Performance Assessment of Image Super Resolution and Denoising Based on Generative Adversarial Network
評估基於對抗生成網路的影像超解析度與去雜訊之表現

Yung-Chun, Chen (陳詠君)
Advisor: Shyh-Hau, Wang (王士豪 教授)

Department of Computer Science and Information Engineering,
National Cheng Kung University

June 5, 2022

OUTLINE

- Introduction
- Materials & Methods
- Results and Discussion
- DEMO
- Conclusion

INTRODUCTION

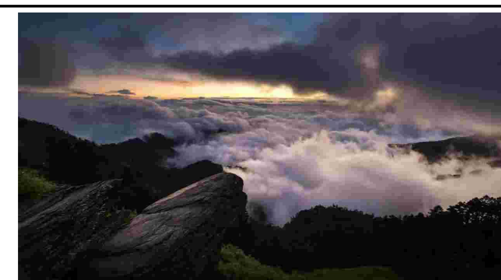
- Why GAN?
  - Ledig et al [1] had proved that Generative Adversarial Network(GAN) has better performance than CNN model in the field of super resolution, so I will use three different Gan-based models to super resolve several images to compare their performance in my project.

[1] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, ´Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi Twitter, “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, 2017.

INTRODUCTION

- Motivation

Image we want :

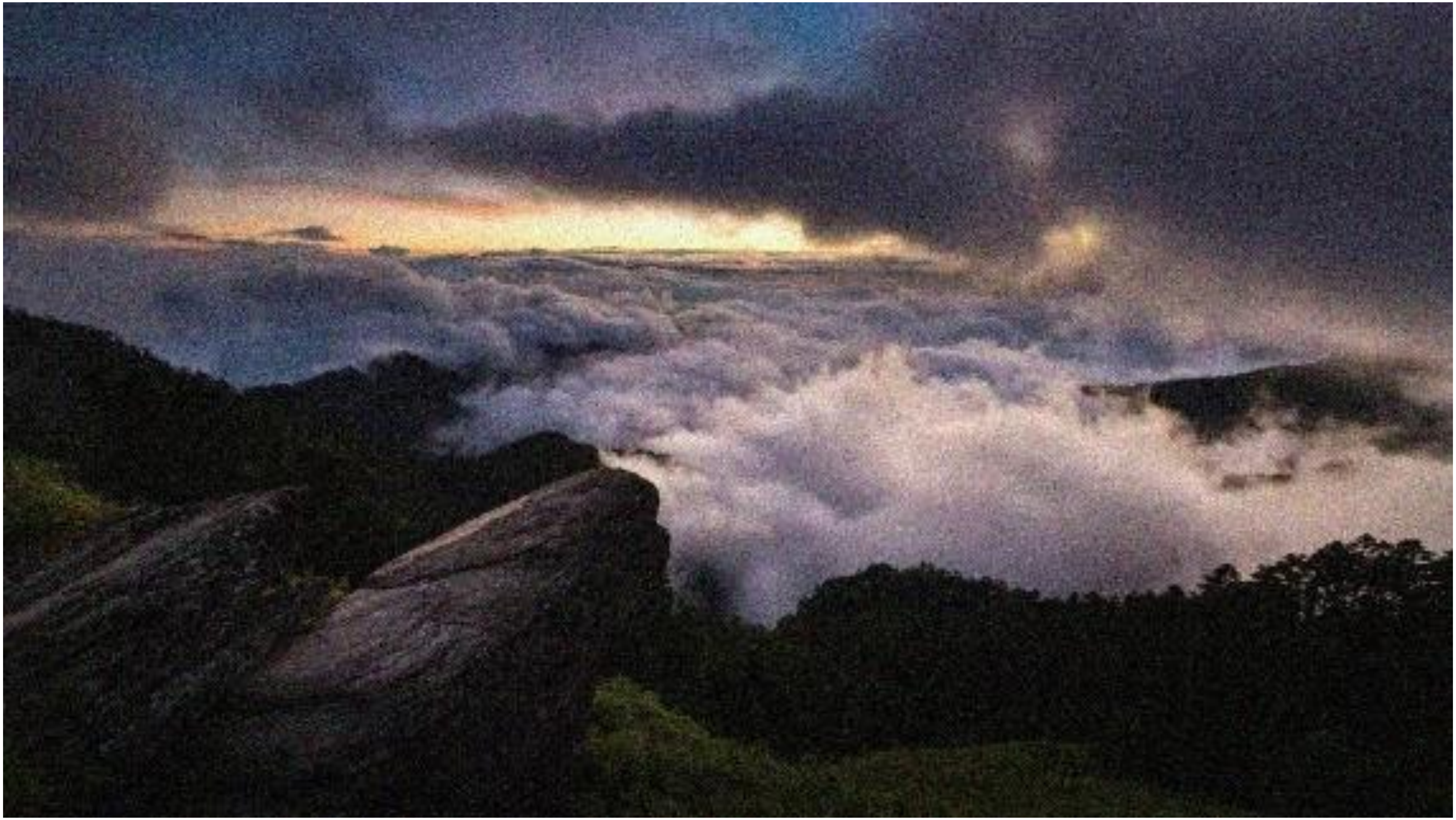
BLUR	
NOISE	
