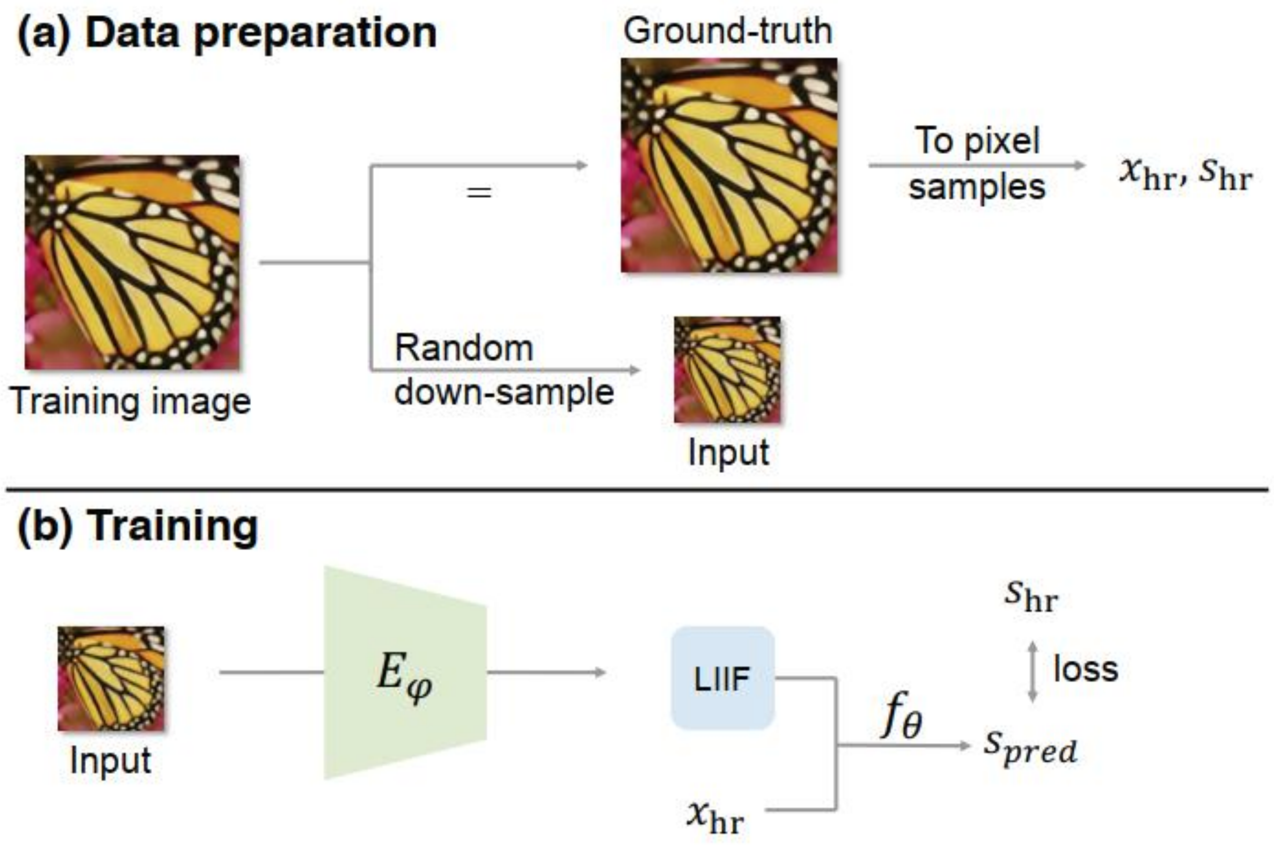
FVL2025第六期学习讲座

基于2D高斯建模的任意尺度超分辨率重建方法

主讲人：詹丹丹

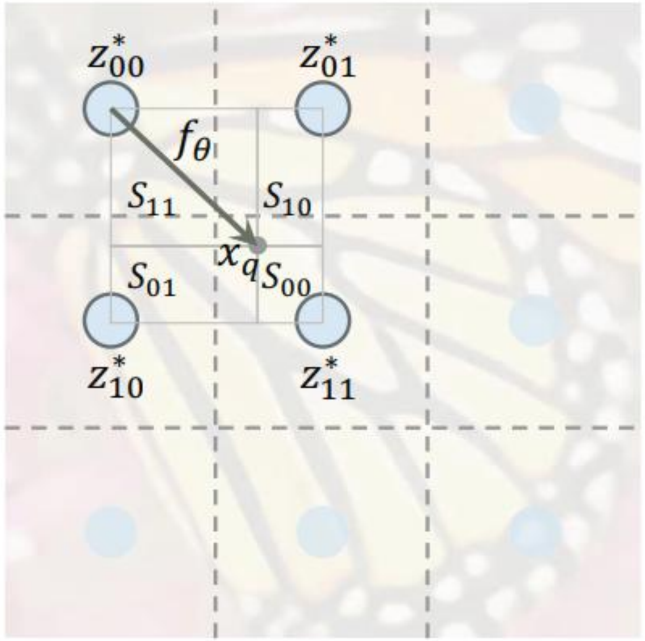
1 背景介绍

- 基于INR的任意尺度超分辨率重建



$$I^{(i)}(x_q) = f_\theta(z^*, x_q - v^*)$$

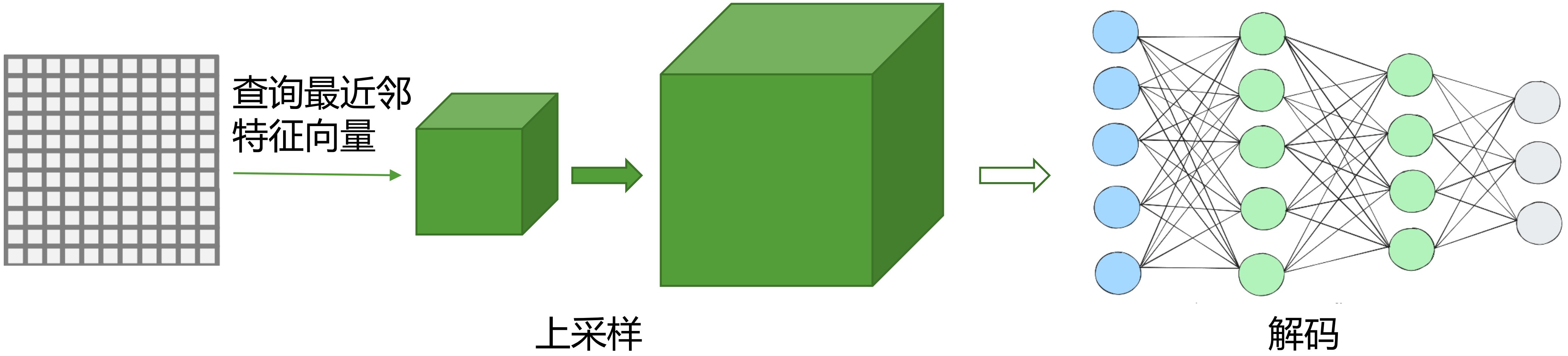
Local ensemble



$$I^{(i)}(x_q) = \sum_{t \in \{00,01,10,11\}} \frac{S_t}{S} \cdot f_\theta(z_t^*, x_q - v_t^*)$$

[1] Y. Chen, S. Liu and X. Wang, "Learning Continuous Image Representation with Local Implicit Image Function," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, 2021, pp. 8624-8634.

