Single Image Super-Resolution

Yujing Song

University of Minnesota, Morris

UMM CSci Senior Seminar Conference, November 17 2018

<ロ><回><回><日><日><日><日><日><日><日><日><日><日><日><日><1/34

What is Image Super-Resolution?

- Single image super-resolution, SR
- From low-resolution to high-resolution



Outline

- Background
 - Images
 - Classical neural network
 - Stochastic gradient descent
 - Convolutional neural network
- 2 Convolutional neural network for super-resolution
 - Patch extraction and representation
 - Non-linear mapping
 - Reconstruction
 - Training and result comparison
- Improvement by Generative Adversarial Network

Conclusion

Images

Background

Classical neural network Stochastic gradient descent Convolutional neural network

・ロト ・回ト ・ヨト ・ヨト

3

4/34

Images

What makes an image?

- Pixels
- Channels

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion

Channel

Images Classical neural network Stochastic gradient descent Convolutional neural network



Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Pixels

0	1	2	2	0
0	1	1	0	0
2	2	0	2	1
0	1	1	0	2
1	2	2	1	1

1	0	2	1	1
1	2	2	1	0
1	1	0	2	1
0	0	0	1	1
2	1	1	2	0

0	2	2	0	2
1	1	1	1	1
0	0	1	2	2
0	0	1	2	2
0	2	2	2	0

Red Component

Green Component

Blue Component

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Image Upscaling



- Bicubic interpolation
- Sparse-coding base method (SC)

Images Classical neural network Stochastic gradient descent Convolutional neural network

Basic structure of neural network



Images Classical neural network Stochastic gradient descent Convolutional neural network

Basic structure of neural network



<ロ> < 回> < 回> < 目> < 目> < 目> 目 のQ(0 9/34

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Activation function

Rectified linear unit (ReLU):

R(z) = max(0, z)

To introduced the non-linearity in neural networks



10/34

Loss function

Images Classical neural network Stochastic gradient descent Convolutional neural network

11/34

Loss function is used to measure the degree of inconsistency between the predicted value and the true value:

• Peak signal-to-noise ratio (PSNR)

$$PSNR = 10\log_{10}\frac{MAX_{I}^{2}}{MSE}$$

• Mean squared error (MSE)

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Mean squared error

- n: number of pixels
- X: pixels of original image
- \widehat{X} : pixels of high-resolution image

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (X_i - \widehat{X}_i)^2$$

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

13/34

Stochastic gradient descent

- W_{j+1} : new weight
- W_j: the weight from last iteration W_j
- η : learning rate

$$W_{j+1} = W_j + \eta \cdot \frac{\partial L}{\partial W_j}$$

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Back-propagation



Kernel

Images Classical neural network Stochastic gradient descent Convolutional neural network

Input image Convolution Kernel Feature map $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

0

RGB and Kernel

0	1	2	2	0
0	1	1	0	0
2	2	0	2	1
0	1	1	0	2
1	2	2	1	1

1	0	2	1	1
1	2	2	1	0
1	1	0	2	1
0	0	0	1	1
2	1	1	2	0

2 0

Green Component

Γ	-1	-1	-1
	0	1	0
	0	1	1

Green Kernel

1	1	1	1	1
0	0	1	2	2
0	0	1	2	2
0	2	2	2	0

0 2

2 2

Blue Component

1	1	0
-1	0	1
-1	1	-1

Blue Kernel

Red	Component

1	0	1
1	1	0
1	0	1

Red Kernel

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Sliding windows

0	2	2	0	2
1	1	1	1	1
0	0	1	2	2
0	0	1	2	2
0	2	2	2	0

0	2	2	0	2
1	1	1	1	1
0	0	1	2	2
0	0	1	2	2
0	2	2	2	0

Images Classical neural network Stochastic gradient descent Convolutional neural network

Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion

Convolution for a single channel



Convolutional neural network for super-resolution Improvement by Generative Adversarial Network Conclusion Images Classical neural network Stochastic gradient descent Convolutional neural network

Convolution for multiple channels



19/34

3

Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Super-resolution by convolutional neural network



Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Patch extraction and representation

- Y: upscaled low-resolution image
- W₁: filters in the first layer
- B_1 : biases in the first layer
- $F_1(Y)$: feature maps in the first layer

$$F_1(Y) = max(0, W_1 * Y + B_1)$$



Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Patch extraction and representation

Suppose the input image Y:

- Size: $A \times B$
- n_1 filters
- Kernel size: $f_1 \times f_1$

After the first step operation:

- n₁ feature maps
- Without negative pixels



Non-linear mapping

- W₂: filters in second layer
- B₂: biases in second layer
- $F_2(Y)$: feature maps in second layer

 $F_2(F_1(Y)) = max(0, W_2 * F_1(Y) + B_2)$



Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Non-linear mapping

In the second operation step:

- n_2 filters
- Kernel size: 1×1

Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Notice:

- Kernel size is 1×1
- Same with fully connected layer



Reconstruction

- W₃: filters in third layer
- B₃: biases in third layer
- F(Y): high-resolution output

 $F(F_2(F_1(Y))) = W_3 * F_2(Y) + B_3$



Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Reconstruction

- *n*₂ kernels
- Kernel size: $f_3 \times f_3$
- Filters act like an averaging filters



Patch extraction and representation

Training and result comparison

Non-linear mapping

Reconstruction

Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

3

27/34

Training

Training set:

- 395,909 images from ImageNet
- 5 million of 33×33 sub-images

Training kernels:

- MSE as loss function
- Stochastic gradient descent

Training process

Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison



(ロ) (部) (言) (言) (言) (で) (28/34)

Patch extraction and representation Non-linear mapping Reconstruction Training and result comparison

Result comparison



Super-resolution by generative adversarial network

Generative adversarial network:

- Generator network: creator
- Discriminator network: classifier

What is improved?



SRCNN

SRGAN

Conclusion

- Neural network
- Convolutional neural network
- Solving super-resolution problem

Acknowledgement

Thank you to my advisor Peter Dolan and Elena Machkasova for guidance and feedback

Discussion

Questions?

<ロト < 回 > < 言 > < 言 > こ き < こ > こ の < C 34/34