RESIZE	
JPEG COMPRESSION	

INTRODUCTION

- Blur

INTRODUCTION

- Noise

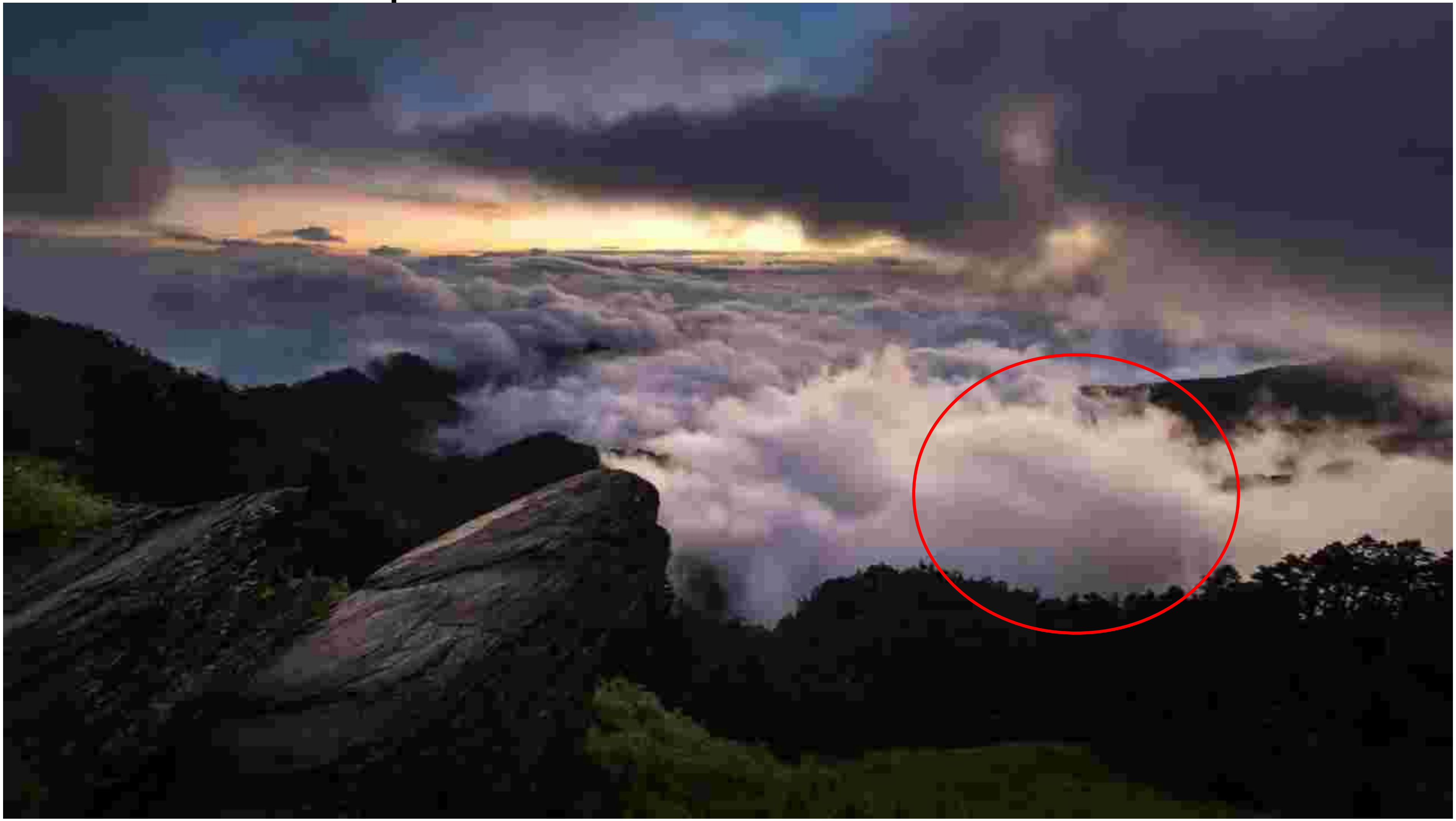


INTRODUCTION

- Resize

INTRODUCTION

- JPEG Compression



INTRODUCTION

- In my project, I will use three model to test their performance on the work of super resolving images, that is:

  1. SRGAN [1],
  2. ESRGAN (focus on deep neural network approaches) [2], and
  3. Real-ESRGAN (focus on image preprocessing approaches) [3].

[1] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, ´Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi Twitter, “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, 2017.

[2] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks”, 2018.

[3] Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan. ”Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data”, 2021.

INTRODUCTION

- Objective
  - Super resolve the images which have poor resolution and bad quality in the real-world scenarios with the variety of real-world training datasets.

