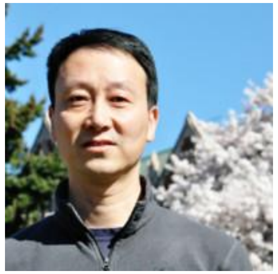
AIM 2021

Real-ESRGAN: Training *Real-World* Blind Super-Resolution with *Pure Synthetic Data*



Xintao Wang[1] Liangbin Xie[2,3] Chao Dong[2,4] Ying Shan[1]

[1] ARC Lab, Tencent PCG

[2]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences

[3]University of Chinese Academy of Sciences

[4]Shanghai AI Laboratory

BasicSR Codes & Models

Single Image Super-Resolution (SR)

- ✗ Most approaches (such as ESRGAN) assume an ideal **Bicubic** downsampling kernel, which is different from real degradations.

Our Goal – Real-World Blind Super-Resolution

✓ We extend the powerful ESRGAN to a practical restoration application – *Real-ESRGAN*.
Real-ESRGAN aims at developing practical algorithms for general image restoration.

Input (Bicubic) | Real-ESRGAN Output | Input (Bicubic) | Real-ESRGAN Output

Challenges

- Unknown and complex degradations
    - Usually, **paired training data** with similar degradations to real scenarios is required to train the networks.

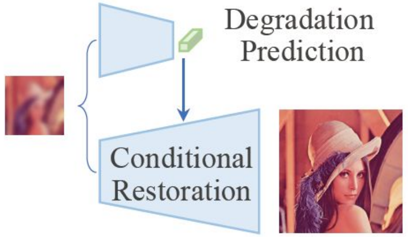
Capture paired data with specific cameras followed by alignments
e.g., RealSR[1]

Directly learn degradation distributions and then synthesize paired training data
e.g., Cycle-in-Cycle GAN[2]

Synthesize paired data with classical operators and generalize trained models to real degradations
As close to real data as possible

- Deal with diverse degraded images in one unified network

Typical blind SR methods: two-branch network
e.g., DAN[3]



Process all degraded images in one network

More discussion

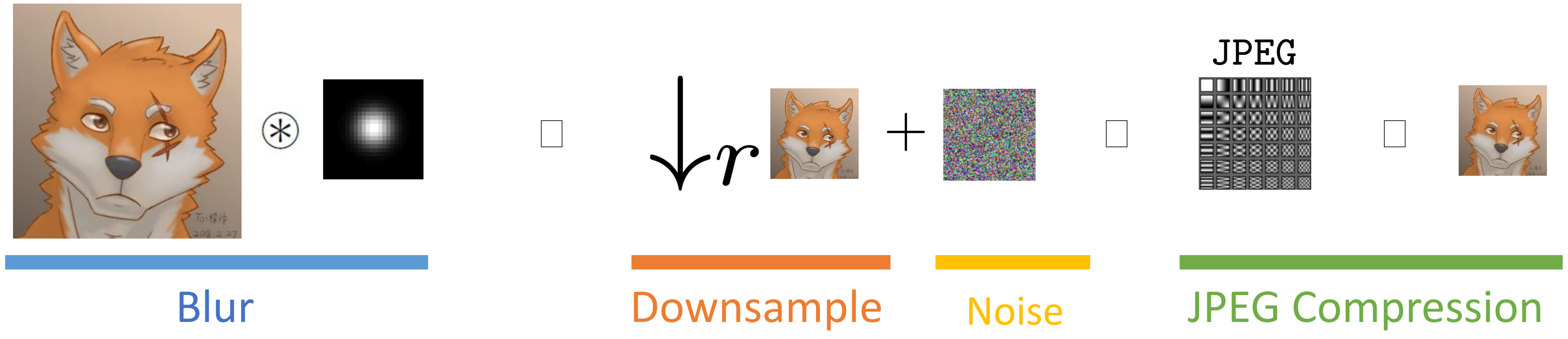
NEURAL INFORMATION PROCESSING SYSTEMS
NeurIPS 2021