- 基于INR的任意尺度超分辨率重建



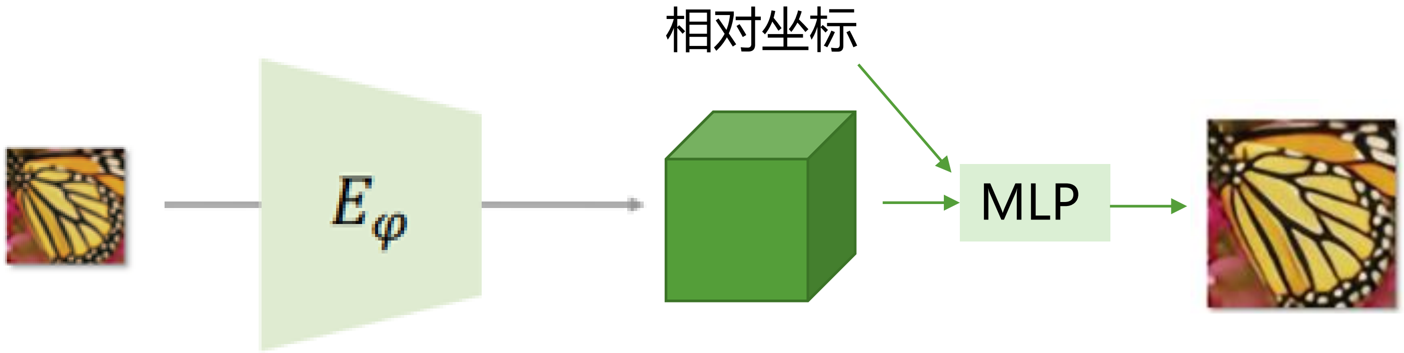

```python
        for vx in vx_lst:
            for vy in vy_lst:
                coord_ = coord.clone()
                coord_[:, :, 0] += vx * rx + eps_shift
                coord_[:, :, 1] += vy * ry + eps_shift
                coord_.clamp_(-1 + 1e-6, 1 - 1e-6)
                q_feat = F.grid_sample(
                    feat, coord_.flip(-1).unsqueeze(1),
                    mode='nearest', align_corners=False)[:, :, 0, :] \
                    .permute(0, 2, 1)
                q_coord = F.grid_sample(
                    feat_coord, coord_.flip(-1).unsqueeze(1),
                    mode='nearest', align_corners=False)[:, :, 0, :] \
                    .permute(0, 2, 1)
                rel_coord = coord - q_coord
                rel_coord[:, :, 0] *= feat.shape[-2]
                rel_coord[:, :, 1] *= feat.shape[-1]
                inp = torch.cat([q_feat, rel_coord], dim=-1)
```

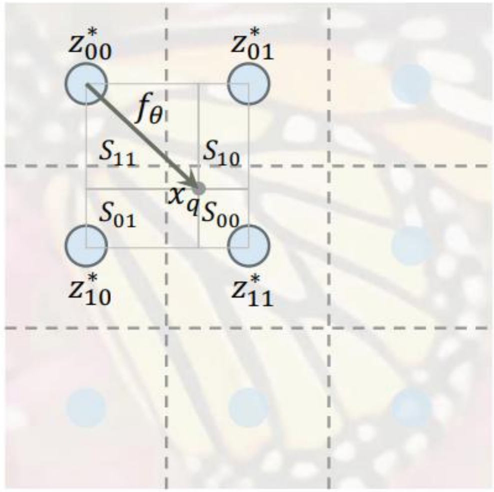
```python
                bs, q = coord.shape[:2]
                pred = self.imnet(inp.view(bs * q, -1)).view(bs, q, -1)
                preds.append(pred)
```

2

GaussianSR: High Fidelity 2D Gaussian Splatting for Arbitrary-Scale Image Super-Resolution

