

Super-Resolution Reconstruction Using Generative Adversarial Network for Wideband Magnetic Resonance Imaging

¹*Feng-Yu Hsu* (許豐育)

²*Chiou-Shann Fuh* (傅楸善)

³*Jyh-Horng Chen* (陳志宏)

¹Graduate Institute of Biomedical Electronics and Bioinformatics,
National Taiwan University, Taiwan

²Department of Computer Science & Information Engineering,
National Taiwan University, Taiwan

³Department of Electrical Engineering,
National Taiwan University, Taiwan

E-mail: R07945009@ntu.edu.tw

fuh@csie.ntu.edu.tw

jhchen@ntu.edu.tw

ABSTRACT

Magnetic Resonance (MR) images with both high resolutions and precise structure details are significant in not only research but also medical domain. However, its relatively long imaging time makes demands of acceleration techniques stay high. Besides, with the rapid advancement of machine learning techniques, learning-based Super-Resolution (SR) methods for MRI have attracted considerable attention. So we try to combine these two field with Single-frequency Excitation WideBand (SE-WB) MRI and super-resolution (SR) Generative Adversarial Network (GAN) to achieve better reconstruction quality.

Keywords: *Deep Learning; Super-Resolution; Magnetic Resonance imaging; Compressed Sensing; Generative Adversarial Network;*

1. Introduction

Since Magnetic Resonance Imaging (MRI) has several advantages like non-radiation and non-ionizing nature. It becomes a widely used imaging technology for visualizing the structure and functioning of the body. But MRI usually takes dozens of minutes. And this shortcoming lower its potential of application. So it's essential to reduce time consuming of it. Compressed Sensing (CS) MRI algorithms[1,2,3,4] becomes a possible solution. And SE-WB MRI [4], proposed by E. L. Wu, is one of the potential methods. It reduces number of phase-encoding steps and take advantage of increased readout sampling rate with separation gradient(s) applied along the phase-encoding direction(s) which makes the horizontal readout trajectory inclines. As a result, it forms parallelogram shaped k -space coverage with the same sampling density as standard imaging. So that it can accelerate the imaging process and lose part of high-frequency signal merely.

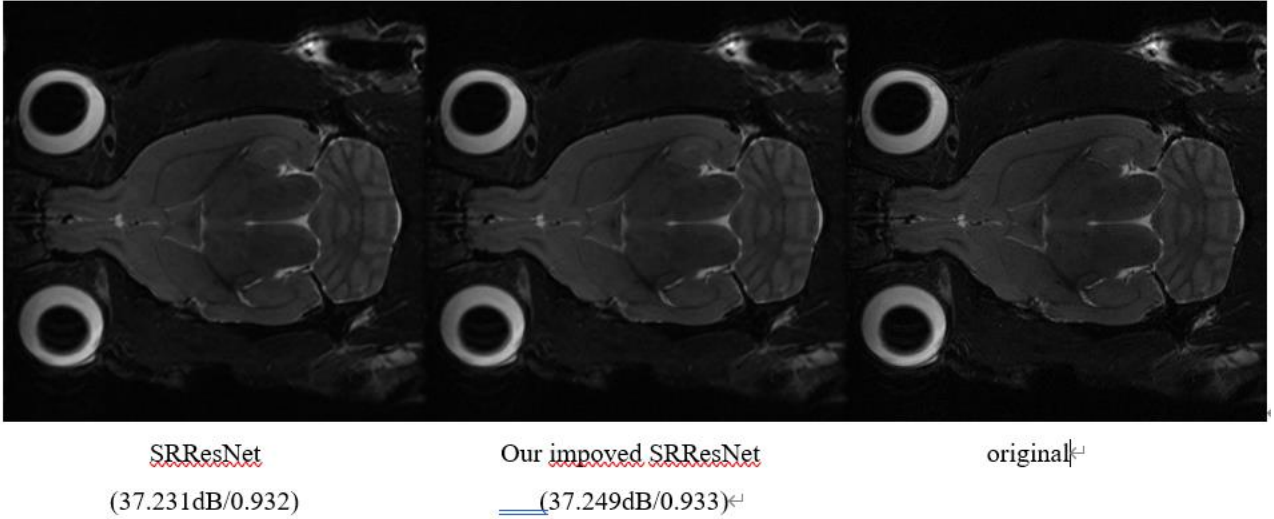


Fig. 1. From left to right: SRResNet optimized for MSE, our improved SRResNet optimized for a more powerful loss inspired by triplet loss, original HR image. Corresponding Peak Signal-to-Noise Ratio (PSNR) and structural similarity (SSIM) are shown in brackets. [2× upscaling]

