

Article

Multi-Branch Deep Residual Network for Single Image Super-Resolution

Peng Liu^{1,2,3,*}, Ying Hong^{1,2} and Yan Liu^{1,2}

¹ Key Laboratory of Information Technology for Autonomous Underwater Vehicles, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China; hongying@mail.ioa.ac.cn (Y.H.); liuyan@mail.ioa.ac.cn (Y.L.)

² Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China

³ University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: liupengucas@mail.ioa.ac.cn

Received: 21 August 2018; Accepted: 25 September 2018; Published: 27 September 2018



Abstract: Recently, algorithms based on the deep neural networks and residual networks have been applied for super-resolution and exhibited excellent performance. In this paper, a multi-branch deep residual network for single image super-resolution (MRSR) is proposed. In the network, we adopt a multi-branch network framework and further optimize the structure of residual network. By using residual blocks and filters reasonably, the model size is greatly expanded while the stable training is also guaranteed. Besides, a perceptual evaluation function, which contains three parts of loss, is proposed. The experiment results show that the evaluation function provides great support for the quality of reconstruction and the competitive performance. The proposed method mainly uses three steps of feature extraction, mapping, and reconstruction to complete the super-resolution reconstruction and shows superior performance than other state-of-the-art super-resolution methods on benchmark datasets.

Keywords: single image super-resolution; deep neural networks; residual networks; peak signal-to-noise ratio; structural similarity index

1. Introduction

Single image super-resolution (SISR) is an important topic in digital image processing and computer vision. SISR aims to recover a high resolution (HR) image from its low resolution (LR) image. Generally, many studies assume that the LR image is down-sampled or bicubicing from HR image with a scale factor. In the past decades, the problem of image super-resolution [1] (SR) has attracted extensive attention. These SR works have been applied to satellite imaging [2], medical imaging [3,4], face recognition [5], and surveillance [6]. Inherently, the SR is a highly ill-posed problem since there is a lot of high-frequency information lost for the LR image. Furthermore, the one-to-many mapping from LR image to HR image has many solutions. Therefore, SR can be considered as an inference problem, which needs to restructure the missing high-frequency data from the low-frequency components.

Thus far, many methods based on deep convolution neural network [7–10] have been proposed for the single image super-resolution and show excellent performance. These approaches apply the back-propagation algorithm [11] to train on large image datasets in order to learn the nonlinear mappings between LR images and HR images. Compared with previous statistics-based [12–15] and patch-based [16–22] models, these techniques provide improved performance for the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM). Due to the characteristics of deep convolution networks, these works will still have some existing defects. Such as the minor changes in network structure and different training methods will cause huge differences in reconstruction results.

Besides, for a better mapping between the LR and HR image, the network layers and parameters often need to be increased, which will cause the network overfitting and insensitivity to initialization weights. Meanwhile, it will increase the training difficulty and the pressure of memory. Obviously, this is not beneficial for the convergence of network and the reconstruction task.

With the development of residual network [23] (ResNet), SRResNe [9] is proposed for single image super-resolution and successfully solves those overfitting and memory issues. However, since SRResNet simply employs the residual network architecture that is proposed for image classification and detection, it doesn't have much modification, which makes it unable to obtain the optimal performance in the process of reconstruction. Therefore, the method of enhanced deep residual networks for single image super-resolution [24] (EDSR) optimizes the ResNet architecture and removes unnecessary modules to simplify the network. Compared with SRResNet, EDSR solves the legacy problems and achieves superior reconstruction performance on all datasets in terms of PSNR and SSIM. However, this improved result is achieved at the cost of a large number of residual blocks and parameters. Roughly 43 M parameters are applied to the network structure of EDSR. In addition, EDSR increases the receptive field of the network structure by increasing the number of filters, which uses 256 filters, so as to achieve the purpose of superior reconstruction from the LR image. This design not only increases the difficulty of training and storage greatly, but also makes the iterative updating of parameters insensitive.