- 动机

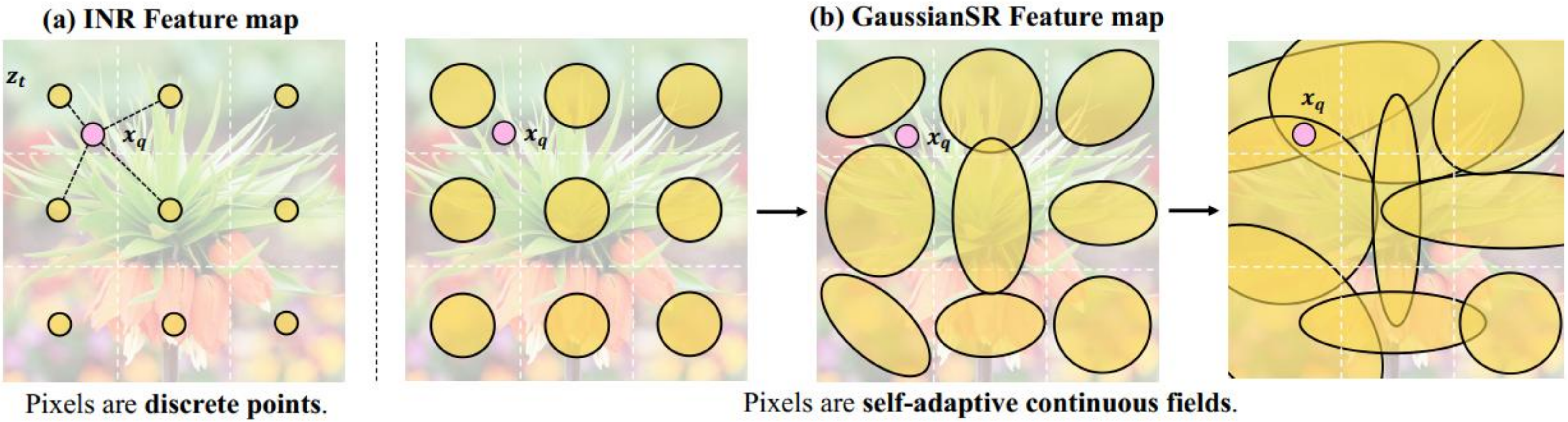




（1）基于INR的任意尺度超分辨率重建方法的性能受限于特征向量的表示能力

（2）Local ensemble会增加计算负担，可能会成为实时应用的瓶颈

- 动机



核心思想：像素值本质上表现出强度变化，这些变化可以通过连续的高斯表示更准确地捕捉

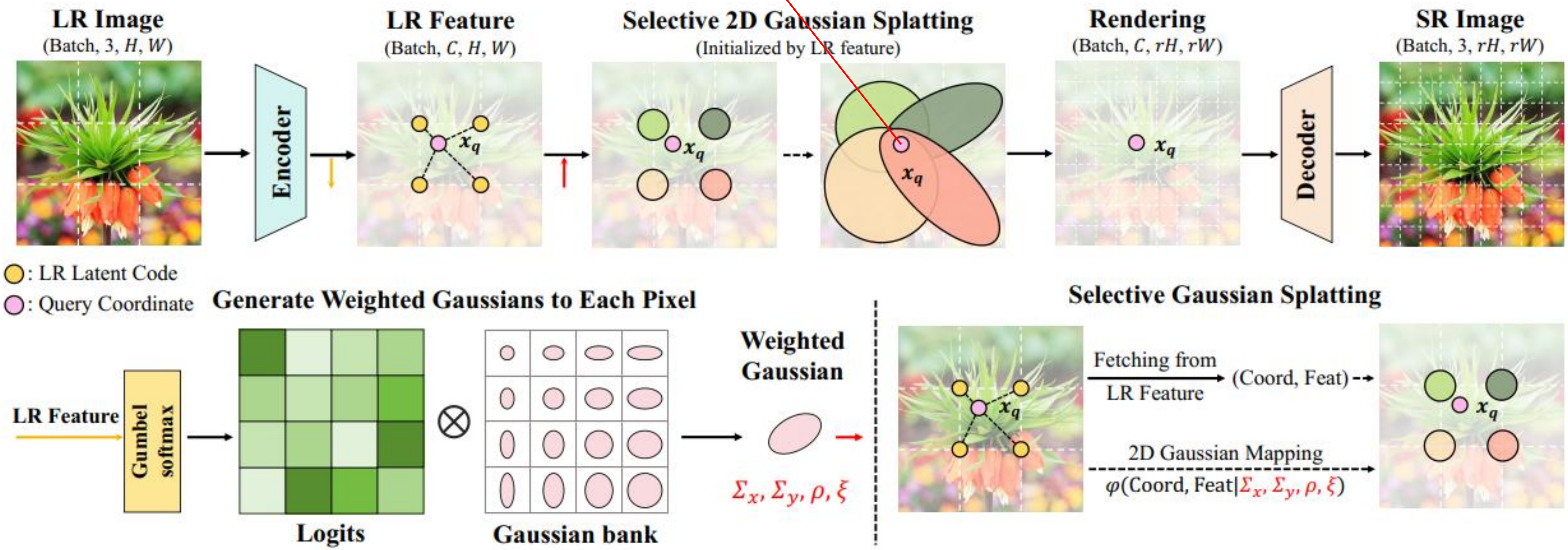
- 方法

$$\boldsymbol{\Sigma}_i = ([\Sigma_{xi}, \rho_i], [\rho_i, \Sigma_{yi}]) \in \mathbb{R}^{2\times 2}$$

$$f_i(p|\mu_i, \boldsymbol{\Sigma}_i) = \frac{1}{2\pi\,|\boldsymbol{\Sigma_i}|} e^{-\frac{1}{2}(p-\mu_i)^{\mathsf{T}}\boldsymbol{\Sigma_i}^{-1}(p-\mu_i)}$$

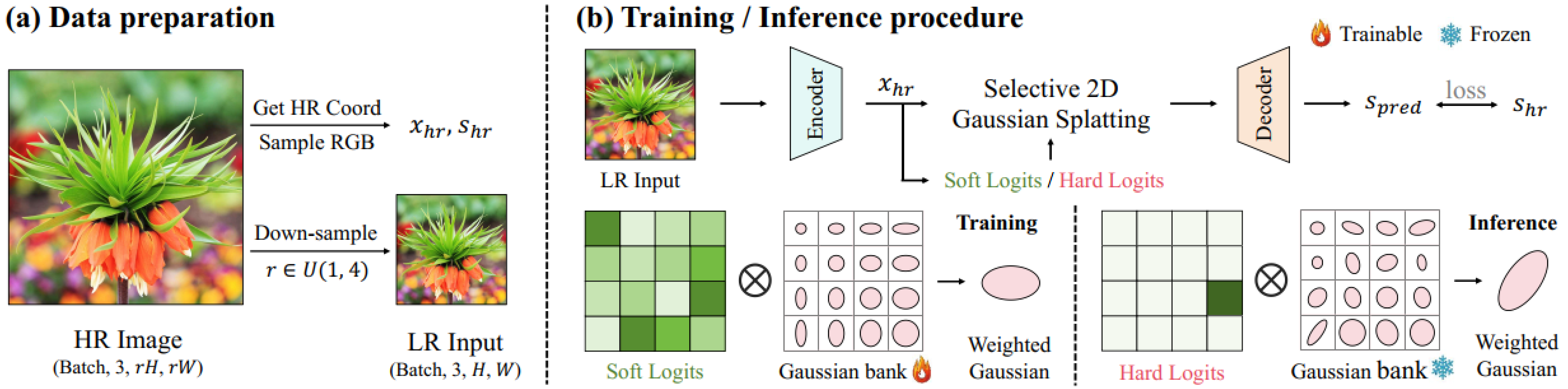
$$\boxed{C_{2DGS}(p|\mathbf{v}, \mu, \boldsymbol{\Sigma}, \xi) = \sum_i f_i(p|\mu_i, \boldsymbol{\Sigma}_i) \cdot c_i}$$

$$c_i = \sigma(\xi) \cdot \mathbf{v}_i$$

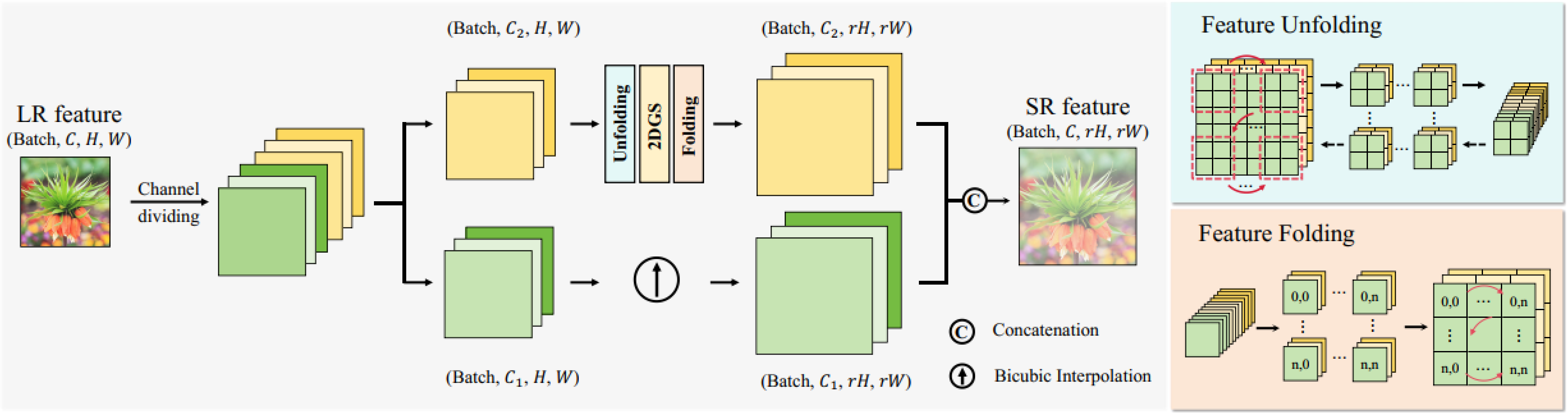


[2] J. Hu, B. Xia, B. Chen, et. al., "GaussianSR: High Fidelity 2D Gaussian Splatting for Arbitrary-Scale Image Super-Resolution," *AAAI*, *2025*.

