Image Super Resolution

Group 2

He Huang, Yixi Liang, Ruiqi Li

CONTENTS

/01 Introduction

Introduction

- Image super resolution: Low-resolution -> High-resolution
- Traditional Methods
- Neural Networks:
  - CNN
  - Autoencoder
  - GAN
- TensorFlow Keras
- Dataset: https://www.kaggle.com/datasets/quadeer15sh/image-super-resolution-from-unsplash

Example Image (Data)

/02 Traditional Method

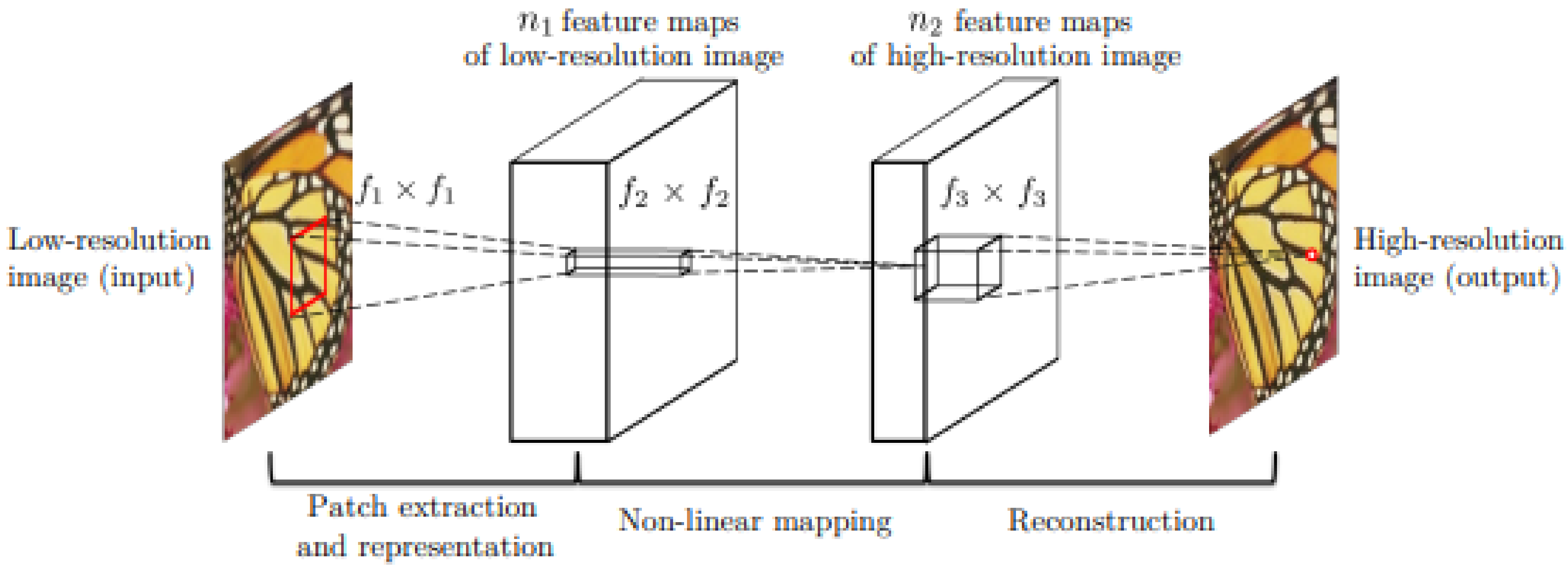
Traditional Methods

- Set rules to fill up missing values of pixels
- From CV2 Package
- Inter Nearest: Fill up missing value with same value of its nearest pixel
- Inter Linear: Uses value of ~4 pixels to fill up the missing

original inter nearest inter linear

/03 CNN Model

SRCNN



Dong, Chao, et al. "Image Super-Resolution Using Deep Convolutional Networks." *ArXiv.org*, 31 July 2015, https://arxiv.org/abs/1501.00092.

1st part: Extraction Representation

$$F_1(\mathbf{Y}) = \max\left(0, W_1 * \mathbf{Y} + B_1\right)$$

- n1 times of convolution operations with kernel of c*f1*f1
- Y: input image
- W1: filters
- B1: biases
- C: number of channels
- *: Convolutional Operation
- ReLU form output

2nd part: Non-Linear Mapping

$$F_2(\mathbf{Y}) = \max\left(0, W_2 * F_1(\mathbf{Y}) + B_2\right)$$

- Transforming n1-dimensional vectors to n2-dimensional vectors
- n2 times of convolution operations with kernel of n1*f2*f2
- W2: n2 filters of n1 * f2 * f2
- B2: n2 dimensions
- ReLU form output
-

3rd part: Reconstruction

$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3$$

- Reconstructing Image
- n2 times of convolution operations with kernel of n1*f2*f2
- W3: c filters of n2 * f3 * f3
- B3: c dimensions
- Loss function: MSE
- $$L(\Theta) = \frac{1}{n}\sum_{i=1}^{n} ||F(\mathbf{Y}_i;\Theta) - \mathbf{X}_i||^2$$