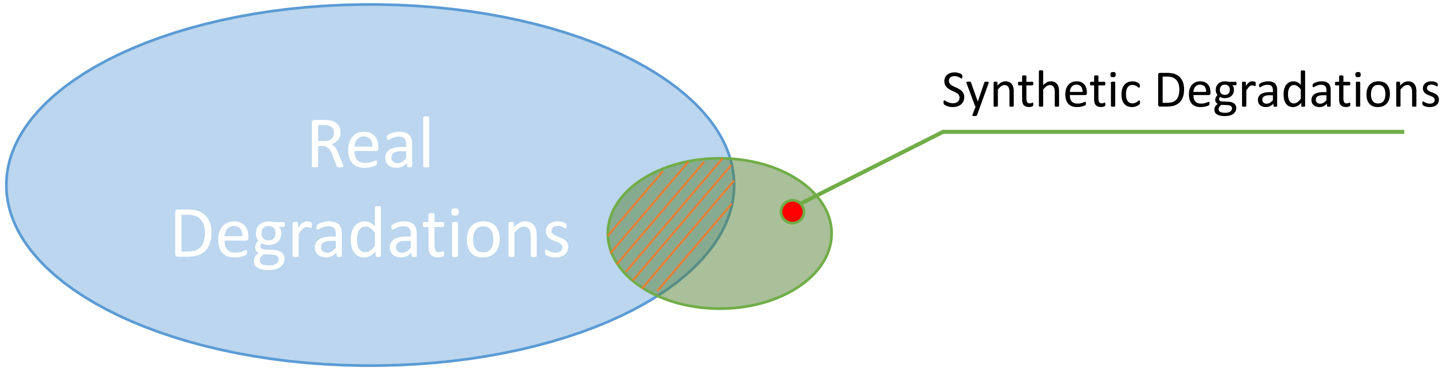
Finding Discriminative Filters for Specific Degradations in Blind Super-Resolution

[1] Cai J, Zeng H, Yong H, et al. Toward real-world single image super-resolution: A new benchmark and a new model. ICCV 2019
[2] Yuan Y, Liu S, Zhang J, et al. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. CVPRW 2018
[3] Luo Z, Huang Y, Li S, et al. Unfolding the alternating optimization for blind super resolution. NeurIPS 2020.
[4] Xie L, Wang X, Dong C, et al. Finding Discriminative Filters for Specific Degradations in Blind Super-Resolution. NeurIPS, 2021.

Classical Degradation Model

$$\boldsymbol{x} = [(\boldsymbol{y} \circledast \boldsymbol{k}_\sigma)\downarrow_r + \boldsymbol{n}_\delta]_{\mathtt{JPEG}_q}$$

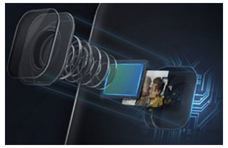


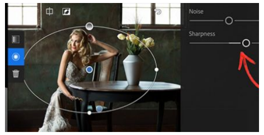


Complicated Combinations of Degradation Processes

The real complex degradations usually come from complicate combinations of different degradation processes, such as imaging system of cameras, image editing, and Internet transmission.