- 方法


(a) Data preparation
Get HR Coord
Sample RGB
x_{hr}, s_{hr}
Down-sample
$r \in U(1, 4)$
HR Image
(Batch, 3, rH, rW)
LR Input
(Batch, 3, H, W)
(b) Training / Inference procedure
Trainable
Frozen
LR Input
Encoder
x_{hr}
Selective 2D Gaussian Splatting
Soft Logits / Hard Logits
Decoder
s_{pred}
loss
s_{hr}
Training
Soft Logits
Gaussian bank
Weighted Gaussian
Inference
Hard Logits
Gaussian bank
Weighted Gaussian

- 方法



LR feature
(Batch, C, H, W)
Channel dividing
(Batch, C_2, H, W)
Unfolding
2DGS
Folding
(Batch, C_2, rH, rW)
(Batch, C_1, H, W)
(Batch, C_1, rH, rW)
SR feature
(Batch, C, rH, rW)
Concatenation
Bicubic Interpolation
Feature Unfolding
Feature Folding
0,0
0,n
n,0
n,n

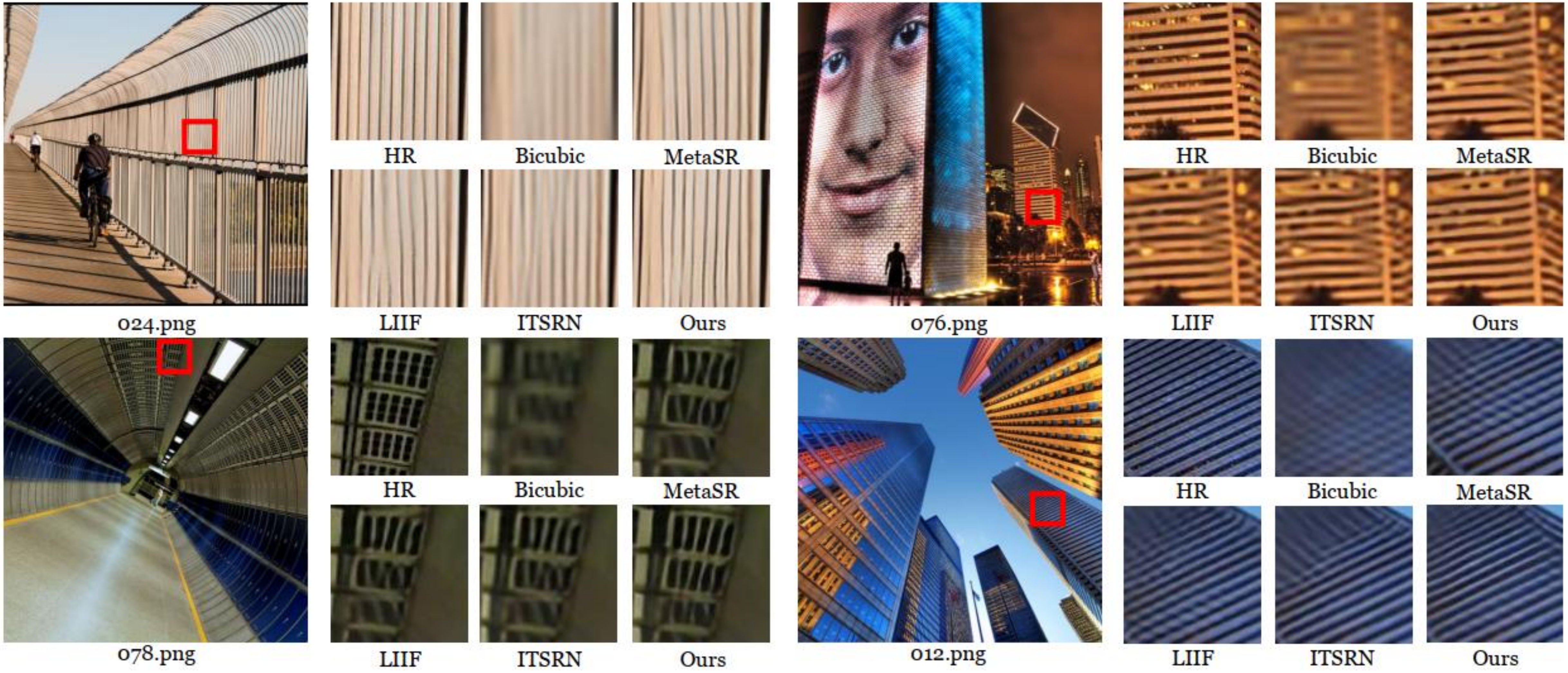
- 实验结果

Methods	General100			BSD100			Urban100			Manga109			DIV2K100		
	×2	×3	×4	×2	×3	×4	×2	×3	×4	×2	×3	×4	×2	×3	×4
Bicubic	32.14	28.56	26.58	28.25	25.96	24.69	25.68	23.07	21.77	29.98	25.68	23.52	31.45	28.42	26.81
EDSR-baseline	38.23	33.93	31.48	32.16	29.09	27.57	31.98	28.15	26.04	38.54	33.45	30.35	34.55	30.90	28.94
EDSR-baseline-MetaSR	38.22	33.93	31.40	32.16	29.09	27.55	32.08	28.12	25.95	38.53	33.51	30.37	34.64	30.93	28.92
EDSR-baseline-LIIF	38.25	33.97	31.53	32.17	29.10	27.60	32.15	28.22	26.15	38.63	33.47	30.54	34.67	30.96	29.00
EDSR-baseline-ITSRN	38.25	33.95	31.48	32.18	29.10	27.58	32.13	28.14	26.06	38.58	33.47	30.47	34.67	30.93	28.97
EDSR-baseline-ALIIF	38.21	33.95	31.48	32.18	29.11	27.60	32.09	28.19	26.14	38.53	33.42	30.47	34.65	30.95	28.99
EDSR-baseline-DIINN†	-	-	-	30.69	27.73	26.22	30.29	26.46	24.49	-	-	-	34.63	30.93	28.98
EDSR-baseline-GaussianSR	**38.31**	**34.02**	**31.55**	**32.20**	**29.13**	**27.61**	**32.25**	**28.28**	**26.19**	**38.64**	**33.57**	**30.54**	**34.71**	**31.00**	**29.03**