INTRODUCTION

Degradation：

◎ $X = D(y) = [(y \oplus k)\downarrow + n]_{JPEG}$

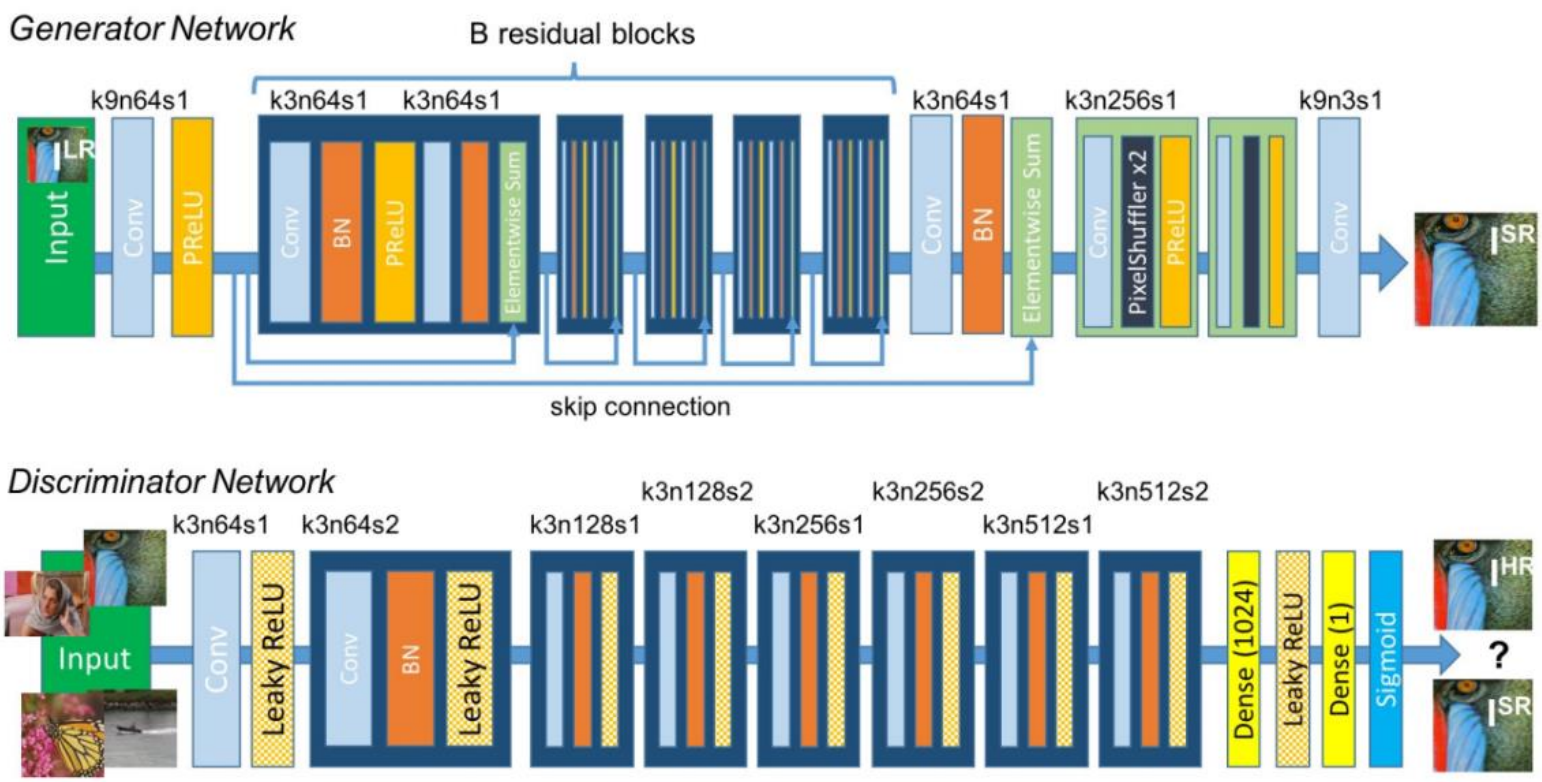
X : low resolution image
y : high resolution image
D : degradation
$\oplus$: convolution

k : blur kernel
$\downarrow$: down-sampling
n : noise
JPEG : JPEG compression

Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan. "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data", 2021.

INTRODUCTION

- SRGAN Architecture



Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, ´Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi Twitter, “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, 2017.