Take Photo





Camera Blur

Sensor Noise

Sharpening Artifacts

JPEG compression

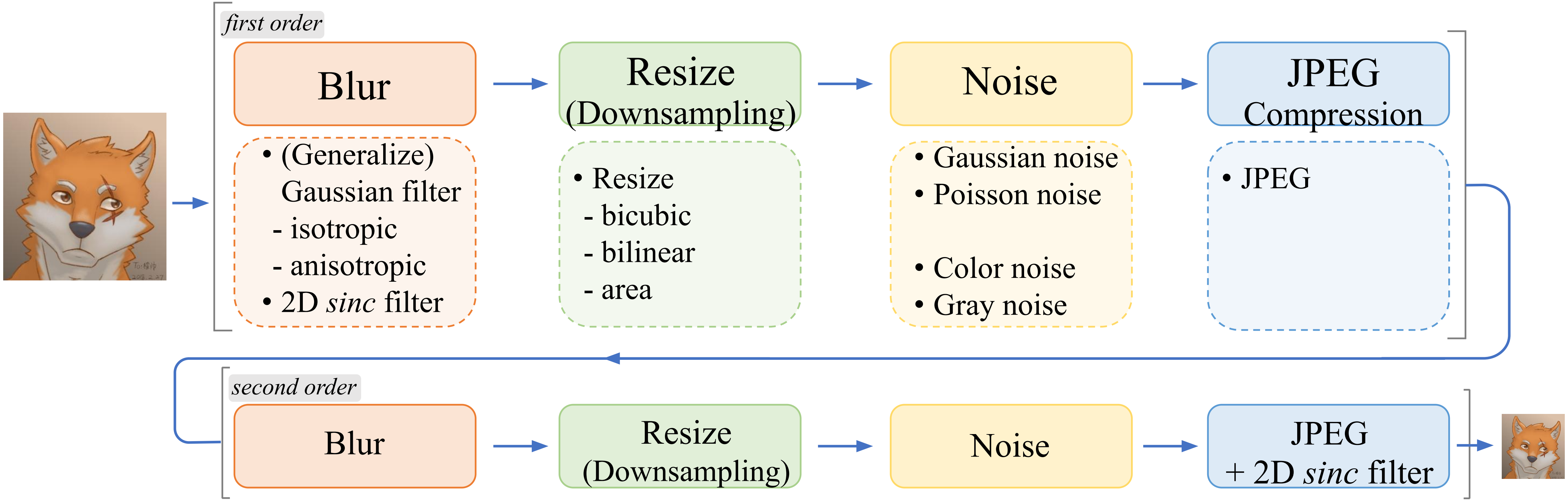
Image Editing

Upload to Social Media

Further Compression
and
Unpredictable Noises

Shared Several Times
on the Internet

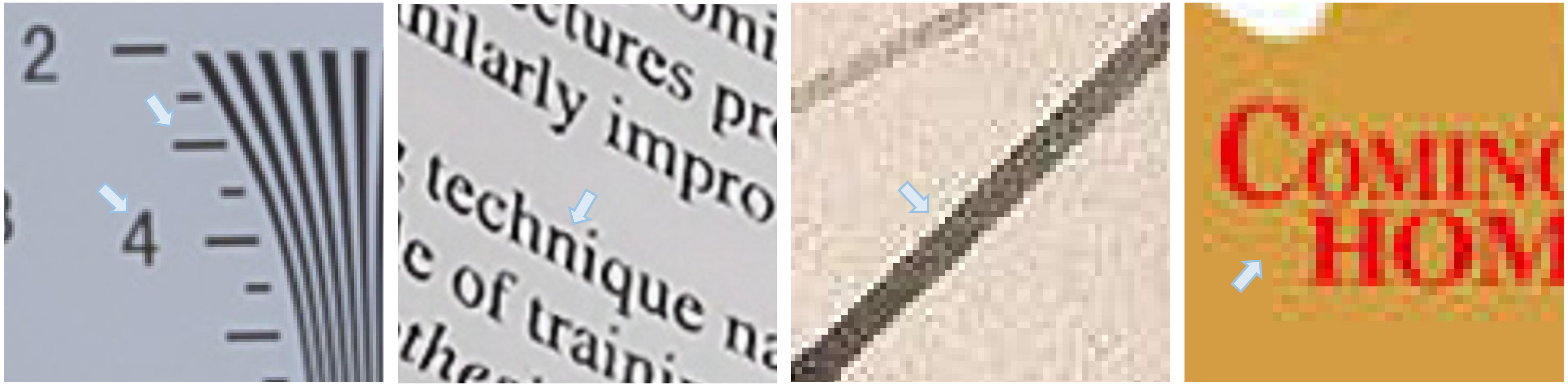