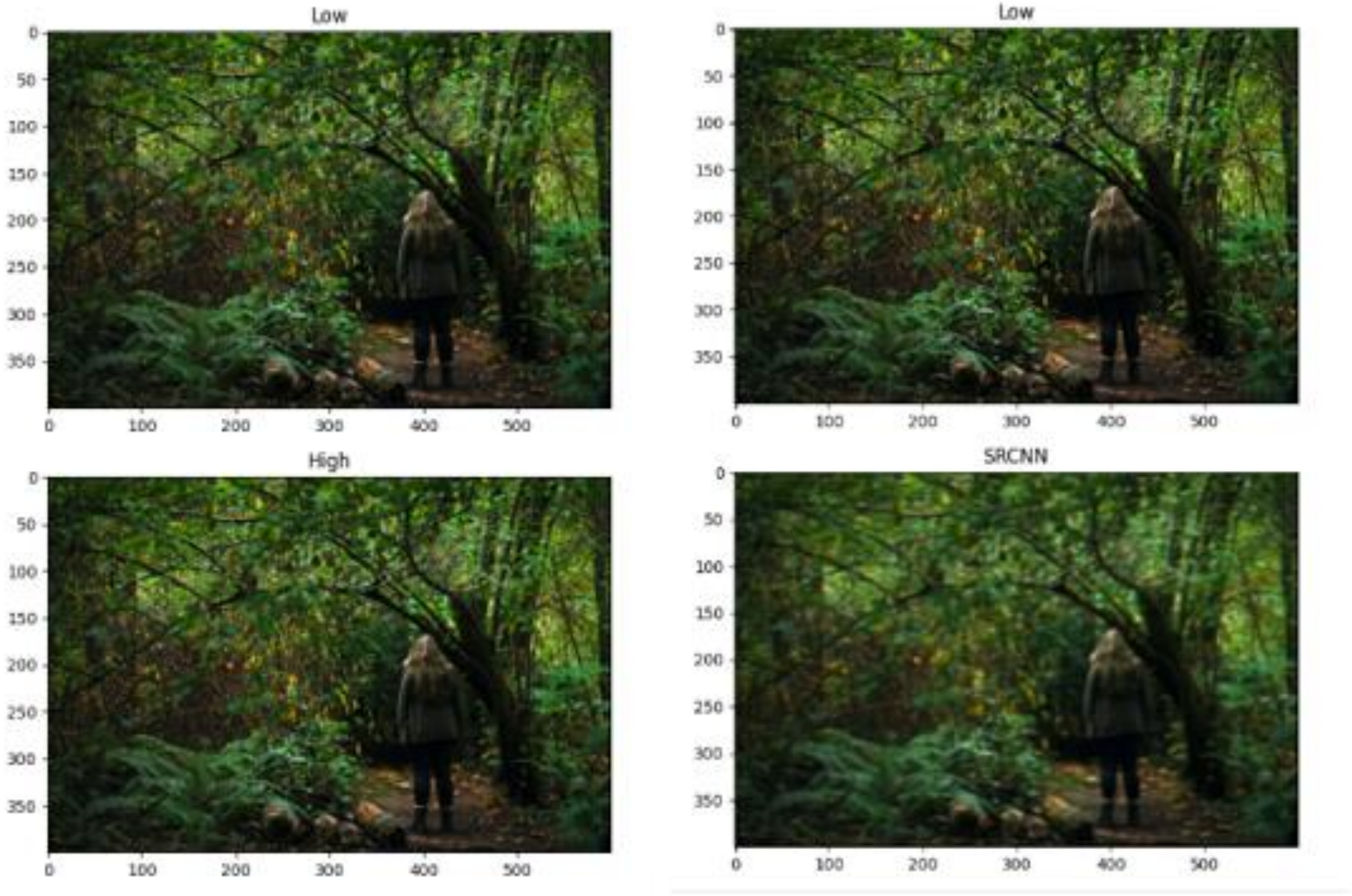
SRCNN Model Info

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, None, None, 64)    15616

 conv2d_1 (Conv2D)           (None, None, None, 32)    2080

 conv2d_2 (Conv2D)           (None, None, None, 3)     2403

=================================================================
Total params: 20,099
Trainable params: 20,099
Non-trainable params: 0
_________________________________________________________________
```