In order to solve these issues, we propose a multi-branch deep residual network. There are two designs in our method to improve the performance and ease the difficulty of training. First, multiple branches network structure is adopted in our approach and all branches are supervised. Each branch has its own reconstruction network. Additionally, the feature maps extracted by each branch are sent to its own reconstruction network to complete the SR task. While this design solves the training problem, it can also enlarge the receptive field of the network, which determines the amount of contextual information that can be exploited to infer missing high-frequency components. Therefore, our network effectively reduces the number of filters.

Second, we propose a new residual network structure. Compared with the existing ResNet structure, our proposed structure can get more effective components to infer the missing high-frequency that determines the authenticity of the reconstructed image. This is proved in the experiments, which shows that the proposed ResNet structure can achieve better reconstruction with significantly fewer parameters. Furthermore, the skip-connection from input to the reconstruction layer is applied in our network, which can largely achieve the data correlation and information sharing between the input and output. Finally, we evaluate our algorithm on the standard benchmark datasets. Compared with a number of current methods, our model shows superior performance.

2. Related Work

2.1. Image Super-Resolution

Thus far, many methods have been proposed to solve the super-resolution problem. They can be categorized into four types—image statistical methods, prediction based methods, edge based methods, and patch-based methods. Early algorithms apply interpolation techniques and statistical methods [25,26] to SR, but work hardly with the lost details and realistic textures. Then, methods based on prediction, which are the first techniques for SISR, are proposed to reconstruct higher resolution images. While those filtering algorithms use bicubicing or linear filtering to oversimplify the SISR problem, methods based on edge preservation [27,28] have been proposed. These approaches not only take advantage in speed, but can also achieve overly smooth texture reconstruction. In addition, many works based on patch [14,26,29,30] are also designed for SR. With the patch redundancies across scales within the image, Glasner et al. [29] proposed an algorithm to drive the research and development of SR. Compared with the other three methods, these methods based on patch exhibit superior performance.

In order to achieve a better reconstruction, a superior mapping function from low resolution to high resolution is necessary. Among the existing techniques, those works based on deep neural networks are considered to have a strong capability to achieve significant mapping in image super-resolution. Dong et al. [10] first adopt convolutional neural network (CNN) architecture to solve the problem of SISR. They use a three-layer deep fully convolutional network to achieve state-of-the-art SR performance. This attempt shows that the CNN model can further improve reconstruction both in terms of speed and quality. Subsequently, various algorithms based on CNN model, including deeply-recursive convolutional network (DRCN) [8], are proposed to study for SR. With the residual network introduced by He et al. [23], the CNN methods can be used to train much deeper network and have better performance. Such as the method of EDSR, which is proposed by Lim et al. [24], uses enhanced deep residual networks for SISR and achieves improved performance. Moreover, Ledig et al. [9] first adopts the generative adversarial networks [31] (GAN) for SISR. In their model, powerful data generation capability of the GAN network and appropriate evaluation function based on the Nash equilibrium have been well applied.

To train network models better, a perceptual loss function is needed. As the common loss function, mean squared error [32,33] (MSE) has been widely used to calculate the pixel-wise error in general image restoration. Meanwhile, MSE is usually applied to compute the peak signal-to-noise ratio (PSNR), which is a major performance measure in reconstruction. Besides, the structural similarity index (SSIM) is another evaluation index for SR. The higher PSNR or SSIM is corresponding to better recovering quality.

2.2. Residual Network in Super-Resolution

Residual network (ResNet) is a kind of improved neural network, which achieves the identity mapping by applying the shortcut connections in the structure. This design cannot only eliminate the degradation during the SGD optimization process, but also enables the data to flow across layers. It is this characteristic that allows the network to be extended deeper. Compared with the single deep convolution neural network, ResNet can avoid the overfitting effectively and achieve better performance. Thus far, ResNet has been widely used in deep networks to deal with the computer vision problems. Meanwhile, these properties provide a guarantee for ResNet to be applied to solve the SR problem. The methods of SRResNet and EDSR, which apply the residual network in their structure, have achieved state-of-the-art performance in SISR.