Moreover, as computing power increasing significantly nowadays, Deep Learning (DL) becomes a proper solution for many issues. Meanwhile, MR image resolution is positively correlated to the time it takes. Therefore, SR DL method could also be applied to accelerate MRI process

Many learning-based methods improve image quality noticeably, especially the state-of-the-art convolutional methods [5,6,7,8,9,10]. For example, Super-Resolution generative adversarial network (SRGAN) [9], proposed by Christian Ledig, reconstruct more high-frequency signal of image than those methods without discriminator. In other words, it provides better visual quality in spite of having lower PSNR compared with its generator called SRResNet. Moreover, a method called Enhanced Deep Residual Network (EDSR) [10] improved from SRResNet provides better accuracy which is a major factor of medical imaging. Therefore, we try to apply EDSR's architecture to SRGAN to figure out whether it provides better accuracy and maintains superior visual quality at the same time. Fig. 1. shows that our improved SRResNet achieve almost the

same performance on PSNR and SSIM with lower parameters.

In summary, we combine WB MRI and SRGAN with EDSR and reconstruct low-resolution WB MR images into high-resolution ones. Meanwhile, we apply GAN architecture to our model trying to obtain images with

high-frequency details restored, an example that was super-resolved with a 2× upscaling is shown below in Fig. 2.

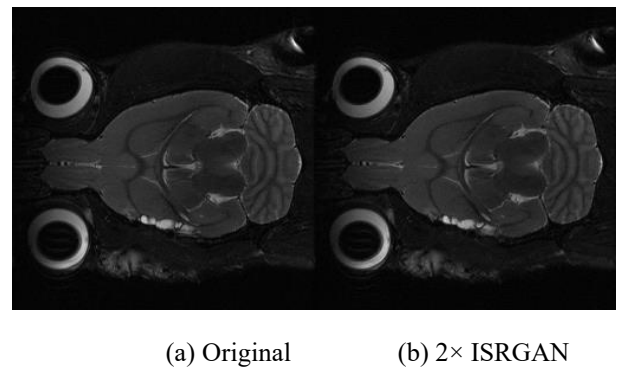


Fig. 2. The 2× ISRGAN super-resolved image (b) is pretty similar with the original one (a) even in the detail structure.

2. Related Works

2.1. Compressed Sensing MRI

As mentioned above, MRI has major limitation on relatively slow data acquisition speed in spite of several advantages. Compressed sensing-based MRI is one of the most effective techniques to accelerate magnetic resonance acquisition by reducing the number of k -space measurements directly acquired by the machine [1,2,3,4]. CS theory shows how accurate or even perfect reconstruction can be achieved via appropriate optimizations to fill in the missing Fourier coefficients of k -space. Recently, CS MRI has been approved by the FDA for some major vendors [11]. Hence more CS MRI methods are expected to be used in clinics, in order to improve patient comfort by speed up the data acquisition process.

SE-WB accelerate the imaging process by increasing k -space coverage and maintaining spatial frequencies with zig-zag trajectory. After processing such as removing the data sampled during the buffer intervals and regriding the k -space according to the trajectory. The k -space became zig-zag form whose unsampled triangular regions were zero-filled.

2.2. Convolutional Neural Network

Recently, with the rapid evolution of machine learning, deep neural networks have become popular among SR studies. Moreover, many computer vision problems is set by specifically designed CNN architectures. Also, many powerful models are proposed. Following is a briefly introduction.

Compared to VDSR [12], proposed by J. Kim et al. SRResNet [9] successfully reduces the influence of time

and memory issue caused by requiring bicubic interpolated images as the input. And it achieves good performance by simply employing the ResNet architecture from He et al. [13] without much modification. However, the model was designed to solve those problems with different attributes like image classification and detection. Therefore, it might be unsuitable to apply ResNet architecture directly to low-level computer vision problems such as super-resolution .