High-Order Degradation Process

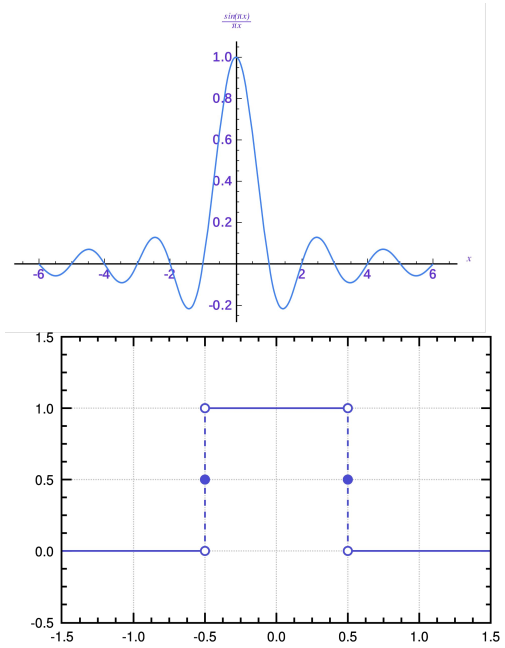


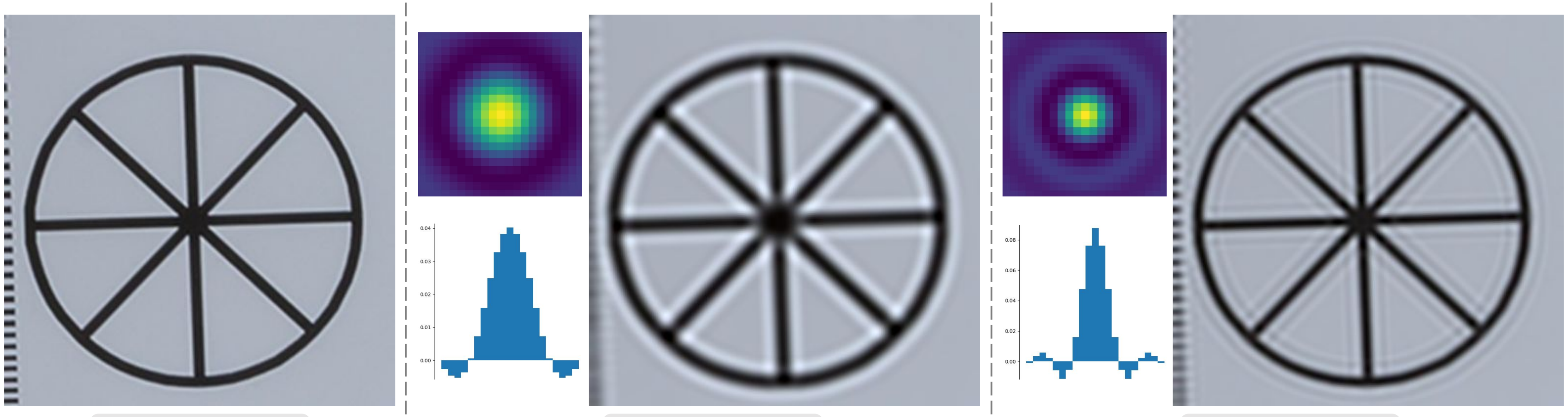
* The “high-order” here is different from that used in mathematical functions. It mainly refers to the implementation time of the same operation.