3. Proposed Methods

3.1. Residual Blocks

Since the residual networks were proposed to solve the computer vision problems such as image classification and detection, they have shown superior performance, especially for the tasks that are from low-level to high-level. SRResNet employs the ResNet architecture directly to complete the SR reconstruction. EDSR improves the performance by adjusting the ResNet structure. In proposed model, the ResNet architecture is further improved and achieves better performance in SR. Compared with original residual network, we optimize the structure of residual network by substituting BN layers with the convolution layers. Further, different from the existing residual network structures that obtain the output from the input and the output of the last layer of the network, the newly proposed ResNet structure combines the input and the output of every layer to get the network output, as shown in Figure 1. The experiment shows that these modifications can not only speed up the convergence of the network, but can also improve the performance substantially in terms of image details and textures.

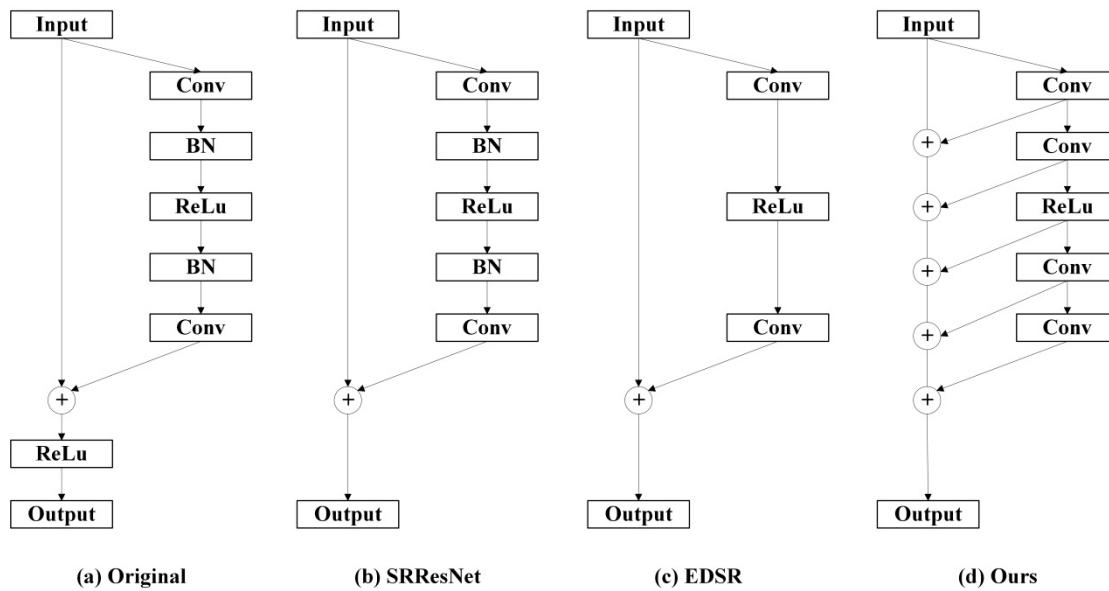


Figure 1. The comparison of different residual networks structure among the Original, SRResNet, EDSR, and our proposed.

3.2. Model Architecture

In this section, the proposed model structure will be introduced, outlined in Figure 2. Our model conceptually consists of three parts: Feature extraction, mapping, and reconstruction. The feature extraction operation gets the features from the input LR image and represents them as a set of feature maps which are ready for the next mapping operation. In order to deliver more information to the next operation, in addition to using multiple filters and residual blocks to operate on the input, we also use a sliding connection to send the input directly to the mapping network. Then, non-linear mapping from LR to HR is worked by the mapping operation, which is the main component that solves the SR task. Obviously, the quality of SR mainly depends on the performance of the mapping network. In proposed model, the mapping network is composed of five branches. Each branch applies residual blocks and filters to achieve the effective mapping from LR to HR feature maps with different parameters. Moreover, the convolution layers are inserted between every two branches, which will result in different sizes of each branch network. This design provides the feasibility to implement multi-scale network and takes advantage of inter-scale correlation. Finally, the reconstruction networks undertake the task of rebuilding the super-resolution image. Since the output of each branch not only differs in size but also in number, every branch has its own independent reconstruction subnetwork. Every reconstruction network combines the output of the corresponding branch network with the original input to restore the SR image. Furthermore, we apply the sub-pixel convolution layers to every reconstruction network to upscale the LR image. Compared with the deconvolution layer or other implementations, which use various forms of upscaling before convolution, the sub-pixel convolution layer is faster in training. According to the structure of MRSR, the final SR image is derived from each reconstruction network.