As described above, EDSR [10] is thence proposed. It removes the Batch Normalization (BN) layers of residual blocks in SRResNet. Since these BN layers normalize the features, they reduce the range flexibility from networks by normalizing the features, so it is better to get rid of them. Moreover, this adjustment saves about 40% of memory usage during training. In other words, this allows architecture to become deeper and wider and maintain lighter model size at the same time. In the paper, the model is designed by setting with $B = 32$, $F = 256$ which represent depth (the number of layers) and width (the number of feature channels) respectively and achieve better performance.

2.3. Loss

Since many SR networks use Mean Square Error (MSE) as their loss function to maximize the Peak Signal-to-Noise Ratio (PSNR). It is proved that L1 loss provides better convergence than L2 loss by comparing performance between SRResNet trained with these two loss function.

In general case, triplet loss is used in tasks like face recognition and classification. There are three parameters which is anchor, positive and negative input. The main concept is making anchor get closer to

positive input and farther to negative input simultaneously. And the distance is usually calculated with MSE. Meanwhile, a margin value between 0 to 1 is used to enhance the ability of classification as the value increase. But this can also bring on higher training difficulty.

3. Material and Method

3.1. Research Purpose

To find and verify most suitable model architecture and parameters for improved SRResNet. Then we adjust loss function to apply to SRGAN in order to achieve better performance.

3.2. Research Method

First, we remove the BN layers of residual blocks in SRResNet. And verify whether it performs as well as origin SRResNet with WB MR images super-reconstruction. If so, the lower hardware load due to smaller structure allows deeper or wider model expansion. Then the second study following is to find the most suitable depth and width of modified SRResNet for WB MR images.

Inspired by triplet loss, we adjust L1 loss into the following form:

$$L(X_A, X_N, X_P) = \sum_{i=1}^N [|x_i^A - x_i^P| - \alpha |x_i^A - x_i^N|]$$

Where x_i^A , x_i^P and x_i^N are the SR images, HR images and WB HR images.

In addition to L1 loss, we add an operand of distance between anchor and negative input to let model learn SR reconstruction better by getting away from negative input. However, the main target is still reconstructing

Low Resolution (LR) images into High Resolution (HR) ones. A weight is added to prevent model from over focus on secondary aim. Furthermore, we remove the margin value of origin triplet loss to avoid rising of training difficulty.

Last, we need to figure out the best weight for the modified loss. Then applying these to the new SRResNet called Improved SRResNet (ISRResNet) as the generator might achieve better performance than the origin SRGAN. The work flow how we train SRGAN is shown in Fig. 4.

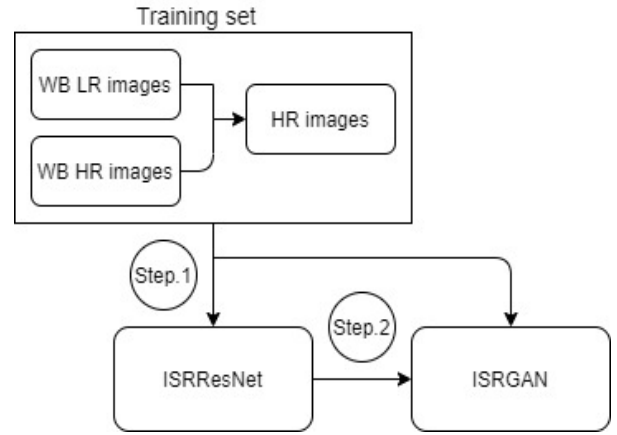


Fig. 4. Training steps. Train ISRResNet to obtain initial weight of generator in ISRGAN first. Then train ISRGAN with the weights.

3.3. Training Set

Because of high time consuming and low reproducibility caused by artifacts and motions etc. It is not only almost impossible to collect Training data from hospitals and most research centers, but also unrealistic to produce by repeated imaging projects. Due to these difficulties of collecting MR images training sets. We use WB-style k -space mask on origin HR MRI k -space to get WB HR k -space. Then we Fourier transform them into images to obtain HR and WB HR training set. And LR training set

can be generated by down-sampling HR training set with bicubic method. All these steps are summarized and shown in Fig. 5.