Methods	General100			BSD100			Urban100			Manga109			DIV2K100		
	×1.5	×2.4	×3.6	×1.5	×2.4	×3.6	×1.5	×2.4	×3.6	×1.5	×2.4	×3.6	×1.5	×2.4	×3.6
Bicubic	34.89	30.12	27.17	30.78	27.09	25.11	27.92	24.25	22.21	33.12	27.50	24.15	34.00	29.75	27.31
EDSR-baseline†	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
EDSR-baseline-MetaSR	42.16	36.14	32.30	35.69	30.60	28.08	36.05	30.07	26.74	42.67	36.18	31.56	38.57	32.78	29.62
EDSR-baseline-LIIF	42.20	36.18	32.37	35.69	30.62	28.11	36.13	30.16	26.89	42.67	36.19	31.59	38.57	32.82	29.67
EDSR-baseline-ITSRN	42.24	36.18	32.34	35.70	30.62	28.10	36.14	30.12	26.81	42.68	36.18	31.55	38.61	32.80	29.64
EDSR-baseline-ALIIF	42.14	36.16	32.33	35.67	30.60	28.10	36.02	30.07	26.84	42.56	36.11	31.53	38.54	32.79	29.65
EDSR-baseline-GaussianSR	**42.24**	**36.23**	**32.40**	**35.73**	**30.64**	**28.13**	**36.27**	**30.23**	**26.94**	**42.72**	**36.25**	**31.60**	**38.64**	**32.85**	**29.70**

2 GaussianSR

- 实验结果



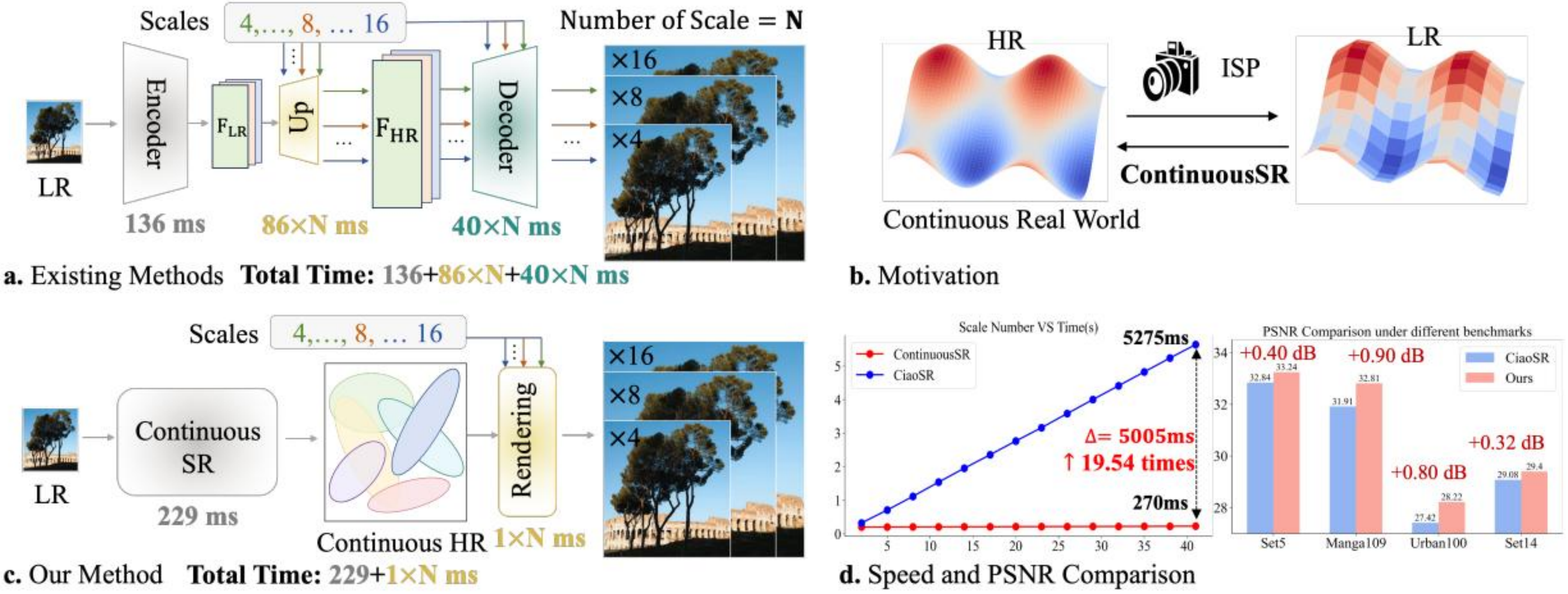
中山大學
SUN YAT-SEN UNIVERSITY

3

Pixel to Gaussian: Ultra-Fast Continuous Super-Resolution with 2D Gaussian Modeling

- 动机

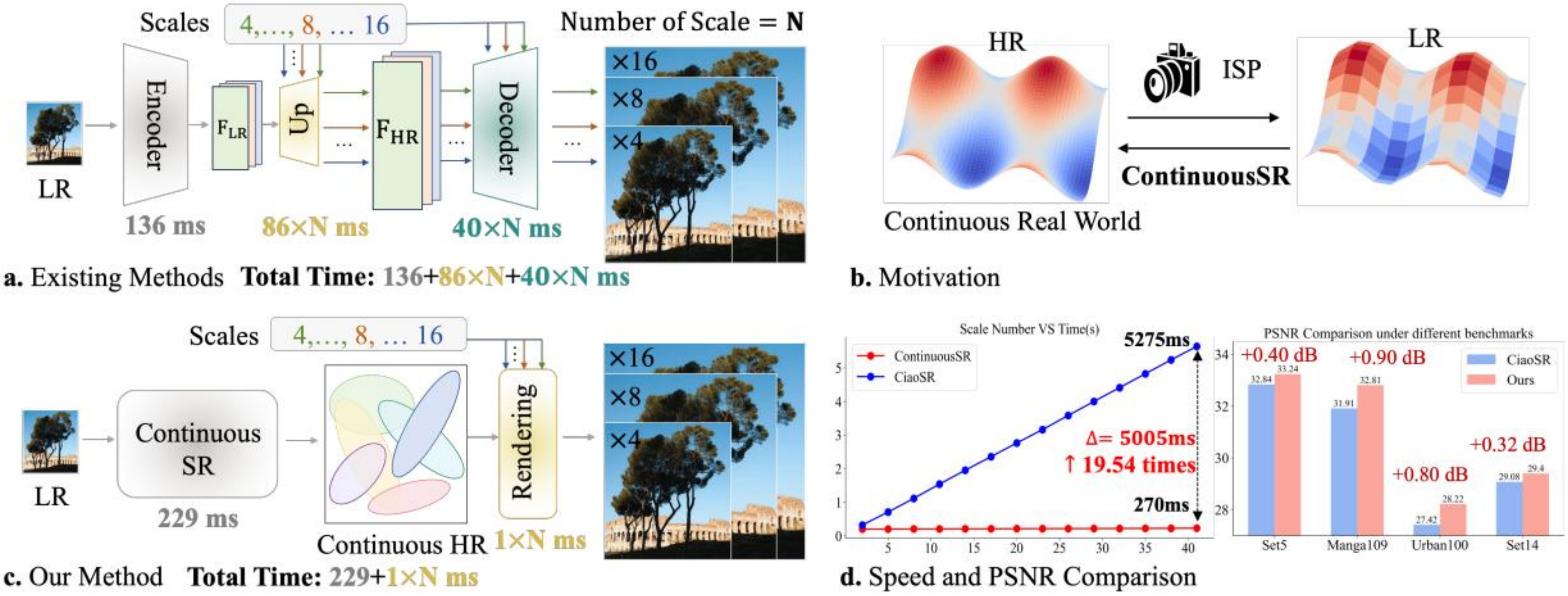
基于INR的方法通常涉及多次上采样和解码步骤，不仅效率低下，还会因为隐函数的表达能力有限而导致重建结果质量下降