In the proposed architecture, the feature extraction modules consist of ten residual blocks and filters with 3×3 kernels that allow more detailed texture information and hidden states to be passed. As the most important component of the model, the mapping network combines the advantages of multi-branch networks and residual networks where kernels are set to 5×5 . By adopting larger kernels, the larger receptive field can be covered in the mapping network. Our model has approximately 50 times more receptive field than DRCN.

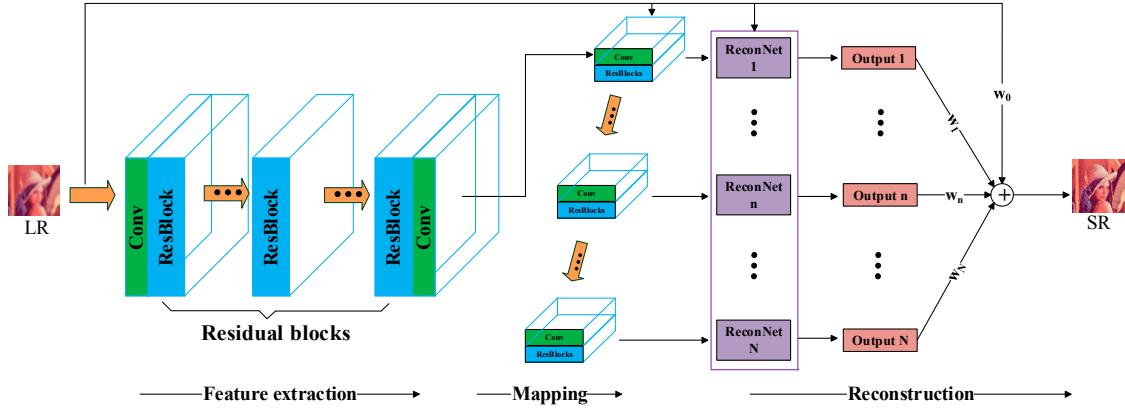


Figure 2. The architecture of our proposed SR network (MRSR), which consists of three parts: Feature extraction, mapping, and reconstruction. The feature extraction is composed of multiple filters and residual blocks. The non-linear mapping between LR and SR adopt the multi-branch network structure and each branch is made up of residual blocks. In the reconstruction, the final output is restored from every branch output and the LR input with different weights.

3.3. Training

As described in Section 3.2, the feature extraction network takes the low-resolution image I^{LR} as input. Assuming that this part of network is model F , we can get the output $F(I^{LR})$, which is the input to the mapping network M . Determined by the multi-branch structure, the output of each branch network $M_n, n = 1, 2, \dots, N$ can be described as:

$$M_n = g_n(g_{n-1} \cdots (g_1(F(I^{LR}), I^{LR}))) \quad (1)$$

where the operator g denotes the function which represents each branch network. Since every branch network has different components, each $g_n, n = 1, 2, \dots, N$ represents different function expression.

The reconstruction network R takes the M_n as input and complete the reconstruction of super-resolution images. Under the branch-supervision, the predictions $R_n, n = 1, 2, \dots, N$ from each reconstruction network are:

$$R_n = h_n(I^{LR}, M_n) \quad (2)$$

Same as the mapping network, the function h completes the reconstruction task of each recursive network. This process can be considered as the inverse operation of feature extraction network in a sense. Following our model, the final output I^{SR} is the weighted average of all intermediate predictions and original input I^{LR} :

$$I^{SR} = \sum_{n=1}^N w_n \cdot R_n + w_0 \cdot I^{LR} \quad (3)$$

Obviously, the w_n represents the weight of every intermediate prediction in reconstruction. Learning from training, these weights will directly determine the final reconstruction quality.