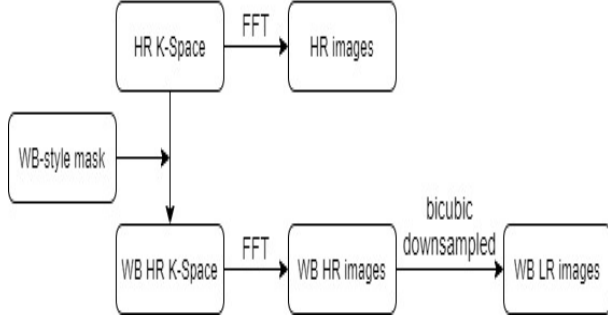


Fig. 5. Steps of creating our HR, WB LR and WB HR training sets where FFT means Fourier Transform.

Through this work flow, our training set is created with 312 RARE images of rat brain whose images are selected from the middle 3 slices of whole brain MR from 104 cases each. And we cover WB-style k -space mask like Fig. 6. on these images' k -spaces to generate simulated WB MR images. Then, we down-sample them with bicubic method to get low-resolution version of WB MR image. All the original images of training sets' resolution are 256*256 pixels.

And we follow the same steps to create 78 images testing set with other 26 cases.

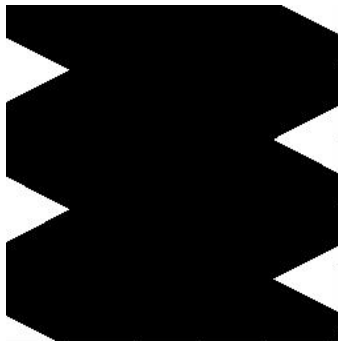


Fig. 6. WB-style k -space mask. We use it to get rid of specific area of HR MRI k -space to obtain WB-style k -space.

4. Results

Refer to table. 1. We can see that ISRResNet performs as well as the original SRResNet with the same size of residual blocks. Also, we test different numbers of residual blocks and convolution feature channels to figure out the best depth and width of ISRResNet. ISRResNet contains 12 residual blocks with 64 convolution feature channels not only perform as well as SRResNet but also saves almost half training time consume. So we decide to choose this setting as our SRGAN's generator. The complete architecture is shown on Fig. 7.

Model name	nb	nf	PSNR	Time Proportion
SRResNet	16	64	35.45	1
ISRResNet	16	64	35.51	0.69
	20	64	35.504	0.8
	24	64	35.517	0.91
	16	96	35.501	1.07
	16	128	35.514	1.33
	12	64	35.511	0.54
ISRResNet_NAC	12	64	35.49	0.57

Table 1. Performance of different setting of residual blocks base on Peak Signal-to-Noise Ratio (PSNR). NAC means layer order in residual block is activate function (ReLU) and then convolution layer. Nb and nf represent numbers of residual blocks and feature channels.

As describe above, we decide to use 12 residual blocks and 64 convolution feature channels in our ISRResNet. Moreover, Table 2 shows ISRResNet achieves better performance with our triplet type loss in spite of MSE loss and why we set 0.1 as our loss weight. Moreover, it improves the original SRResnet as well. This indicates that our loss can be more powerful function for other models.