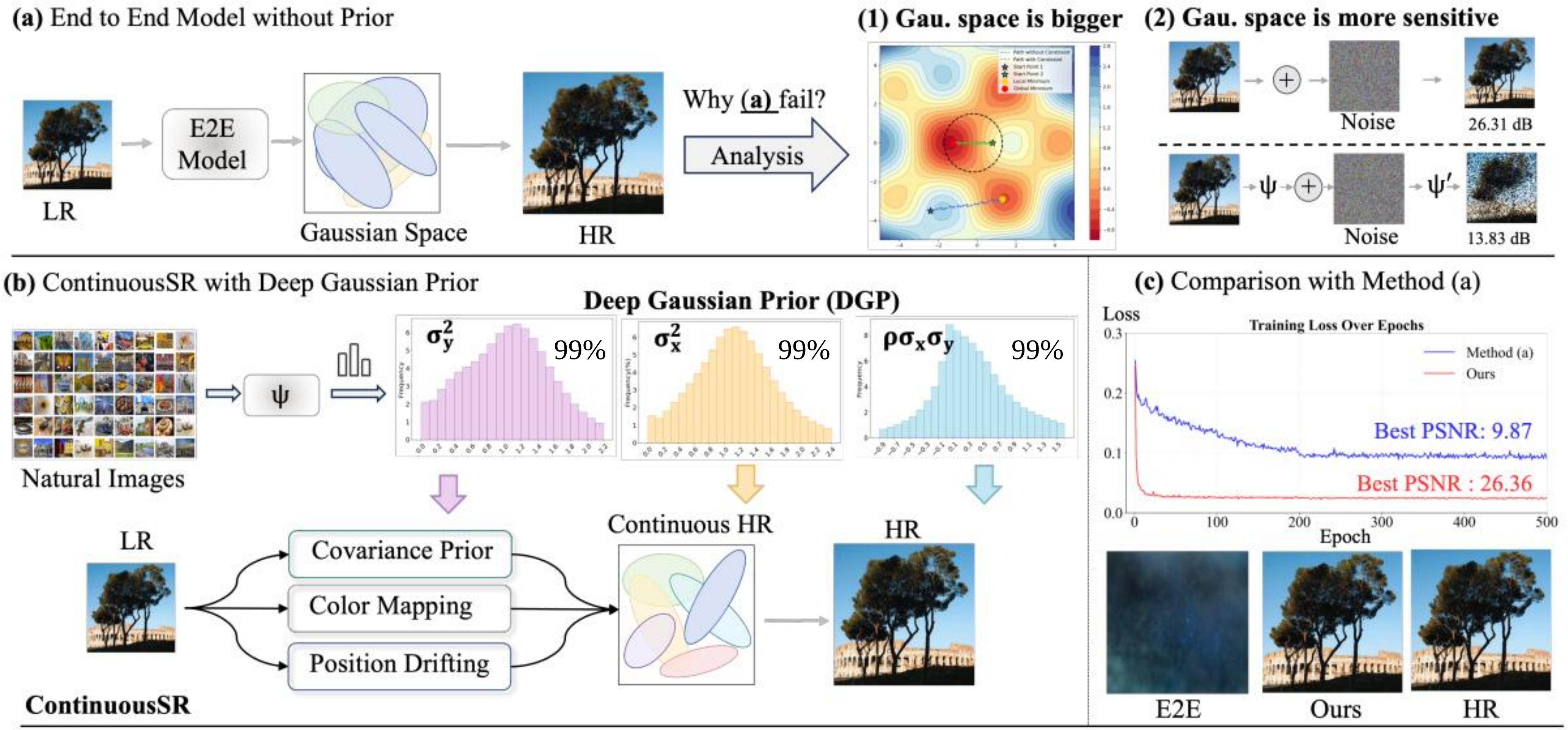
[3] L. Peng, A. Wu, W. Li, et. al., "Pixel to Gaussian: Ultra-Fast Continuous Super-Resolution with 2D Gaussian Modeling," *arXiv:2503.06617*, *2025*.

- 动机

数字图像由真实连续场景通过采样、量化和一系列后续处理得到，通过学习该过程的逆过程，即对高质量的连续函数进行显式建模，可以直接根据实际需求采样得到目标分辨率结果，而不需要进行多次上采样和解码操作