In order to achieve superior performance, besides the ingenious network architecture, a loss function, which is as accurate as possible, is also necessary. Here, the training loss function, which can find optimal parameters for a proposed network, will be introduced. First, the difference between super-resolution image and high-resolution is the most intuitive expression. Combined with L2 loss, which is generally preferred since it minimizes the MSE, the part of loss function is defined as:

$$\ell_1^{SR} = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H (I_{(i,j)}^{SR} - I_{(i,j)}^{HR})^2 \quad (4)$$

where W and H are the image size. The I^{SR} and I^{HR} represent the reconstructed image and reference image respectively. According to proposed model, every branch network needs to be supervised:

$$\ell_2^{SR} = \frac{1}{NWH} \sum_{n=1}^N \sum_{i=1}^W \sum_{j=1}^H (R_{n(i,j)} - I_{(i,j)}^{HR})^2 \quad (5)$$

Furthermore, based on the ideas of Johnson et al. [34] and Bruna et al. [35], a pre-trained 16 layer VGG network is adopted to compute the VGG loss ℓ_{VGG}^{SR} , which is closer to perceptual similarity. The ℓ_{VGG}^{SR} is based on the ReLU activation layers of the VGG network, described as Zisserman [36]. Each activation layer in the VGG network will get different feature maps for the two different inputs of the reconstructed image I^{SR} and the truth image I^{HR} . Then, we define the ℓ_{VGG}^{SR} as the Euclidean distance between these feature maps:

$$\ell_{VGG}^{SR} = \frac{1}{XY} \sum_{i=1}^X \sum_{j=1}^Y \left\| \psi_{(i,j)}(I^{SR}) - \psi_{(i,j)}(I^{HR}) \right\|^2 \quad (6)$$

where $\psi_{(i,j)}$ denotes the i -th feature map of j -th activation layer of the VGG network. X and Y are the dimensions of the feature maps. For the final loss function, we have:

$$\ell^{SR} = \alpha \times \ell_1^{SR} + (1 - \alpha) \times \ell_2^{SR} + \beta \times \ell_{VGG}^{SR} \quad (7)$$

where α and β represents the weight of the partial loss, which are setting between 0 and 1. Based on a series of experiments, we find that high α makes the model stable and easy to converge.

4. Experiments

4.1. Datasets

Fairly, we implement all experiments on the standard datasets which have been widely used for other image restoration tasks. These datasets mainly include: Set5 [37], Set14 [38], BSD100 [39], and Urban100 [40]. Meanwhile, our model is also being tested on a newly proposed high-quality images dataset DIV2K [41], which contains 1000 training and testing images and is the official dataset of the NTIRE2017 Super-Resolution Challenge.

4.2. Training Details

For all experiments, the RGB HR image is down-sampled by using a bicubic kernel with a scale factor r . This is a common method to obtain the LR image, which has been applied in other state-of-the-art methods for an SR problem. The HR images used in all experiments are cropped to 96×96 with a batch-size of 16. In training, our network is trained with an initialized learning rate of 10^{-4} and 10^6 update iterations on a GPU GTX1080. For optimization, we use the ADAM (Adaptive Moment Estimation) with the setting $\beta_1 = 0.9$, $\beta_2 = 0.99$, and $\epsilon = 10^{-8}$.

In terms of network architecture, we use 60 residual blocks totally. The weights in layers are initialized by the same method. For the biases, we set to zero. To achieve a fair comparison with other SR reconstruction methods, we take the PSNR and SSIM as the performance metrics to evaluate the experiment results.

4.3. Comparisons with State-of-the-Art Methods

In this section, the qualitative and quantitative comparisons with other state-of-the-art SR approaches will be introduced. We mainly compare the performance of MRSR with results from Bicubic, A+ [21], SRCNN, DRCN, SRResNet, EDSR, and the most recent work WSD [42]. Benchmark, the public code for these algorithms and the same technique to obtain the LR images from HR images were used in experiments. For a good visual comparison, we also adopt same method to deal with the luminance components.

Besides, compared with EDSR, the new ResNet structure uses fewer filters and parameters. The proposed model only uses 128 filters that account for 50% of EDSR. Not only that, the parameters has been reduced to about 12 M, which is 0.28 times compared with EDSR, as shown in Table 1. Consequently, the GPU memory usage and the training difficulty can be dramatically reduced. Moreover, the optimized ResNet architecture exhibits better performance with less computation.

