Towards Realistic Data Generation for Real-World Super-Resolution

Long Peng1,2∗, Wenbo Li2†, Renjing Pei2, Jingjing Ren3, Jiaqi Xu4, Yang Wang1,5†, Yang Cao1, Zheng-Jun Zha1

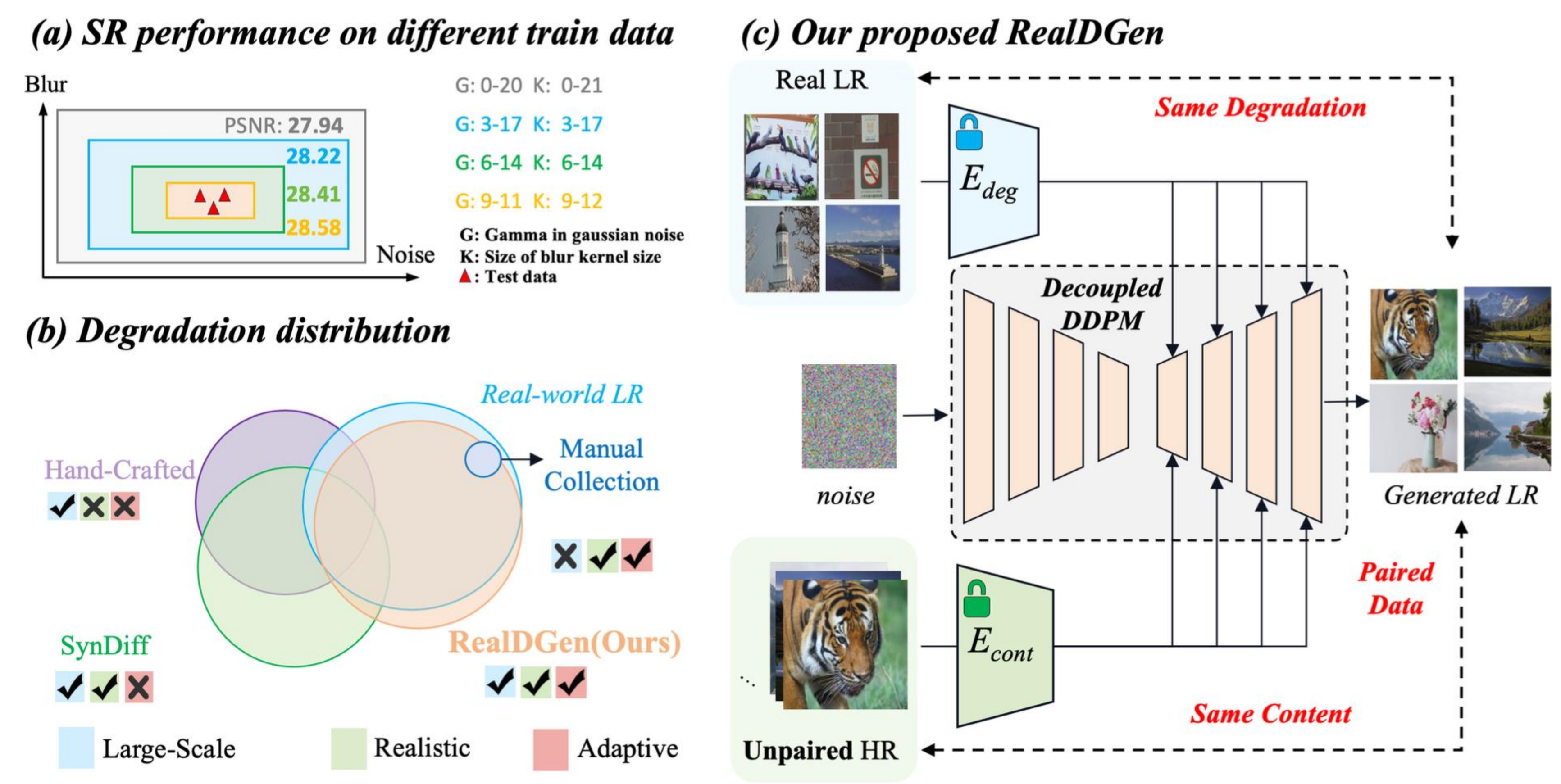
1 USTC, 2 Huawei Noah's Ark Lab, 3 HKUST(GZ), 4 CUHK, 5 Chang'an University

{longp2001@mail.,ywang120@}ustc.edu.cn,liwenbo50@huawei.com. † corresponding authors.