```
1/1 - 2s - loss: 0.0020 - mean_squared_error: 0.0020 - 2s/epoch - 2s/step
1/1 - 0s - loss: 0.0036 - mean_squared_error: 0.0036 - 107ms/epoch - 107ms/step
1/1 - 0s - loss: 0.0015 - mean_squared_error: 0.0015 - 108ms/epoch - 108ms/step
1/1 - 0s - loss: 0.0011 - mean_squared_error: 0.0011 - 97ms/epoch - 97ms/step
1/1 - 0s - loss: 0.0028 - mean_squared_error: 0.0028 - 98ms/epoch - 98ms/step
1/1 - 0s - loss: 0.0019 - mean_squared_error: 0.0019 - 95ms/epoch - 95ms/step
1/1 - 0s - loss: 0.0022 - mean_squared_error: 0.0022 - 96ms/epoch - 96ms/step
1/1 - 0s - loss: 0.0023 - mean_squared_error: 0.0023 - 96ms/epoch - 96ms/step
```

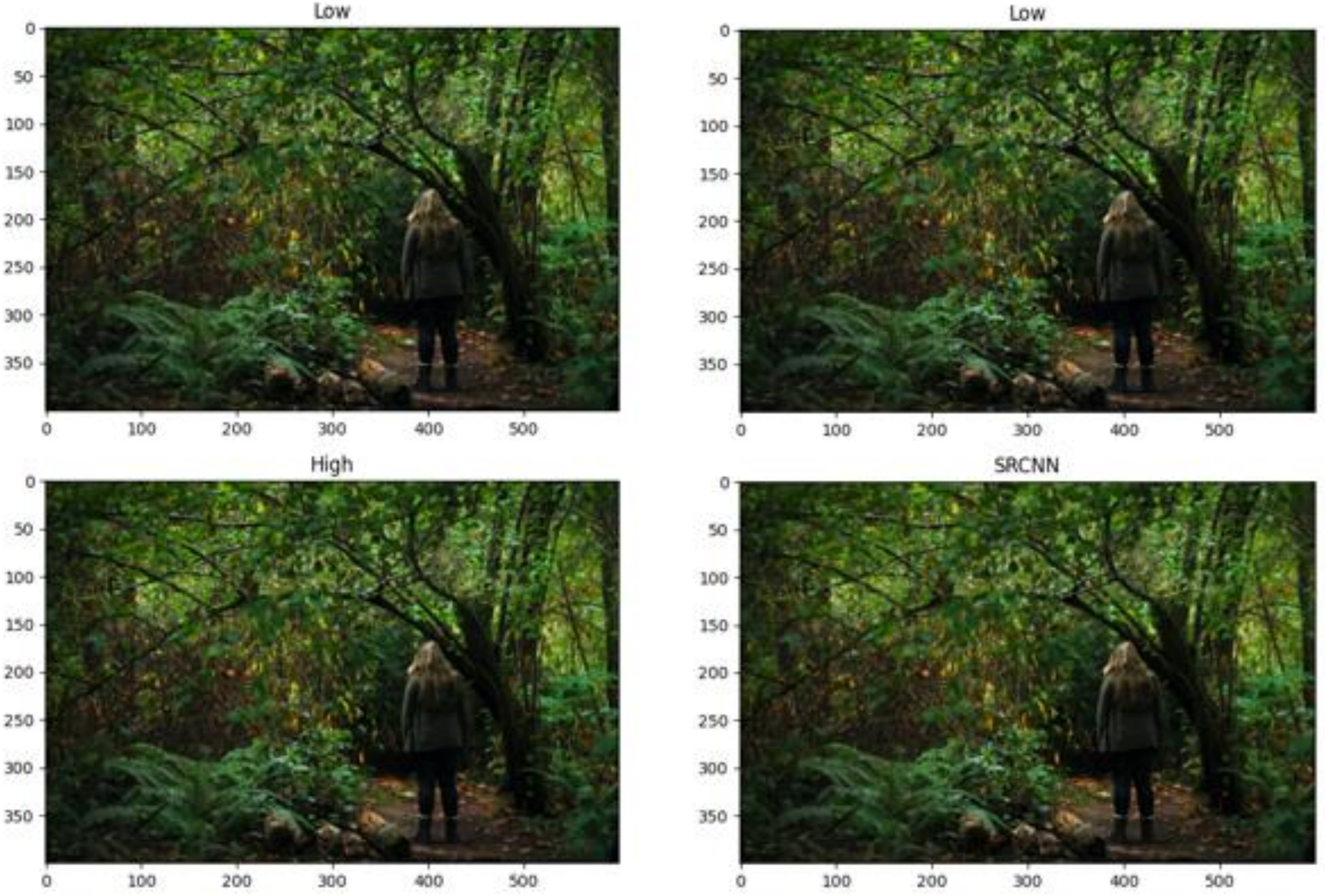
SRCNN Results

Epoch = 1



SRCNN Results

Epoch = 20



/04 Autoencoder

What is Autoencoder?

Autoencoder is an unsupervised artificial neural network that is trained to copy its input to output. In the case of image data, the autoencoder