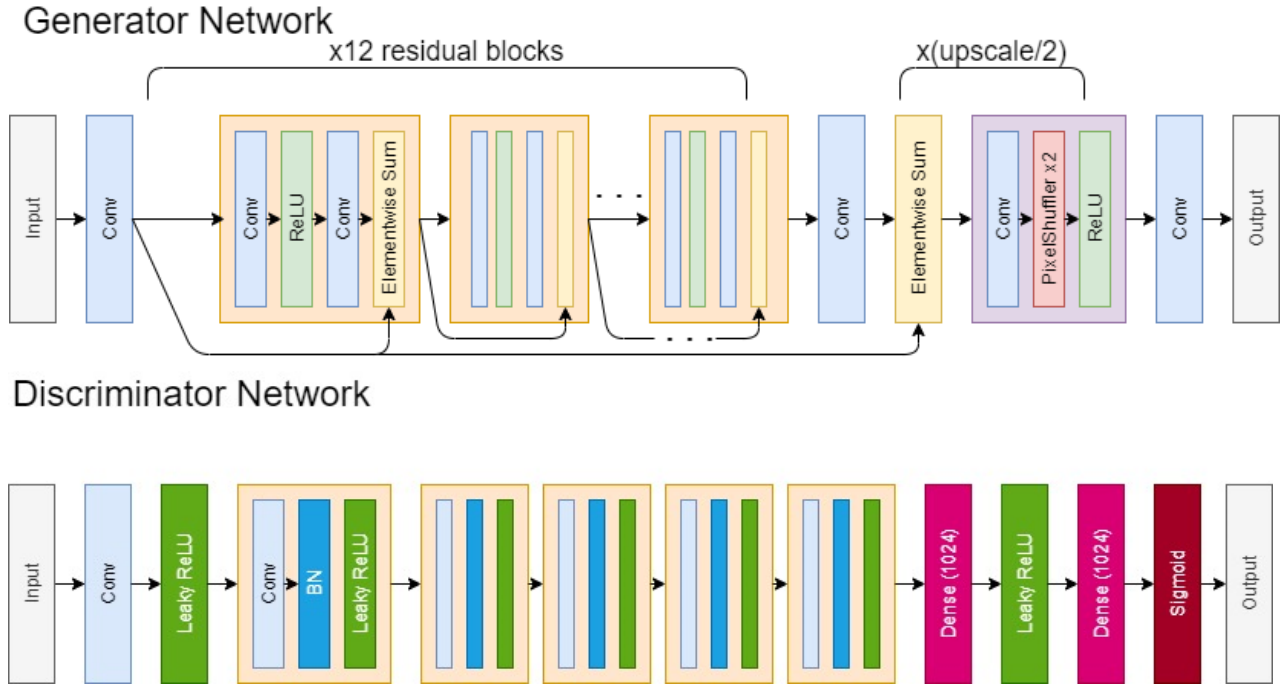


Fig. 7. Architecture of Generator and Discriminator Network with corresponding numbers of residual blocks and dense layer width.

4. Conclusion

Model	PSNR	SSIM
ISRResNet ($\alpha=0.3$)	35.627	0.907
ISRResNet ($\alpha=0.2$)	35.647	0.907
ISRResNet ($\alpha=0.1$)	35.648	0.907
ISRResNet with L1 loss	35.51	0.906
SRResNet ($\alpha=0.1$)	35.659	0.908
SRResNet with L2 loss	35.45	0.906

Table 2. Performance of ISRResNet and SRResNet with different loss setting. α represents weight of our triplet type loss.

Finally, our model architecture is defined and shown as Fig. 7. And Fig. 8 shows performance of various models on our dataset.

Since lots of people try to design new kind of blocks or arrange various blocks and connections, even just widen or deepen models to achieve better performance. There's still other parameter can be optimized for particular tasks. As we refer to triplet loss and join negative training set. It improves SRResNet without complicated adjustment. This might be a feasible solution for other models.