INTRODUCTION

- SRGAN loss function: solve the min-max problem consists of two parts:

  (1) minimize generator loss

  (2) maximize discriminator loss

$$\boxed{\min_{\theta_G} \max_{\theta_D} E_{I^{HR} \sim p_{train}(\mathrm{I}^{\mathrm{HR}})}\left[\log D_{\theta_D}(I^{HR})\right] + E_{I^{LR} \sim p_G(I^{LR})}[\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR})))]}$$

INTRODUCTION

- Perceptual loss (compare whether two different images look similar): The perceptual loss function consists of two parts: (1) the content loss $I^{SR}_{\frac{VGG}{i,j}}$ and (2) the adversarial loss I^{SR}_{Gen}.

$$I^{SR} = \underset{\text{content loss}}{\boxed{I^{SR}_{VGG}}} + \underset{\text{adversarial loss}}{\boxed{10^{-3} I^{SR}_{Gen}}}$$

(1) Content loss

$$I^{SR}_{\frac{VGG}{i,j}} = 1/(W_{i,j}H_{i,j}) \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\emptyset_{i,j}(I^{HR})_{x,y} - \emptyset_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

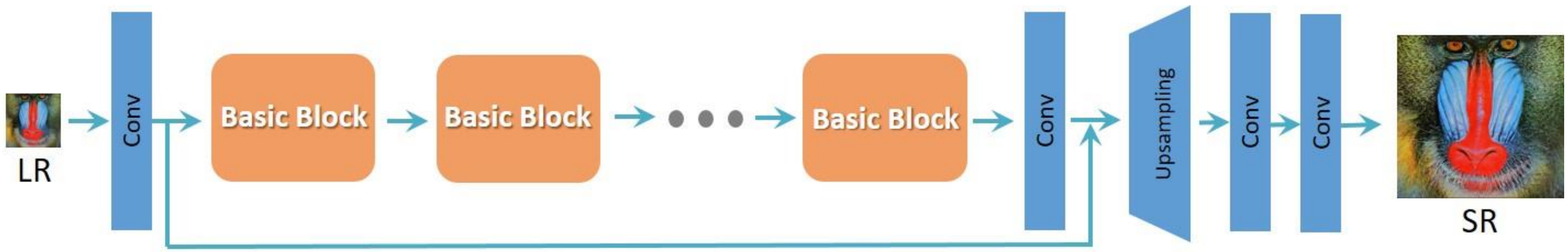
Restored image

(2) Adversarial loss

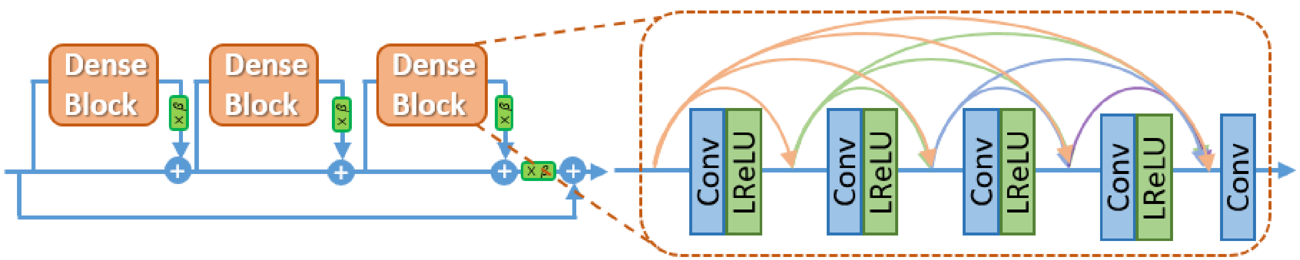
$$\mathrm{I}^{SR}_{Gen} = \sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

INTRODUCTION

- ESRGAN Architecture



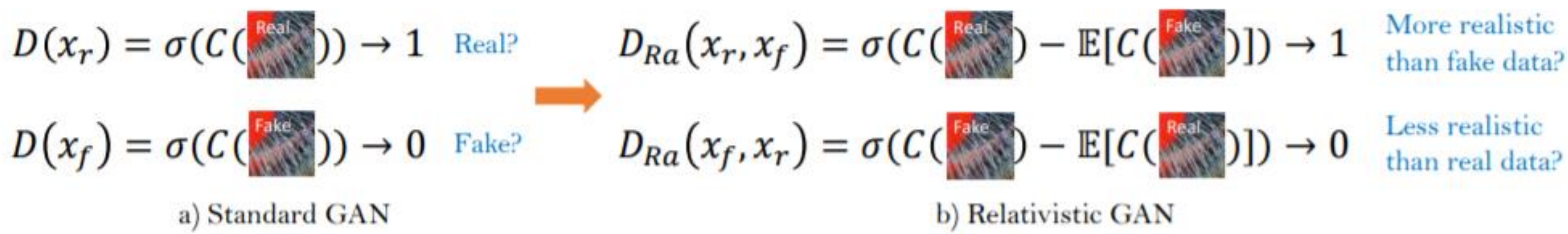
Residual in Residual Dense Block (RRDB)



Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks”, 2018.

INTRODUCTION

- Discriminator tries to predict the probability that a real image x_r is relatively more realistic than a fake one x_f.



- Improve the perceptual loss by using the features before activation, which could provide stronger supervision for brightness consistency and texture recovery.

Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks”, 2018.