will first encode the image into a lower-dimensional representation, then decodes that representation back to the image. Encoder-Decoder automatically consists of the following two structures:

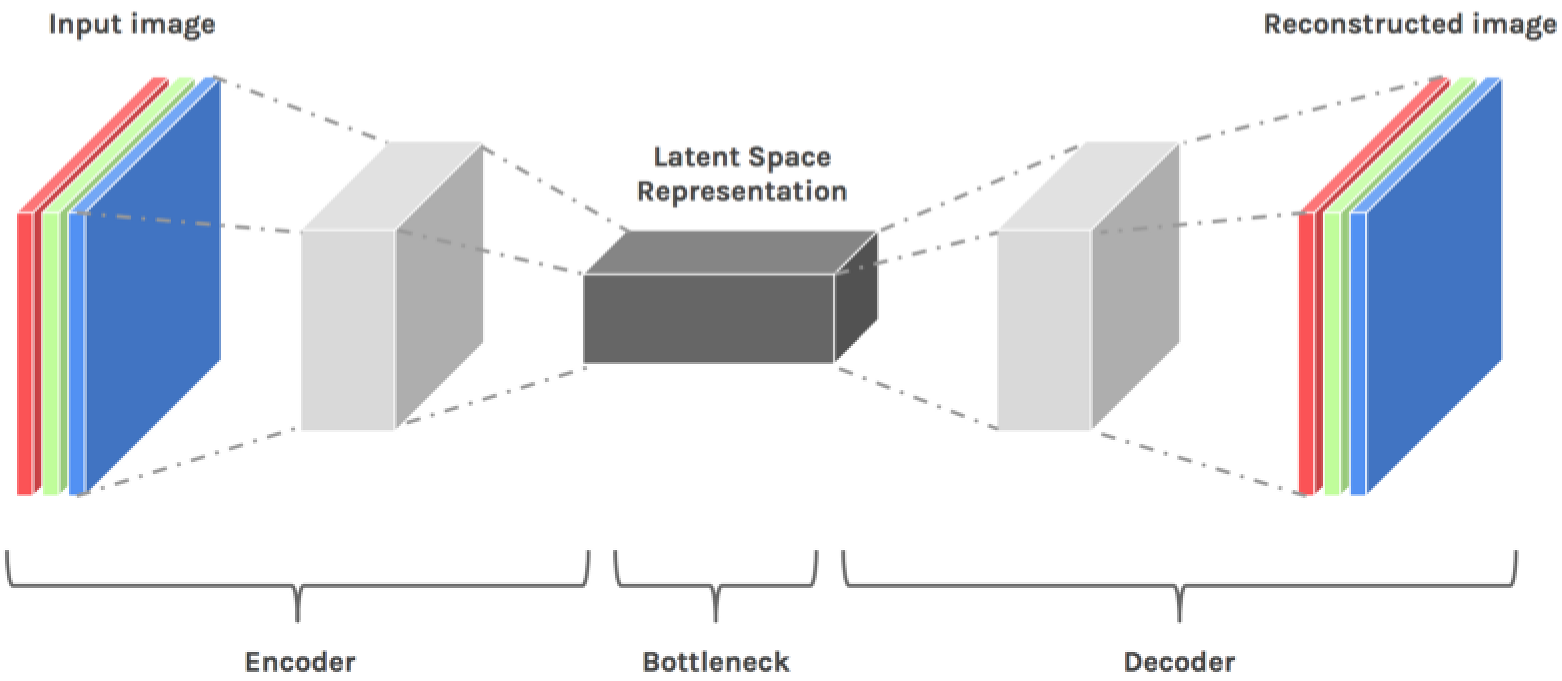
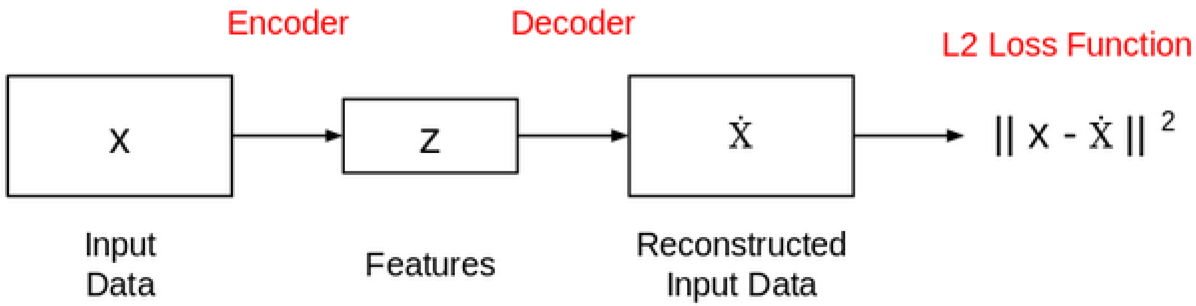
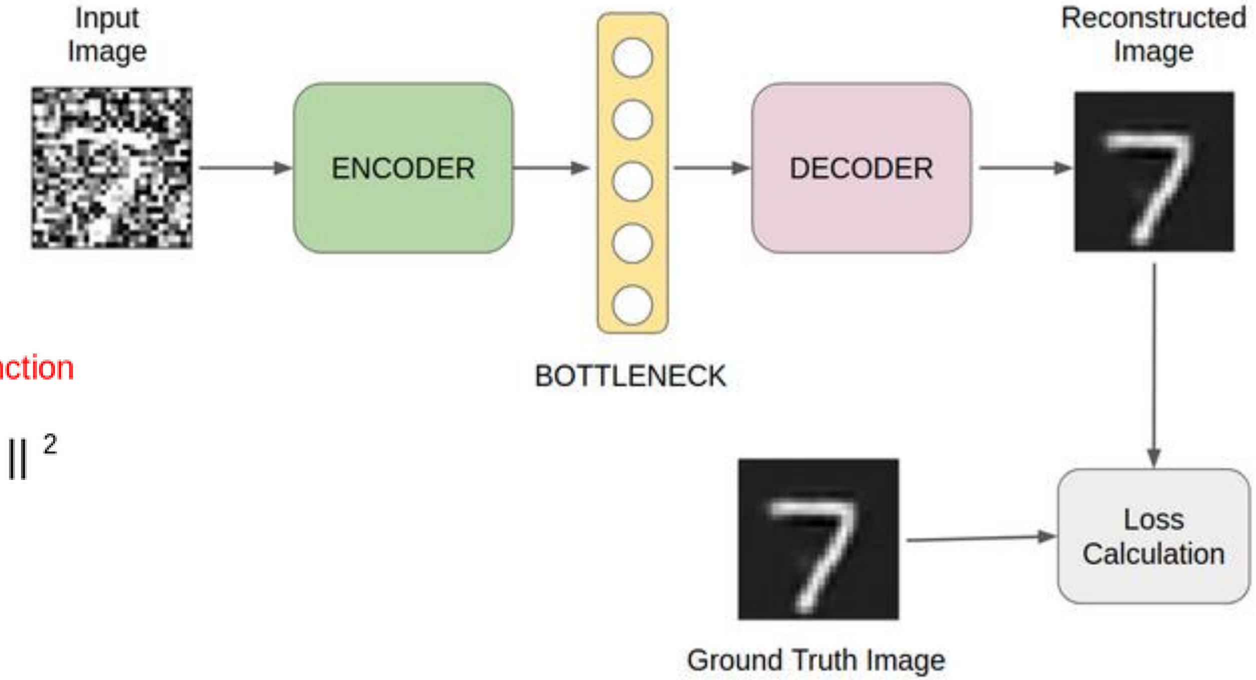
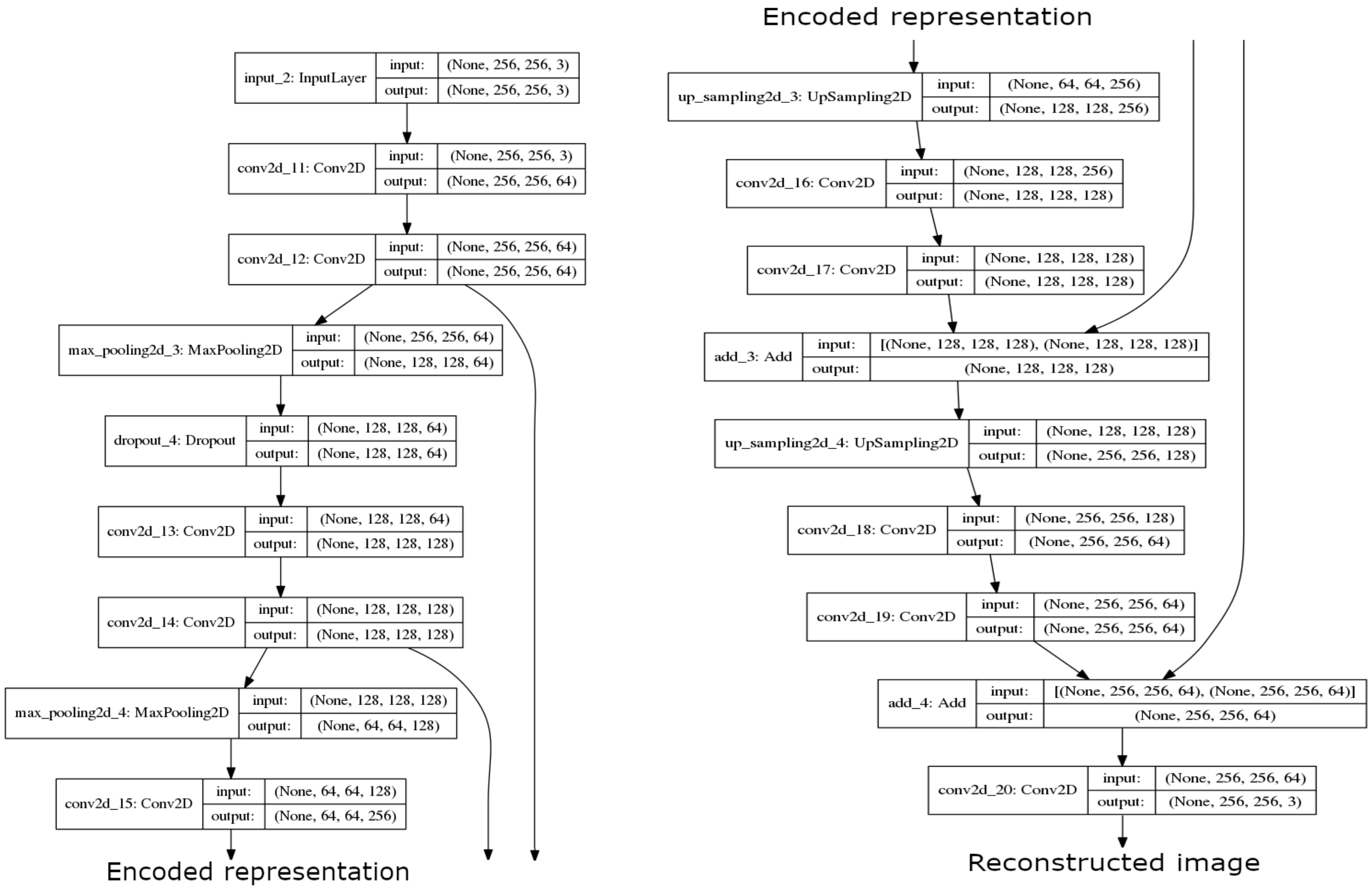


Image Denoising Autoencoder

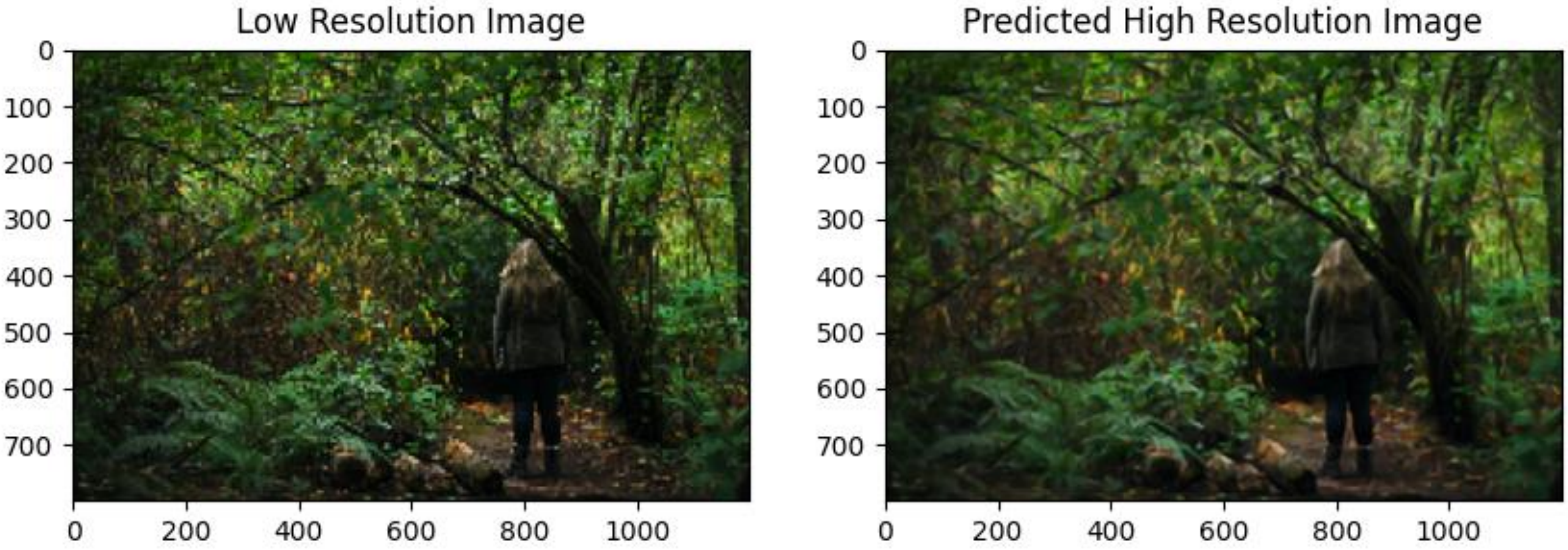