Sinc Filter for Ringing and Over-shoot Artifacts

Real Samples







Input

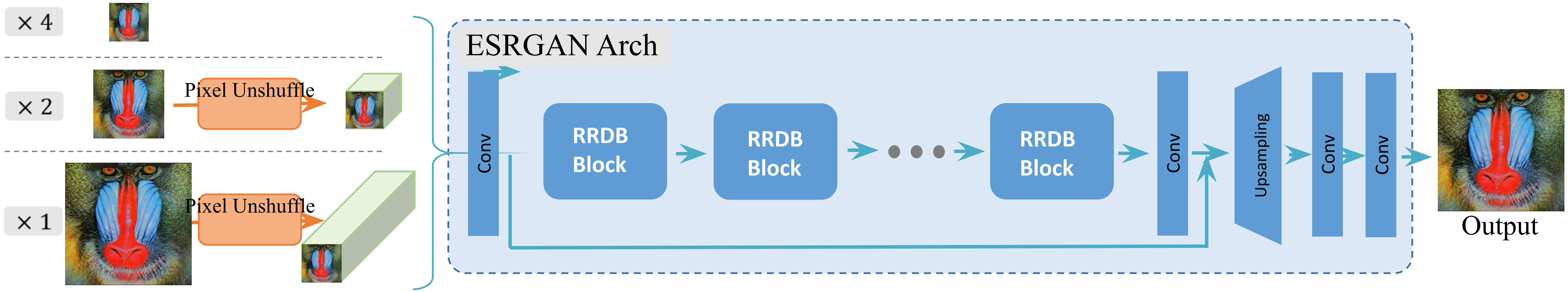
$\omega_c = \pi/5$

$\omega_c = \pi/3$

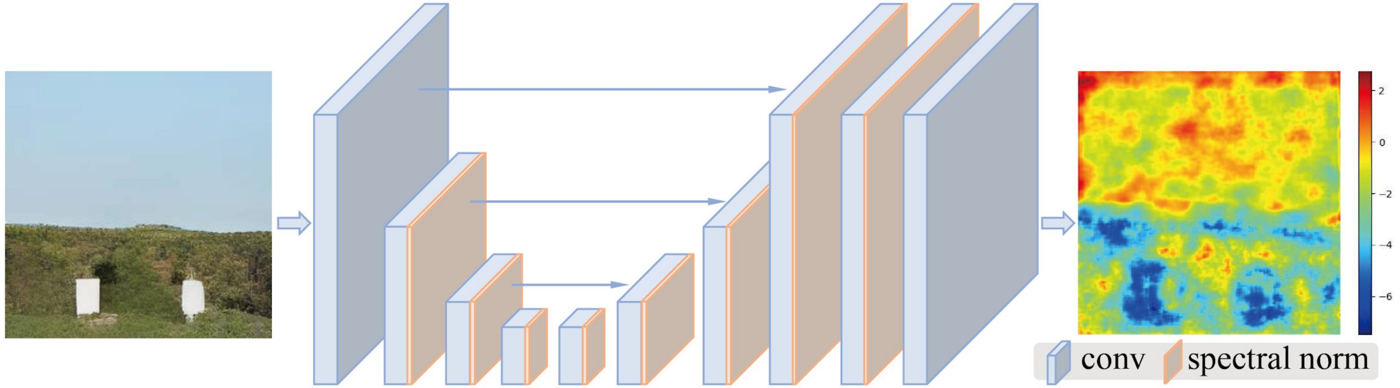
Real-ESRGAN Architecture

ESRGAN → Real-ESRGAN

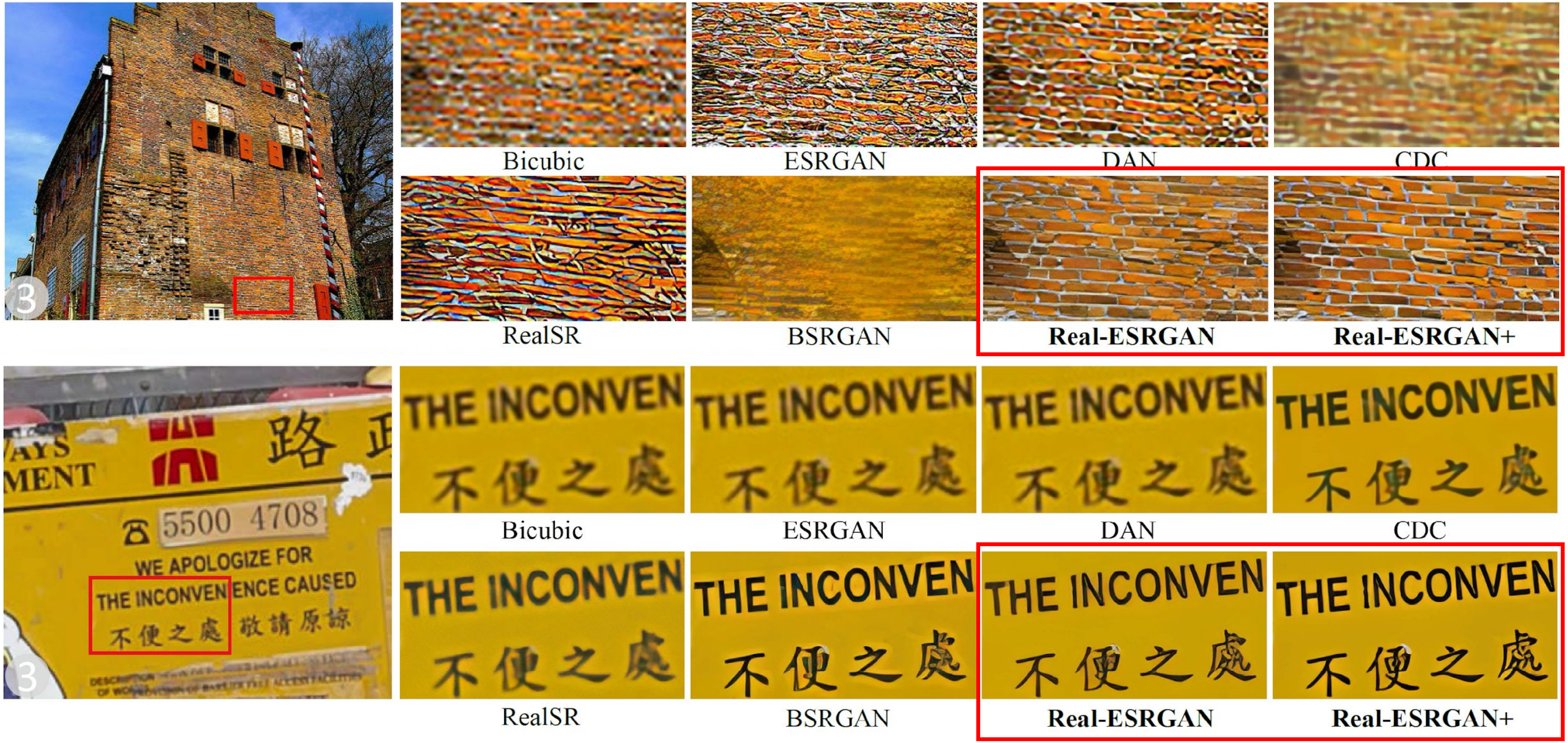
