubic	33.66	36.10	36.24	30.24	32.87	31.83	29.56	32.95	30.11	26.88	27.25	27.05	30.80	32.91	31.12
MetaSR [9]	38.22	40.66	40.93	34.00	37.51	36.17	32.36	36.95	33.22	33.12	33.62	33.30	39.32	41.30	39.59
ArbSR [33]	38.26	40.69	40.97	34.09	37.53	36.28	32.39	36.93	33.23	33.14	33.55	33.25	39.27	41.32	39.56
LIIF [4]	38.17	40.64	41.00	33.97	37.45	36.25	32.32	36.93	33.14	32.87	33.52	33.20	39.21	41.32	39.52
LTE [18]	38.33	40.75	41.20	34.25	37.79	36.56	32.44	37.05	33.26	33.50	34.11	33.83	39.58	41.69	39.89
Ours	38.35	40.76	41.16	34.45	38.83	36.59	32.46	37.11	33.29	33.62	34.15	33.86	39.44	41.67	39.81
Ours†	38.41	40.83	41.21	34.62	38.05	36.82	32.50	37.18	33.33	33.83	34.45	34.11	39.67	41.87	39.97

	×6*	×5.5*	×6.25*	×6*	×4.25*	×5.25*	×6*	×4.75*	×6.75*	×6*	×5.75*	×6.5*	×6*	×5.25*	×6.75*
Bicubic	24.17	24.46	23.70	23.15	24.17	23.23	23.69	24.25	23.11	20.82	20.73	20.55	21.53	21.83	20.95
MetaSR [9]	29.09	29.96	28.55	26.55	28.47	27.10	25.91	26.92	25.24	24.04	24.37	23.60	27.02	28.19	25.99
ArbSR [33]	28.45	29.24	27.66	26.22	28.46	26.89	25.74	26.89	25.16	23.70	23.81	23.23	26.18	27.59	24.93
LIIF [4]	29.15	30.04	28.73	26.64	28.50	27.38	25.98	26.96	25.53	24.20	24.44	23.79	27.34	28.54	26.37
LTE [18]	29.50	30.20	29.03	26.86	28.55	27.46	26.09	26.98	25.61	24.62	24.79	23.95	27.84	28.96	26.41
Ours	29.41	30.24	28.97	26.79	28.72	27.49	26.07	27.03	25.63	24.66	24.86	24.15	27.97	29.14	26.69
Ours†	29.51	30.38	29.12	26.90	28.78	27.63	26.11	27.10	25.66	24.83	25.10	24.36	28.04	29.36	26.85

More results can be found in our paper & Suppl.

Results



Scale Vector (1×dim)
Cosine Similarity
×1.25
×2.95
×3.00
×3.05
×5.00
0.9230
0.9575
0.9544
0.8664
2.5
-1.0
(a)
0.05
-0.17
×4.25
(b)

上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

Thank You!

饮 水 思 源 爱 国 荣 校