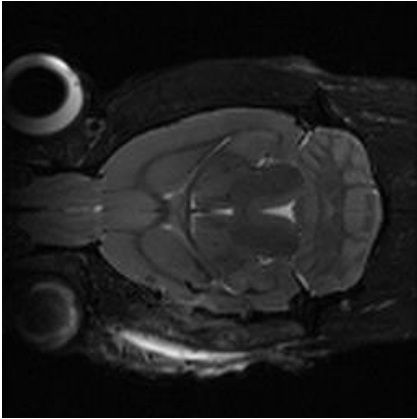
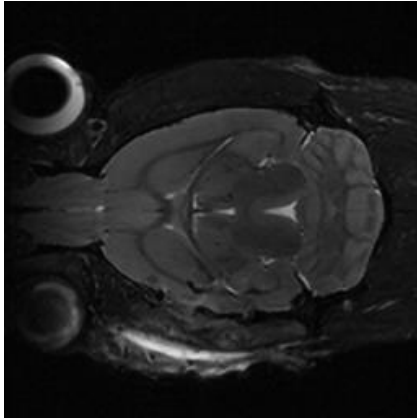
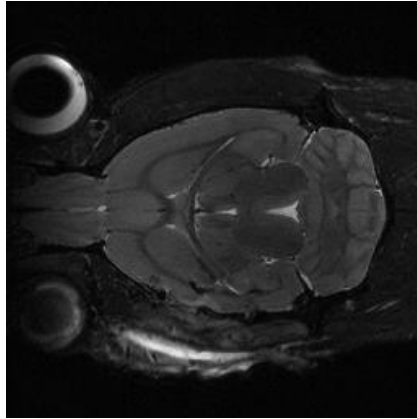
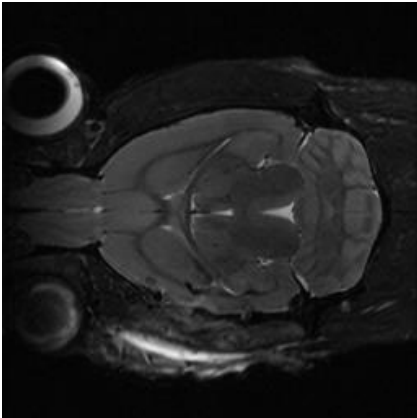
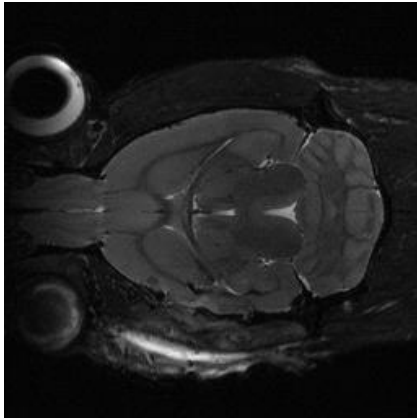
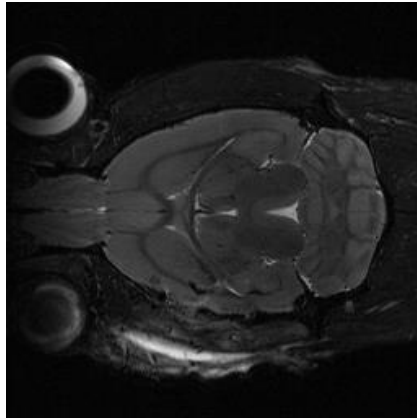
ESPCN (34.16dB/0.892)	SRResNet (35.45dB/0.906)	SRGAN (33.404dB/0.85)
		
(35.181dB/0.909)	(35.935dB/0.921)	(34.231dB/0.87)
ISRResNet (35.498dB/0.906)	ISRGAN (33.531dB/0.854)	HR image
		
(35.973dB/0.922)	(34.386dB/0.87)	

Fig 8. Performance of models based on PSNR and SSIM.

References

- [1] Hisamoto Moriguchi, and Jeffrey L. Duerk, Bunched phase encoding (BPE): A new fast data acquisition method in MRI, 2006.
- [2] E. L. Wu, J-H. Chen, and T-D. Chiueh, Wideband MRI: A New Dimension of MR Image Acceleration, 2009
- [3] Kim MO1, Lee J, Zho SY, and Kim DH, Accelerated MR whole brain imaging with sheared voxel imaging using aliasing separation gradients, 2013
- [4] Edzer L. Wu, Yun-An Huang, Tzi-Dar Chiueh, and Jyh-Horng Chen, Single-frequency excitation wideband MRI (SE-WMRI), 2015.
- [5] J. Kim, J. K. Lee, K. M. Lee, "Accurate image super-resolution using very deep convolutional networks", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1646-1654, 2016.
- [6] Y. Tai, Y. Jian, X. Liu, "Image super-resolution via deep recursive residual network", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017.
- [7] X. J. Mao, C. Shen, Y. B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections", *Proc. 30th*

Int. Conf. Neural Inf. Process. Syst., pp. 2802-2810, 2016.

[8] W. Yang et al., "Deep edge guided recurrent residual learning for image super-resolution", IEEE Trans. Image Process., vol. 26, no. 12, pp. 5895-5907, Dec. 2017.

[9] Christian Ledig, Lucas Theis, Ferenc Huszar, and Jose Caballero, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, 2017.

[10] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, Enhanced Deep Residual Networks for Single Image Super-Resolution, 2017

[11] J. A. Fessler, "Medical image reconstruction: A brief overview of past milestones and future directions" in arXiv:1707.05927, 2017

[12] J. Kim, J. Kwon Lee, and K. M. Lee. Accurate image superresolution using very deep convolutional networks. In CVPR 2016

[13] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR 2016