Generator



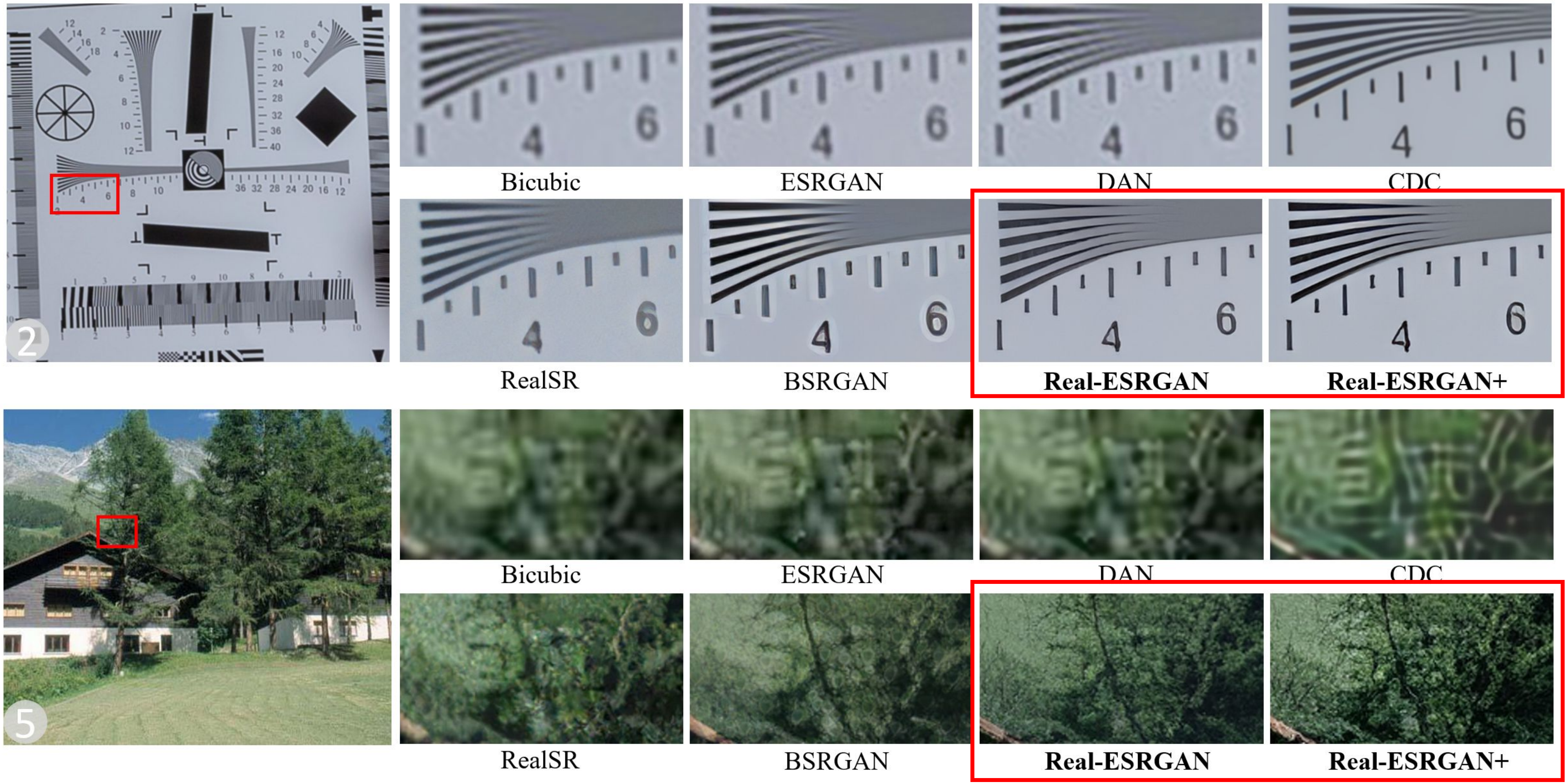
Discriminator



Qualitative Comparisons



Qualitative Comparisons



Quantitative Comparisons

- We provide a non-reference image quality assessment – NIQE *for reference.*
- Existing metrics for perceptual quality cannot well reflect the actual human perceptual preferences on the fine-grained scale.
- Though our Real-ESRGAN+ does not optimize for NIQE scores, it still produces lower NIQE scores on most testing datasets.

Table 1: NIQE scores on several diverse testing datasets with real-world images. The lower, the better.