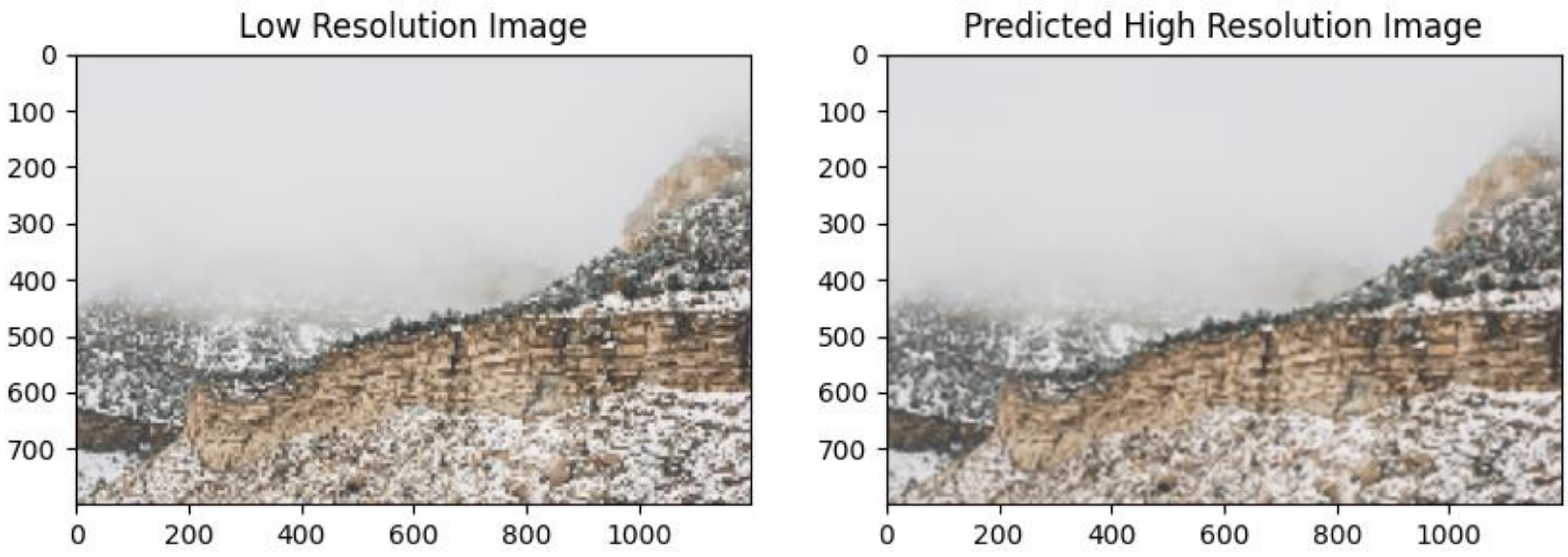
Enhance Image Resolution using Autoencoder



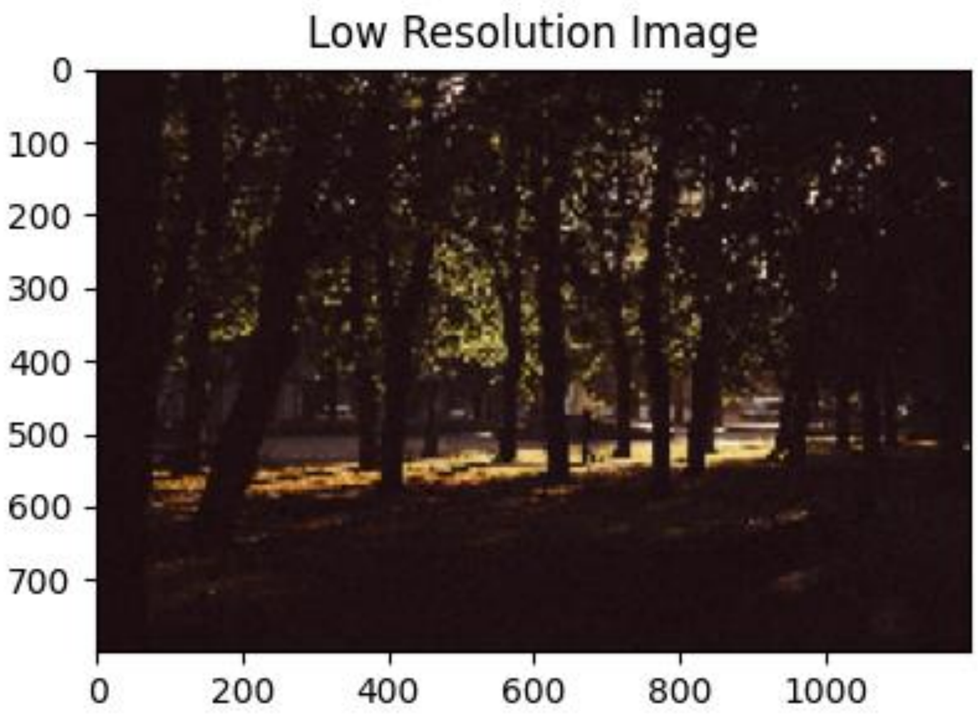
Result of autoencoder

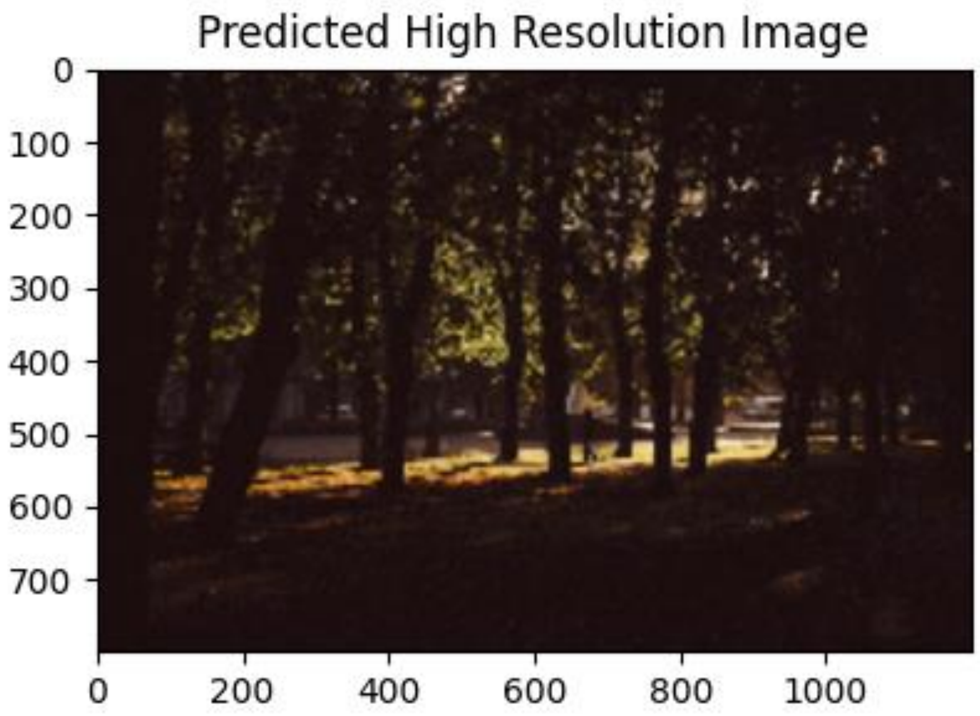


Result of autoencoder



Result of autoencoder



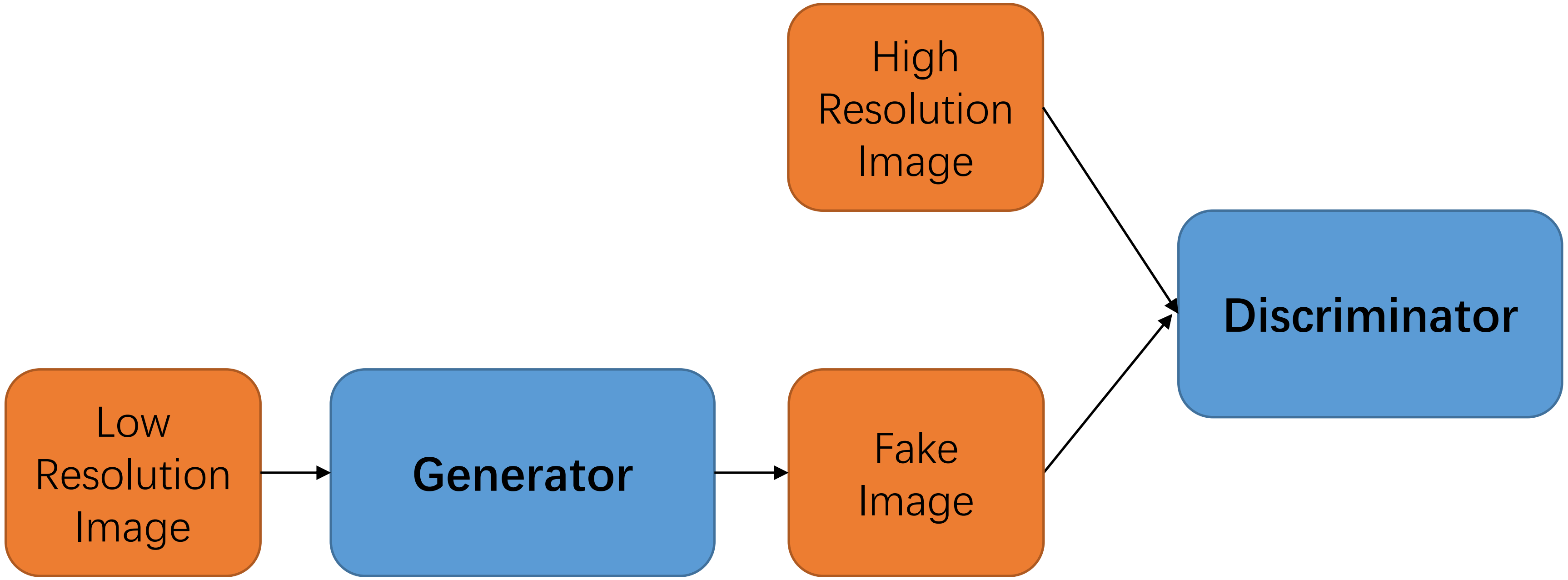


/05

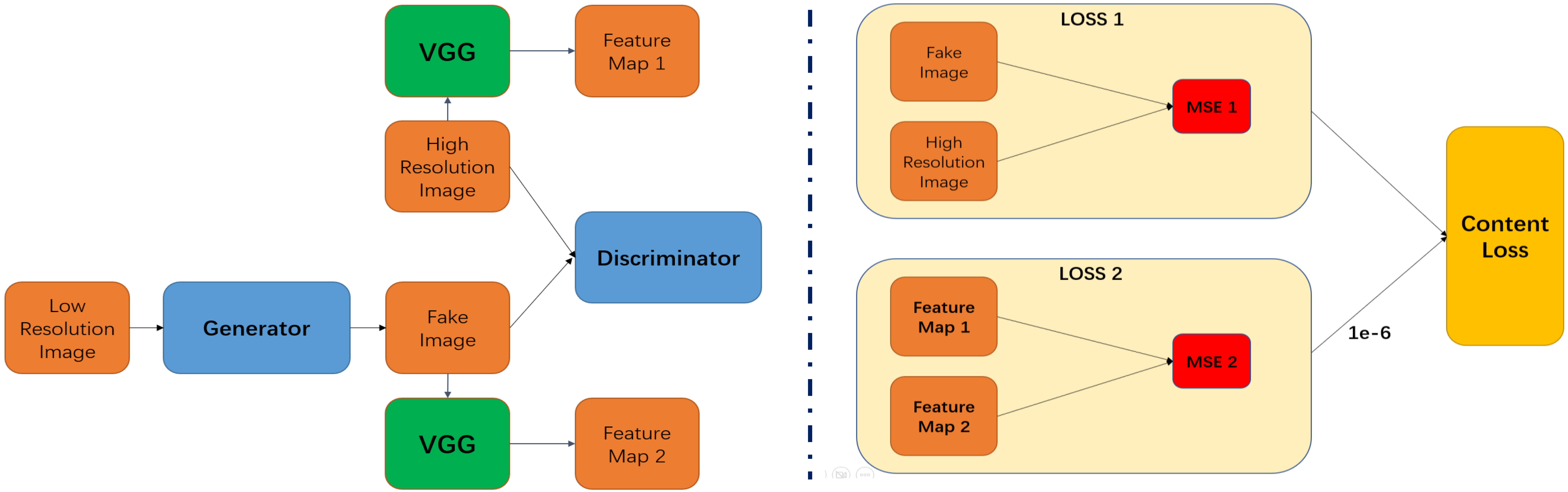
GAN

Generative Adversarial Network

How GAN works in image super resolution



SRGAN(Super Resolution Generative Adversarial Network)