• 方法

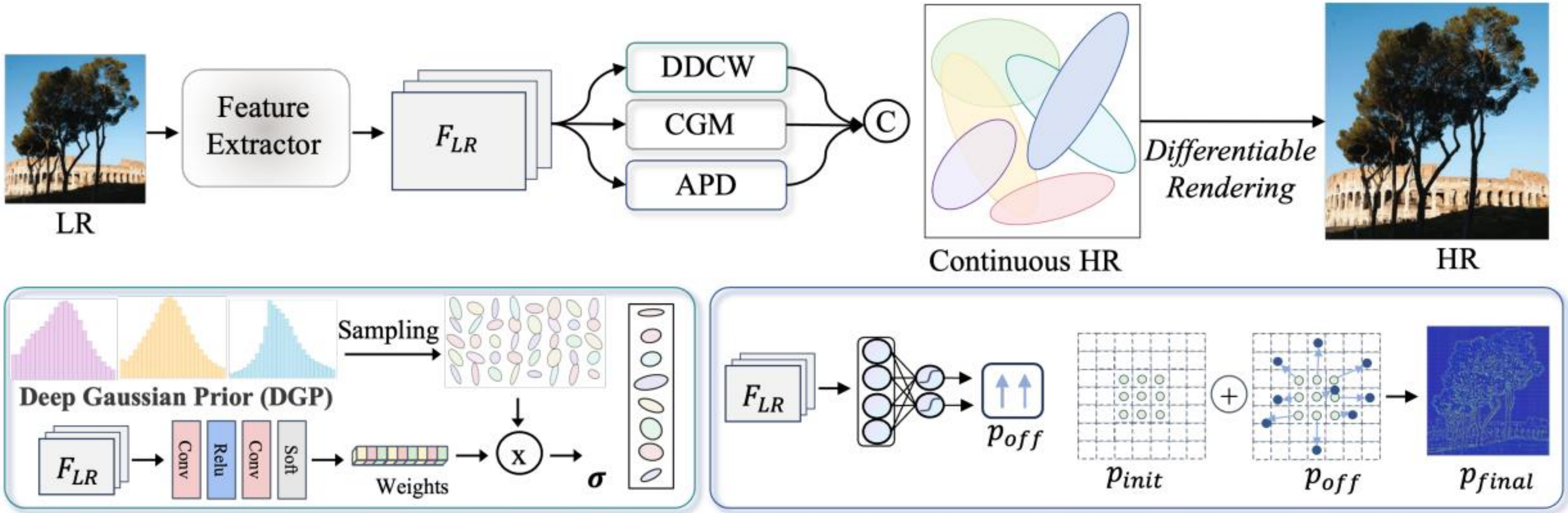


$$\Sigma = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}, \mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, c_{rgb} = \begin{bmatrix} c_r \\ c_g \\ c_b \end{bmatrix} \qquad G_i(x, y, c_{rgb}, \Sigma) = c_{rgb} \frac{1}{2\pi|\Sigma_i|} \exp\left(-\frac{1}{2} d^\top \Sigma_i^{-1} d\right)$$

3 ContinuousSR

- 方法

$$f_c(x,y) = \sum_{i=1}^{N} G_i(x,y)$$



DGP-Driven Covariance Weighting (DDCW)

Adaptive Position Drifting (APD)

$$\sigma_{i,x}^2, \sigma_{i,y}^2 \sim \mathcal{P}(\sigma_x^2), \mathcal{P}(\sigma_y^2); \rho_i\sigma_{i,x}\sigma_{i,y} \sim \mathcal{P}(\rho\sigma_x\sigma_y)$$

$$\mathcal{K} = \{G_i(\begin{bmatrix} \sigma_{i,x}^2 & \rho_i\sigma_{i,x}\sigma_{i,y} \\ \rho_i\sigma_{i,x}\sigma_{i,y} & \sigma_{i,y}^2 \end{bmatrix})\}_{i=1}^{N} \qquad G_{\text{target}} = \sum_{i=1}^{N} w_i \cdot G_i.$$

$$P_{\text{off}} = \text{Tanh}(\mathcal{M}_{\text{pos}}(\mathcal{F}_{\text{LR}}))$$

$$P_{\text{final}} = P_{\text{init}} + P_{\text{off}}.$$

3 ContinuousSR

- 实验结果

average runtime across 45 different scales, ranging from x4 to x48

PSNR↑	Methods	×4	×6	×8	×10	×12	×16	×18	×20	×32	×48	AT
Urban100 [26]	MetaSR [24]	26.76	24.31	22.92	22.02	21.31	20.35	19.96	19.65	18.38	17.48	41.4
	LIIF [11]	26.68	24.20	22.79	21.84	21.15	20.19	19.80	19.51	18.30	17.45	110.0
	LTE [34]	27.24	24.62	23.17	22.23	21.50	20.47	20.06	19.77	18.47	17.52	151.8
	SRNO [63]	26.98	24.43	23.02	22.06	21.36	20.35	19.95	19.67	18.39	17.51	65.7
	CiaoSR [3]	27.42	24.84	23.34	22.34	21.60	20.54	20.11	19.77	18.45	17.51	341.5
	MambaSR [66]	27.02	24.44	23.01	22.06	21.36	20.34	19.95	19.65	18.29	17.48	90.5
	GaussianSR [23]	26.20	23.76	22.35	21.38	20.66	19.68	19.31	19.03	17.86	17.07	321.4
	Ours	**28.22**	**25.43**	**23.87**	**22.86**	**22.08**	**20.95**	**20.54**	**20.21**	**18.77**	**17.70**	**4.6**
DIV2K [1]	MetaSR [24]	29.33	27.03	25.66	24.69	23.94	22.82	22.39	22.01	20.42	19.25	123.5
	LIIF [11]	29.27	26.99	25.60	24.63	23.89	22.77	22.34	21.94	20.36	19.19	480.6
	LTE [34]	29.50	27.20	25.81	24.84	24.09	22.94	22.50	22.12	20.50	19.31	1407.5
	SRNO [63]	29.42	27.12	25.74	24.77	24.03	22.90	22.46	22.06	20.47	19.27	390.9
	CiaoSR [3]	29.59	27.28	25.89	24.91	24.15	22.99	22.54	22.16	20.50	19.30	1857.8
	MambaSR [66]	29.36	27.08	25.70	24.74	23.99	22.87	22.44	22.05	20.46	19.27	398.3
	GaussianSR [23]	29.03	26.73	25.29	24.23	23.44	22.26	21.81	21.42	19.90	18.76	4962.8
	Ours	**29.80**	**27.47**	**26.07**	**25.08**	**24.33**	**23.18**	**22.74**	**22.35**	**20.68**	**19.45**	**4.7**
LSDIR [35]	MetaSR [24]	26.54	24.64	23.54	22.79	22.24	21.42	21.09	20.80	19.62	18.68	50.4
	LIIF [11]	26.49	24.59	23.49	22.75	22.21	21.40	21.09	20.75	19.59	18.65	226.4
	LTE [34]	26.73	24.78	23.65	22.88	22.33	21.48	21.15	20.85	19.66	18.71	451.5
	SRNO [63]	26.65	24.72	23.61	22.85	22.30	21.45	21.12	20.83	19.64	18.69	163.6
	CiaoSR [3]	26.80	24.84	23.69	22.92	22.35	21.48	21.14	20.84	19.63	18.67	1289.3
	MambaSR [66]	26.62	24.69	23.59	22.83	22.28	21.44	21.11	20.82	19.64	18.70	197.7
	GaussianSR [23]	26.25	24.39	23.28	22.49	21.92	21.06	20.74	20.45	19.29	18.38	1284.3
	Ours	**27.14**	**25.07**	**23.91**	**23.13**	**22.54**	**21.89**	**21.35**	**21.06**	**19.79**	**18.82**	**4.6**