Motivation and Contribution

An ideal data generation system for Real-world image Super-Resolution should meet the following criteria:

- I) Large-scale, to satisfy the extensive data requirements for training deep learning models.
- II) Realistic, to enable Real SR models to accurately learn the characteristics of real-world degradation.
- III) Adaptive, to flexibly generate data with arbitrary given degradation patterns, improving generalization in target domains.

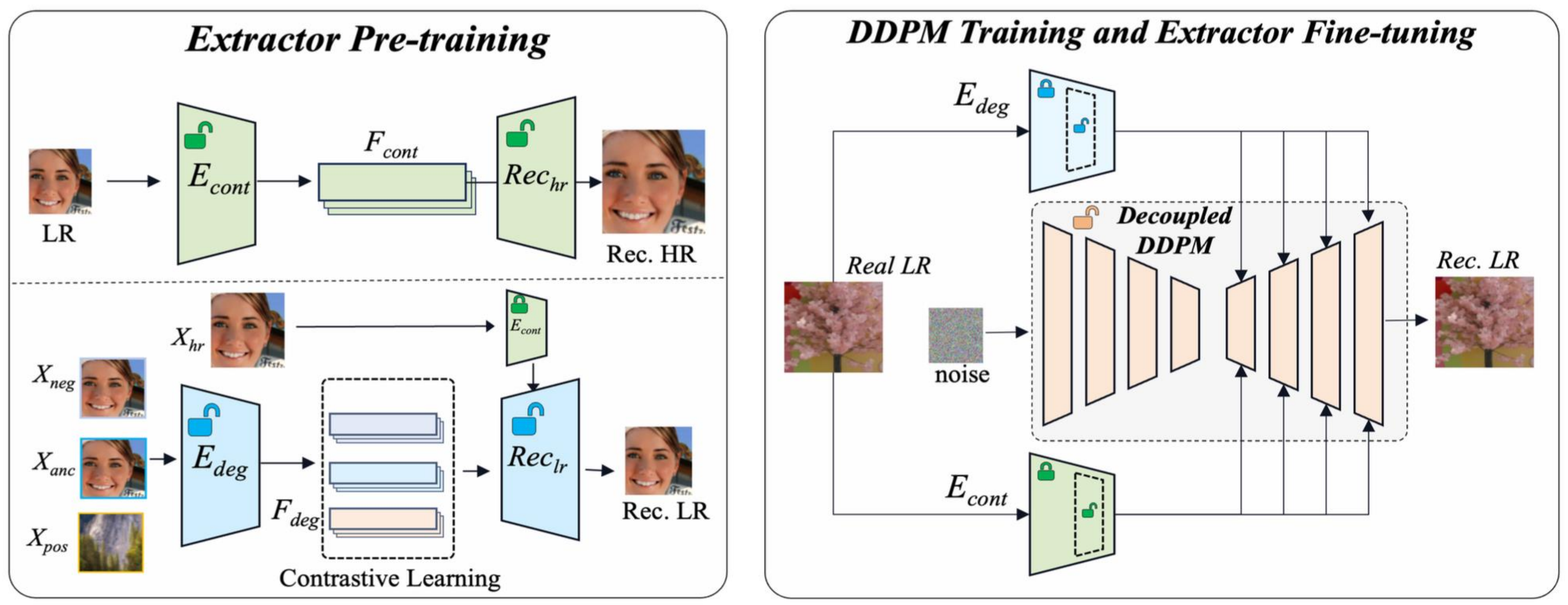


- We propose a novel unsupervised Realistic Decoupled Data Generator (RealDGen) to adaptively generate large-scale, realistic, and diverse data for real-world super-resolution.
- We introduce well-designed content and degradation extraction strategies and a novel content-degradation decoupled diffusion model to generate realistic LR with arbitrary unpaired LRand HR conditions.
- Compared with previous methods, our method significantly advances the generalization ability of popular SR models, achieving the best performance on real-world benchmarks.

Pre-training and Fine-tuning

RealDGen

Pipeline



Algorithm 1 Decoupled DDPM Training

repeat
  $\mathbf{x}_{lr} \sim q(\mathbf{x}_{lr})$
  $t \sim \text{Uniform}(\{1, \ldots, T\})$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
  $F_{cont} = E_{cont}(\mathbf{x}_{lr})$
  $F_{deg} = E_{deg}(\mathbf{x}_{lr})$
  $\mathbf{c} = \mathcal{M}(F_{cont}, F_{deg})$
  Take gradient descent step on
  $$\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, \mathbf{c}, t) \right\|^2$$
until converged