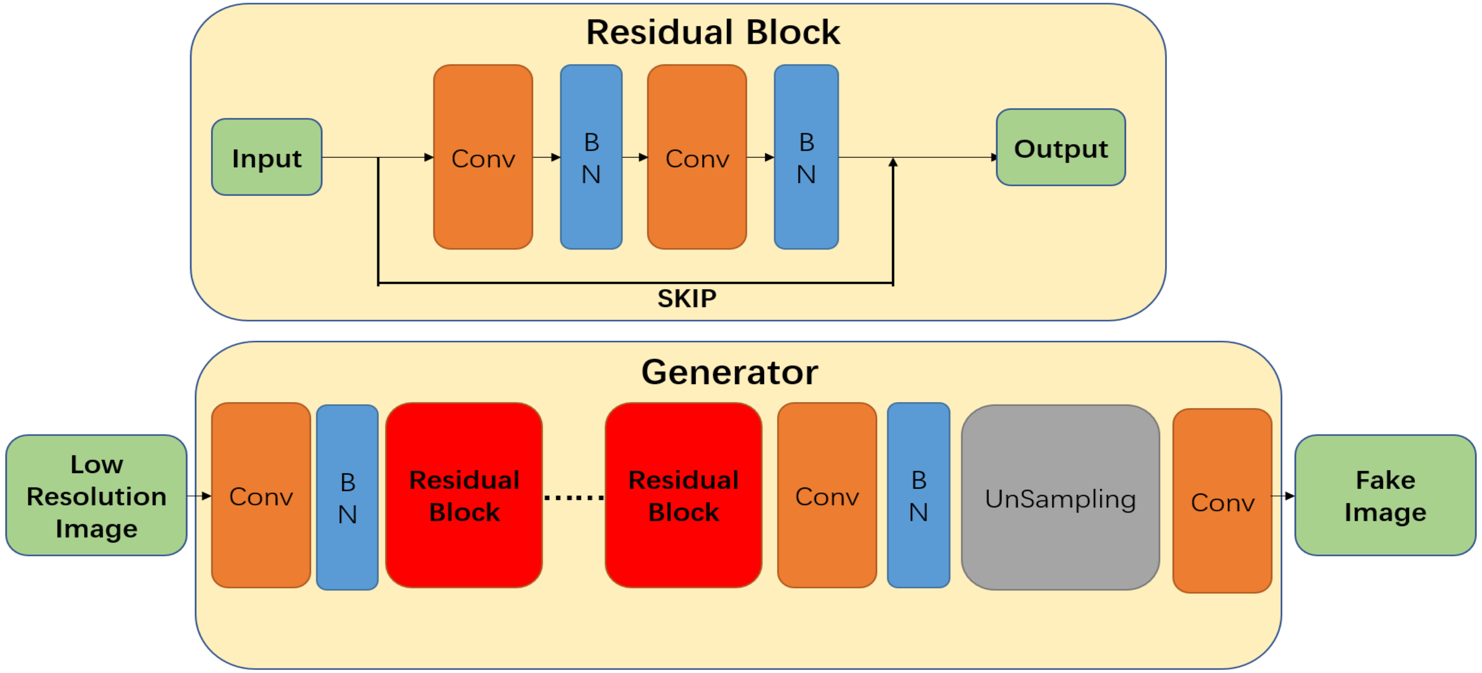
low-resolution images and high-resolution images are strictly equivalent in shape and position, and all we need to do is to add more texture on the low-resolution images

Feature map contains image texture

TOTAL LOSS = Content Loss + GAN Loss

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690)

Our SRGAN



Activation function in Generator: PRelu

Discriminator is a normal CNN binary classification model
Activation function in Discriminator: LeakyRelu

SRGAN Result

Epoch 5

Epoch 15

Original Low
Resolution Image

Epoch 25

Epoch 35

Epoch 90

Epoch 100

After 150 epochs

/06 Summary

Summary

1. Tradition methods in cv2 perform bad on vary low resolution image
2. Neural Networks’ performances are great

- SRCNN
- Autoencoder
- SRGAN

Future Improvement

1. Use more efficient structure
2. Use different types of image
3. Image augmentation

Thanks