INTRODUCTION

- Loss function

Discriminator loss

$$L_D^{Ra} = -E_{x_r}\left[\log\left(D_{Ra}\left(x_r, x_f\right)\right)\right] - E_{x_f}[\log(1 - D_{R_a}(x_f, x_r))]$$

Generator loss

$$L_G = L_{percep} + \lambda L_G^{Ra} + \eta L_1$$

Adversarial loss

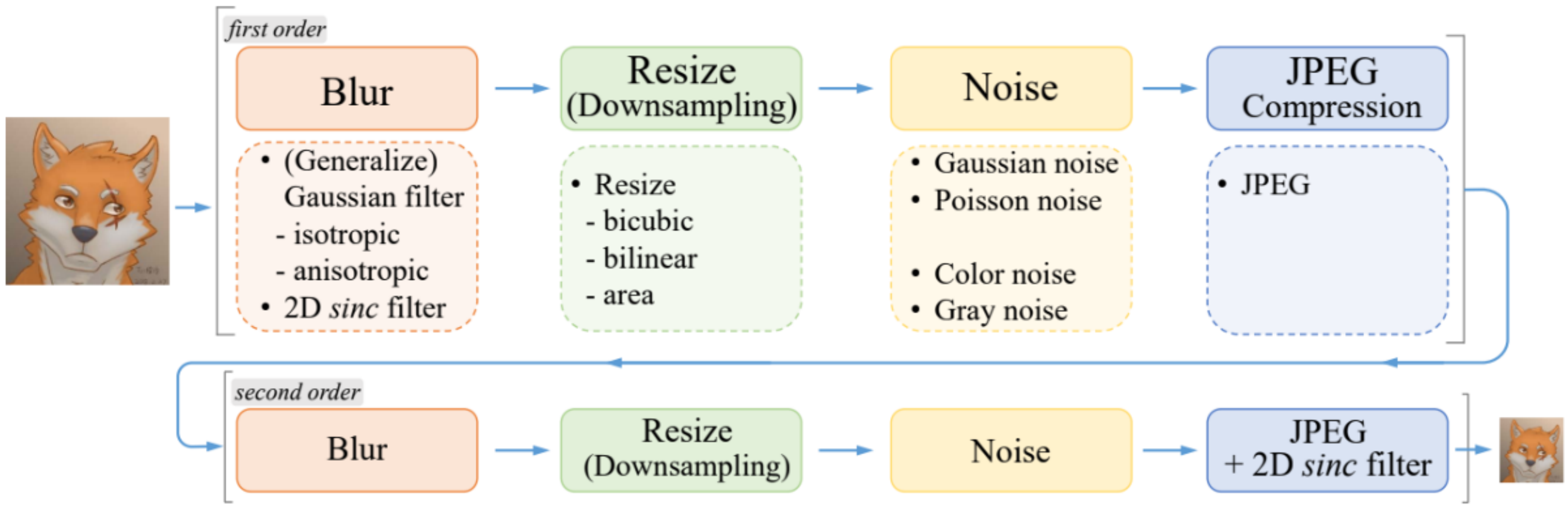
$$L_G^{Ra} = -E_{x_r}\left[\log\left(1 - D_{Ra}\left(x_r, x_f\right)\right)\right] - E_{x_f}[\log(D_{Ra}(x_f, x_r))]$$

Content loss(L_1)

$$L_1 = E_{x_i}\|G(x_i) - y\|_1$$

INTRODUCTION

- Real-ESRGAN
  - Training Process：Training the model with input images that are originally HR images, then degraded into LR images with the combinations of blur, resize, noise, JPEG compression two times.

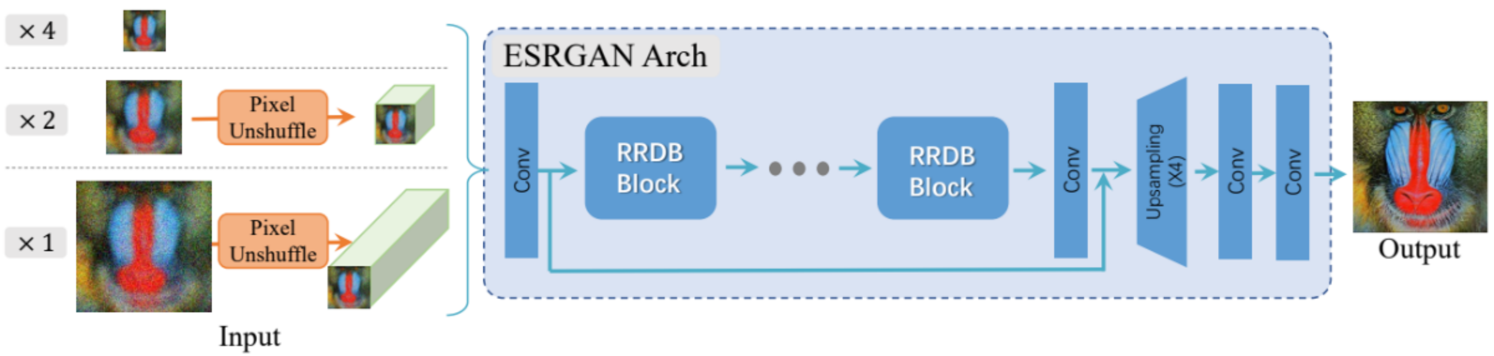


Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan. "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data", 2021.