	Bicubic	ESRGAN [10]	DAN [7]	RealSR [5]	CDC [11]	BSRGAN [12]	**Real-ESRGAN**	**Real-ESRGAN+**
RealSR-Canon [2]	6.1269	6.7715	6.5282	6.8692	6.1488	5.7489	4.5899	**4.5314**
RealSR-Nikon [2]	6.3607	6.7480	6.6063	6.7390	6.3265	5.9920	5.0753	**5.0247**
DRealSR [11]	6.5766	8.6335	7.0720	7.7213	6.6359	6.1362	4.9796	**4.8458**
DPED-iphone [4]	6.0121	5.7363	6.1414	5.5855	6.2738	5.9906	5.4352	**5.2631**
OST300 [9]	4.4440	3.5245	5.0232	4.5715	4.7441	4.1662	2.8659	**2.8191**
ImageNet val [3]	7.4985	**3.6474**	6.0932	3.8303	7.0441	4.3528	4.8580	4.6448
ADE20K val [13]	7.5239	3.6905	6.3839	**3.4102**	6.9219	3.9434	3.7886	3.5778

Qualitative Comparisons with Sliding Bar

Optimize for Anime Images

Qualitative Comparison with Sliding Bar for Amine Images

Naruto: 891x469 3564x1876

Qualitative Comparison with Sliding Bar for Amine Images

Limitations

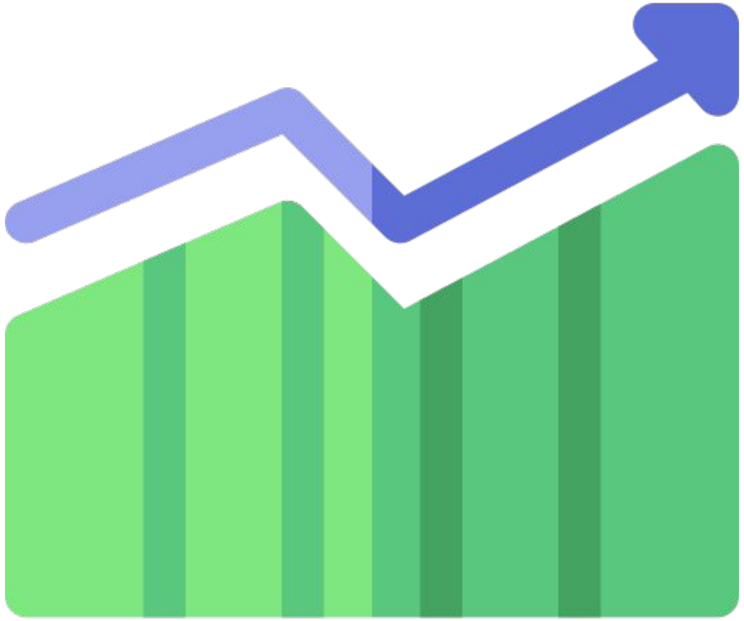
- Twisted lines
- Unpleasant artifacts caused by GAN training
- Unknown and out-of-distribution degradations

Beginnings of Practical Restoration

- Real-ESRGAN aims at developing practical algorithms for general image restoration

- JI Xiaozhong *etc.*
  - *RealSR:* Real-World Super-Resolution via Kernel Estimation and Noise Injection
- ZHANG Kai *etc.*
  - *BSRGAN*: Designing a Practical Degradation Model for Deep Blind Image Super-Resolution
- LIANG Jinyun *etc*.
  - *SwinIR*: Image Restoration Using Swin Transformer

- ……



Open Source

GitHub

- Full training and testing codes
- **Colab Demo** for Real-ESRGAN Open in Colab
- Portable **Windows / Linux / MacOS executable files** for Intel/AMD/Nvidia GPU, which is based on Tencent ncnn

BasicSR Codes & Models

We also incorporate the face restoration method – **GFPGAN**, to improve the face performance

Thanks for Watching

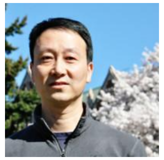
AIM 2021

Real-ESRGAN: Training *Real-World* Blind Super-Resolution with *Pure Synthetic Data*


Xintao Wang


Liangbin Xie


Chao Dong


Ying Shan

BasicSR

Codes & Models