Algorithm 2 Data Generation

$\mathbf{x}_{lr} \sim q(\mathbf{x}_{lr}), \mathbf{x}_{hr} \sim p(\mathbf{x}_{hr})$
$\mathbf{c} = \mathcal{M}(E_{deg}(\mathbf{x}_{lr}), E_{cont}(\mathbf{x}_{hr}))$
$\tau \sim \text{Uniform}(\{1, \ldots, T\})$
$\mathbf{x}_t = (\sqrt{\bar{\alpha}_t}\mathcal{D}(\mathbf{x}_{lr}) + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}), \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = \tau, \ldots, 1$ **do**
  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
  $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{c}, t)\right) + \sigma_t \mathbf{z}$
end for
return $\mathbf{x}_0$

Experimental Results

PSNR-oriented Training	RealSR PSNR↑	RealSR SSIM↑	DRealSR PSNR↑	DRealSR SSIM↑	SmartPhone PSNR↑	SmartPhone SSIM↑
SwinIR (Real-ESRGAN)	24.395	0.7760	26.944	0.8308	27.395	0.8338
SwinIR (BSRGAN)	25.852	0.7808	27.985	0.8308	28.049	0.8407
SwinIR (SynDiff)	25.589	0.7687	28.301	0.8309	28.566	0.8453
SwinIR (Ours)	**26.094**	**0.7822**	**28.721**	**0.8341**	**28.737**	**0.8489**
RRDB (Real-ESRGAN)	24.579	0.7614	27.131	0.8193	27.841	0.8378
RRDB (BSRGAN)	25.406	0.7685	27.523	0.8017	28.029	0.8278
RRDB (SynDiff)	25.488	0.7691	28.078	0.8257	28.303	0.8426
RRDB (Ours)	**26.238**	**0.7747**	**28.727**	**0.8340**	**28.754**	**0.8507**
HAT (Real-ESRGAN)	24.893	0.7726	27.339	0.8215	27.781	0.8336
HAT (BSRGAN)	25.997	0.7816	28.135	0.8273	28.137	0.8369
HAT (SynDiff)	25.790	0.7584	28.506	0.8286	28.508	0.8471
HAT (Ours)	**26.140**	**0.7832**	**28.802**	**0.8345**	**28.767**	**0.8489**
SwinIR-L (Real-ESRGAN)	24.367	0.7723	27.018	0.8244	27.581	0.8409
SwinIR-L (BSRGAN)	25.651	0.7800	27.813	0.8301	28.118	0.8437
SwinIR-L (SynDiff)	25.281	0.7516	28.170	0.8244	28.474	0.8487
SwinIR-L (Ours)	**26.025**	**0.7810**	**28.869**	**0.8328**	**28.868**	**0.8522**

Perceptual-oriented Training	RealSR LPIPS↓	RealSR FID↓	DRealSR LPIPS↓	DRealSR FID↓	SmartPhone LPIPS↓	SmartPhone FID↓
SwinIR (Real-ESRGAN)	0.3037	69.965	0.3219	39.175	0.4053	78.242
SwinIR (BSRGAN)	0.2945	79.833	0.3023	38.541	0.3043	76.871
SwinIR (SynDiff)	0.3835	103.179	0.3801	54.588	0.3129	83.485
SwinIR (Ours)	**0.2536**	**69.736**	**0.2660**	**38.257**	**0.2964**	**74.778**
RRDB (Real-ESRGAN)	0.3480	82.056	0.3551	39.310	0.3480	77.573
RRDB (BSRGAN)	0.3041	77.412	0.3127	36.528	0.3381	78.812
RRDB (SynDiff)	0.4004	98.798	0.4017	56.573	0.3511	88.431
RRDB (Ours)	**0.2972**	**76.973**	**0.3077**	**36.259**	**0.3125**	**76.723**
HAT (Real-ESRGAN)	0.3066	79.209	0.3219	41.862	0.4022	87.950
HAT (BSRGAN)	0.2852	80.192	0.2835	41.723	0.3049	81.247
HAT (SynDiff)	0.3332	93.763	0.3465	50.808	0.3171	85.587
HAT (Ours)	**0.2457**	**67.573**	**6.2587**	**41.319**	**0.2816**	**76.873**
SwinIR-L (Real-ESRGAN)	0.3108	79.491	0.3234	42.986	0.4021	87.531
SwinIR-L (BSRGAN)	0.3013	84.195	0.2978	43.246	0.3146	83.444
SwinIR-L (SynDiff)	0.3793	98.646	0.3748	52.464	0.3190	80.708
SwinIR-L (Ours)	**0.2795**	**75.779**	**0.2862**	**42.542**	**0.3047**	**78.632**