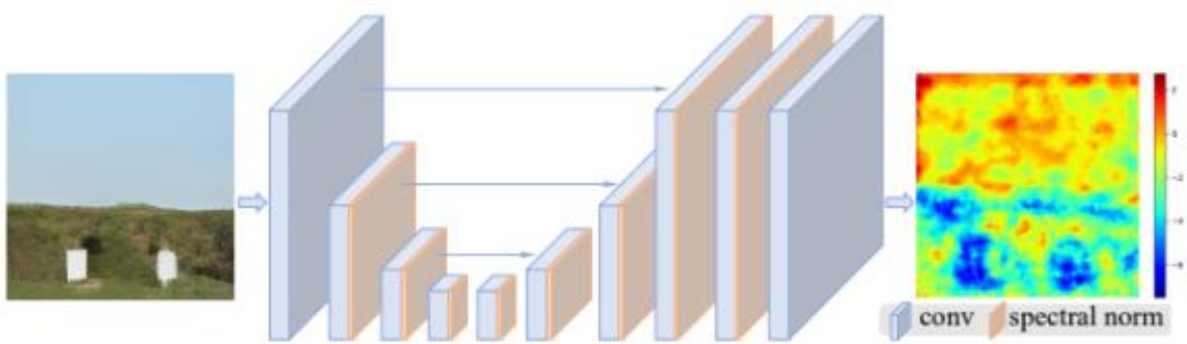
INTRODUCTION

- Method
  - Training Process

generator



discriminator



Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan. "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data", 2021.

MATERIALS & METHODS

- Development Tools & Environment
  - Environment：Windows 10, Matlab (for calculate NIQE)
  - Language：Python3.7
  - Package：Pytorch, Tensorflow, opencv, PyQt5

MATERIALS & METHODS

Datasets

1. DIV2K
(bicubic : x2,x3,x4,x8
unknown : x2,x3,x4)
-Train data : 800 HR/LR images
-Validation data : 100 HR/LR images
-Test data : 100 HR/LR images

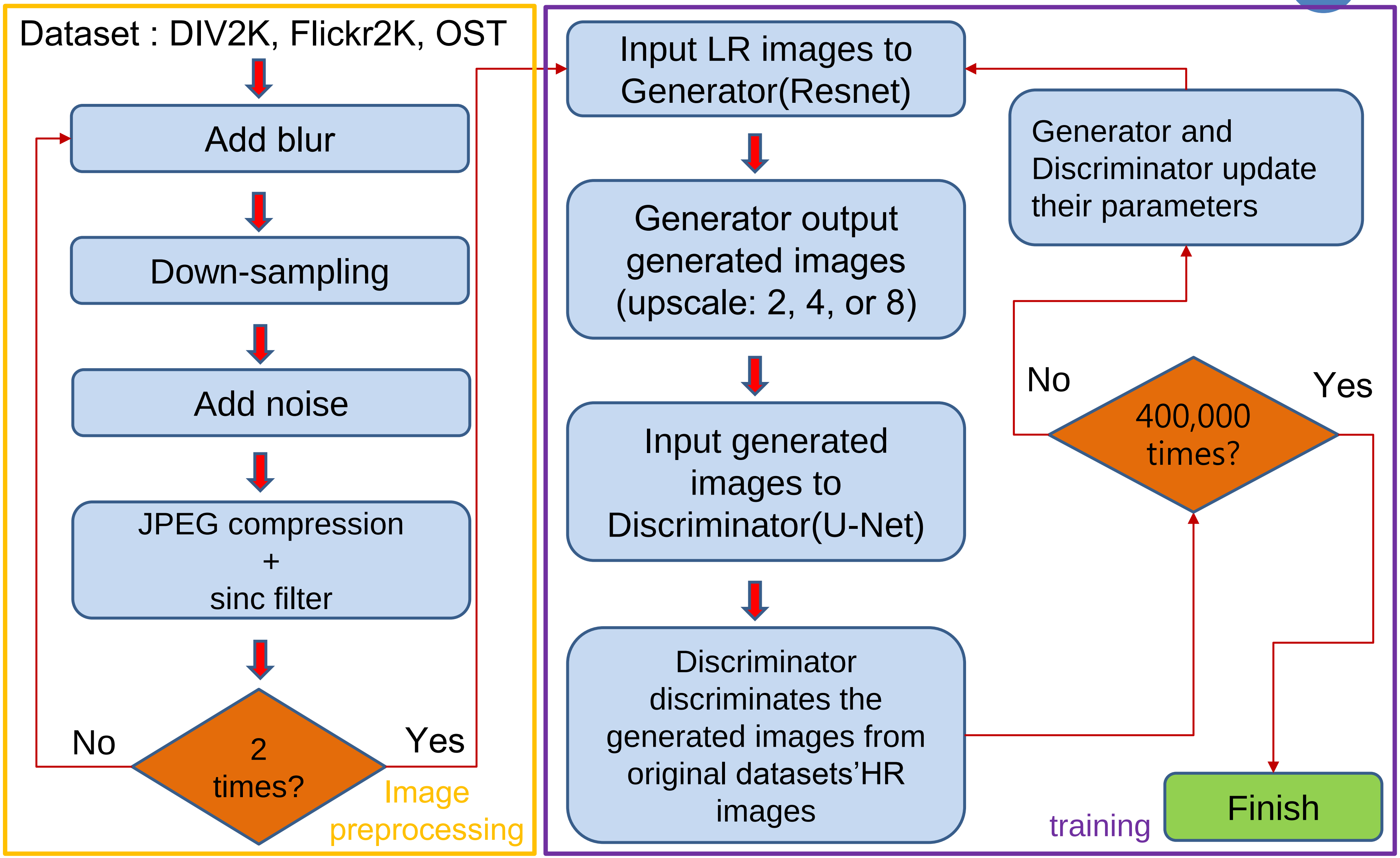
2. Flickr2K
(bicubic : x2,x3,x4
unknown : x2,x3,x4)
-2650 HR/LR images