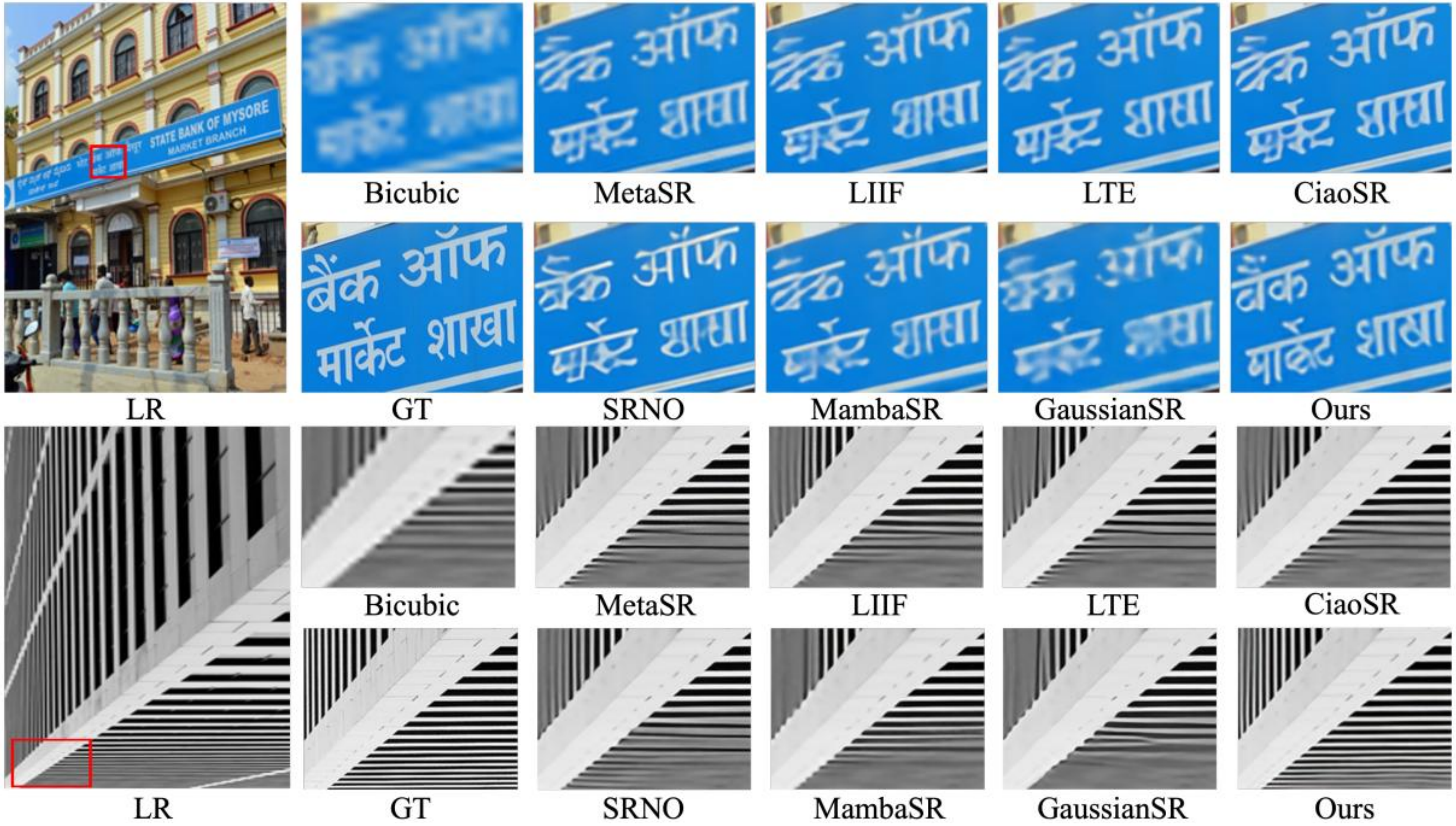
- 实验结果

Metrics	Methods	×4	×6	×8	×10	×12	×16	×18	×20	×32	×48
SSIM↑	LIIF [11]	0.7911	0.6861	0.6148	0.5642	0.5270	0.4790	0.4617	0.4503	0.4106	0.3918
	LTE [34]	0.8069	0.7045	0.6321	0.5810	0.5422	0.4900	0.4710	0.4588	0.4145	0.3931
	GaussianSR [23]	0.7751	0.6633	0.5867	0.5334	0.4967	0.4521	0.4369	0.4277	0.3969	0.3835
	CiaoSR [3]	0.8110	0.7126	0.6415	0.5887	0.5503	0.4974	0.4777	0.4637	0.4168	0.3921
	Ours	**0.8292**	**0.7343**	**0.6624**	**0.6089**	**0.5683**	**0.5097**	**0.4893**	**0.4746**	**0.4216**	**0.3958**
DISTS↓	LIIF [11]	0.1611	0.2178	0.2589	0.2926	0.3209	0.3659	0.3835	0.3990	0.4678	0.5322
	LTE [34]	0.1570	0.2126	0.2541	0.2872	0.3157	0.3611	0.3799	0.3960	0.4695	0.5362
	GaussianSR [23]	0.1740	0.2374	0.2890	0.3296	0.3631	0.4109	0.4302	0.4466	0.5157	0.5713
	CiaoSR [3]	0.1533	0.2074	0.2453	0.2771	0.3049	0.3510	0.3701	0.3863	0.4513	**0.4998**
	Ours	**0.1356**	**0.1901**	**0.2299**	**0.2601**	**0.2860**	**0.3324**	**0.3504**	**0.3670**	**0.4439**	0.5144
FID↓	LIIF [11]	4.76	24.87	50.05	77.54	102.47	145.25	164.41	179.44	256.95	311.09
	LTE [34]	3.84	21.01	45.15	70.56	92.85	136.54	156.24	170.92	253.52	296.27
	GaussianSR [23]	5.64	29.30	57.00	89.25	120.02	166.20	181.59	202.18	264.27	315.97
	CiaoSR [3]	3.74	20.48	43.25	58.60	92.58	133.77	151.70	168.89	247.84	294.49
	Ours	**2.91**	**16.50**	**37.09**	**58.83**	**78.72**	**116.30**	**130.32**	**143.52**	**216.21**	**281.27**

Memory Usage	×4	×6	×8	×12	×16
LIIF	4.12	6.49	9.79	19.27	**OOM**
CiaoSR	12.17	22.83	**OOM**	**OOM**	**OOM**
Ours	**2.48**	**2.49**	**2.50**	**2.52**	**2.54**

3 ContinuousSR

- 实验结果



3 ContinuousSR

- 消融实验

Ablation studies on proposed APD, DDCW, and DGP.

DDCW	APD	PSNR	P_{init}	P_{off}	PSNR	K	PSNR
✓		10.5	✓		27.8	$\mathcal{K}_1$	27.7
	✓	12.3		✓	10.5	$\mathcal{K}_2$	27.1
✓	✓	**28.2**	✓	✓	**28.2**	$\mathcal{K}_{DCP}$	**28.2**

[0, 1] [0, 10]

Number of K.

	PSNR	SSIM
100	28.12	0.8277
500	28.19	0.8286
730	28.22	0.8292

Number of N

	PSNR	SSIM
1	28.01	0.8254
4	28.22	0.8292
9	28.18	0.8284

Ablation on P_{off}

	PSNR	SSIM
0.5	28.03	0.8259
1	28.22	0.8292
2	28.17	0.8283

高斯核位置分布

THANKS

感谢聆听