3. OST
(Scale and Degradation : decide by yourself)
-train data : 10,324 HR images
-test data : 300 HR images

High resolution image

low resolution image counterpart

MATERIALS & METHODS



RESULTS & DISCUSSION

- original image → resized + noised image using bilinear interpolation (scale: 4) and gaussian noise.

ORIGINAL (1280 * 720)

bilinear interpolation
+
Gaussian noise

LOW RESOLUTION (320 * 180)

RESULTS & DISCUSSION

- restore LR image to SR image using SRGAN, ESRGAN, and Real-ESRGAN (scale: 4)

ORIGINAL (1280 * 720)

SRGAN (1280 * 720)

ESRGAN (1280 * 720)

Real-ESRGAN (1280 * 720)

	SRGAN	ESRGAN	Real-ESRGAN
PSNR	19.0711	15.0668	28.6179
SSIM	0.1715	0.0710	0.7967
LPIPS	0.448	1.007	0.269
NIQE	5.5502	7.0807	2.9377
Time Cost (s)	19.9649	19.4307	77.8971

RESULTS & DISCUSSION

- original image → resized + blur image using nearest neighbor interpolation (scale: 4) and average filter(5 * 5).

ORIGINAL (1280 * 720)

Nearest neighbor
interpolation
+
Average filter(5 * 5)

LOW RESOLUTION (320 * 180)

RESULTS & DISCUSSION

- restore LR image to SR image using SRGAN, ESRGAN, and Real-ESRGAN (scale: 4)

ORIGINAL (1280 * 720)

SRGAN (1280 * 720)

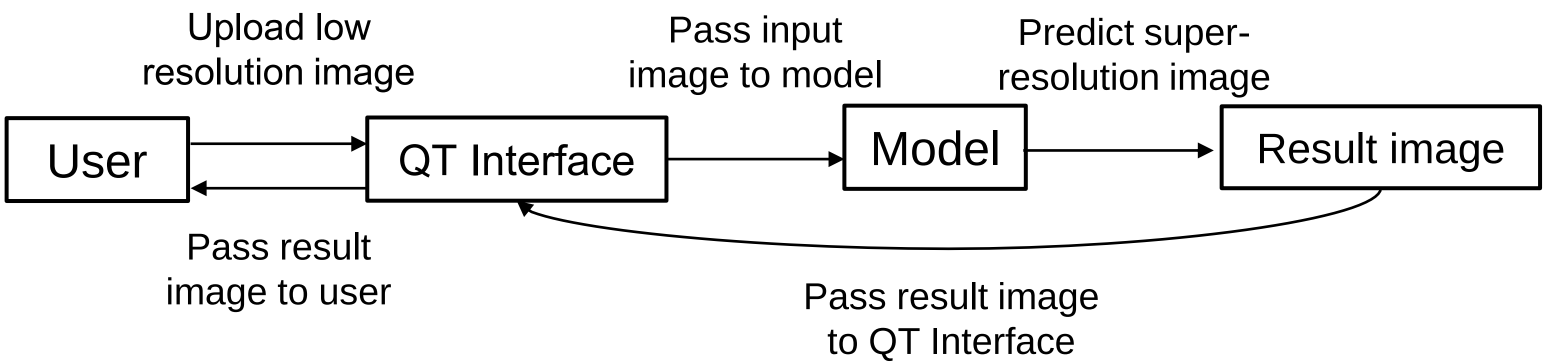
ESRGAN (1280 * 720)

Real-ESRGAN (1280 * 720)

	SRGAN	ESRGAN	Real-ESRGAN
PSNR	19.0711	28.6165	27.6920
SSIM	0. 1715	0.8041	0.7625
LPIPS	0.892	0.459	0.276
NIQE	5.8473	4.8999	3.0430
Time Cost (s)	23.1587	18.6940	72.9102

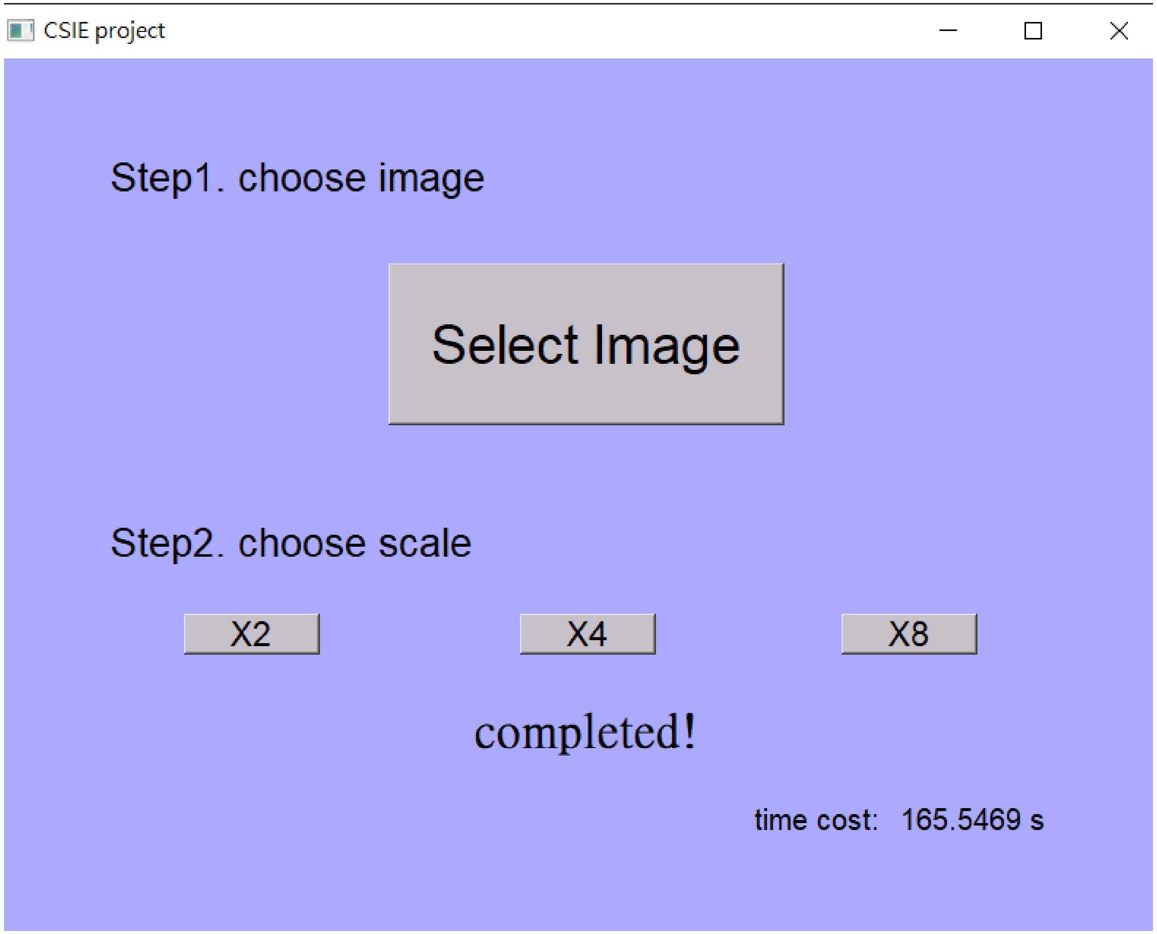
DEMO

System architecture:



DEMO

- QT Interface



DEMO

OST test data, unknown noise, using Real-ESRGAN

LR input (512 * 680)

Output (2048 * 2720)

Time：220.3236 s

NIQE：1.9948

DEMO

My image, unknown noise, using Real-ESRGAN

LR input (401 * 500)

Output (1604 * 2000)

Time：165.5469 s

NIQE：1.5731

CONCLUSION

- Both subjectively and objectively, Real-ESRAGN can remove the unknown degradation factor and perform well in comparison to others in the real-world super-resolution task.

- Image preprocessing approaches model (Real-ESRGAN) can perform well in removing unknown noise than deep neural network approaches model (ESRGAN).

- However, it looks like synthesis picture, not looks like the real photo because Real-ESRGAN is trained with pure synthesis input images.

THANK YOU FOR YOUR ATTENTION