The summary quantitative results on several datasets are presented in Table 2. As can be seen from the table, our model exhibits superior performance than existing methods in all datasets and scale factors in terms of PSNR and SSIM. In addition, visual comparisons of the super-resolution images are shown in Figures 3–5. It can be seen intuitively from the figures that our reconstructed images show higher quality regardless of details or textures and exhibit more realistic outputs compared with the previous works. In Figure 6, the proposed approach is also compared with DRCN quantitatively. Moreover, the comparison between our algorithm and the most recent work WSD, which use wiener filter in similarity domain to achieve the reconstruction of single image super-resolution, is shown in Figure 7. Obviously, the SR result has been improved greatly.

Table 1. Parameters comparison.

		SRResNet	EDSR	MRSR
Residual blocks		16	32	60
Filters		64	256	128
parameters		1.5 M	43 M	12 M

Table 2. Benchmark test results (PSNR(dB)/SSIM) on datasets Set5, Set14, B100, Urban100, and DIV2K for scale factor $\times 2$, $\times 3$, $\times 4$. Bold indicates the best performance.

Dataset	Scale	Bicubic	A+	SRCNN	DRCN	SRResNet	EDSR	WSD	MRSR (ours)
Set5	$\times 2$	33.66/0.9299	36.54/0.9544	36.66/0.9542	37.63/0.9588	-/-	38.11/0.9601	37.21/-	38.32/0.9610
	$\times 3$	30.39/0.8682	32.58/0.9088	32.75/0.9090	33.82/0.9226	-/-	34.65/0.9282	33.50/-	34.88/0.9305
	$\times 4$	28.42/0.8104	30.28/0.8603	30.48/0.8628	31.53/0.8854	32.05/0.8910	32.46/0.8968	31.39/-	32.97/0.9004
Set14	$\times 2$	30.24/0.8688	32.28/0.9056	32.42/0.9063	33.04/0.9118	-/-	33.92/0.9195	32.83/-	34.25/0.9214
	$\times 3$	27.55/0.7742	29.13/0.8188	29.28/0.8209	29.76/0.8311	-/-	30.52/0.8462	29.72/-	30.83/0.8539
	$\times 4$	26.00/0.7027	27.32/0.7491	27.49/0.7503	28.02/0.7670	28.53/0.7804	28.80/0.7876	27.98/-	29.42/0.7984
B100	$\times 2$	29.56/0.8431	31.21/0.8863	31.36/0.8879	31.85/0.8942	-/-	32.32/0.9013	30.29/-	32.67/0.9081
	$\times 3$	27.21/0.7385	28.29/0.7835	28.41/0.7863	28.80/0.7963	-/-	29.25/0.8093	26.95/-	29.54/0.8112
	$\times 4$	25.96/0.6675	26.82/0.7087	26.90/0.7101	27.23/0.7233	27.57/0.7354	27.71/0.7420	25.16/-	28.23/0.7556
Urban100	$\times 2$	26.88/0.8403	29.20/0.8938	29.50/0.8946	30.75/0.9133	-/-	32.93/0.9351	-/-	33.31/0.9384
	$\times 3$	24.46/0.7349	26.03/0.7973	26.24/0.7989	27.15/0.8076	-/-	28.80/0.8653	-/-	29.12/0.8705
	$\times 4$	23.14/0.6577	24.32/0.7183	24.52/0.7221	25.14/0.7510	26.07/0.7839	26.64/0.8033	-/-	27.17/0.8129
DIV2K validation	$\times 2$	31.01/0.9393	32.89/0.9570	33.05/0.9581	-/-	-/-	35.03/0.9695	-/-	35.46/0.9731
	$\times 3$	28.22/0.8906	29.50/0.9116	29.64/0.9138	-/-	-/-	31.26/0.9304	-/-	31.53/0.9382
	$\times 4$	26.66/0.8521	27.70/0.8736	27.78/0.8753	-/-	-/-	29.25/0.9017	-/-	29.73/0.9076

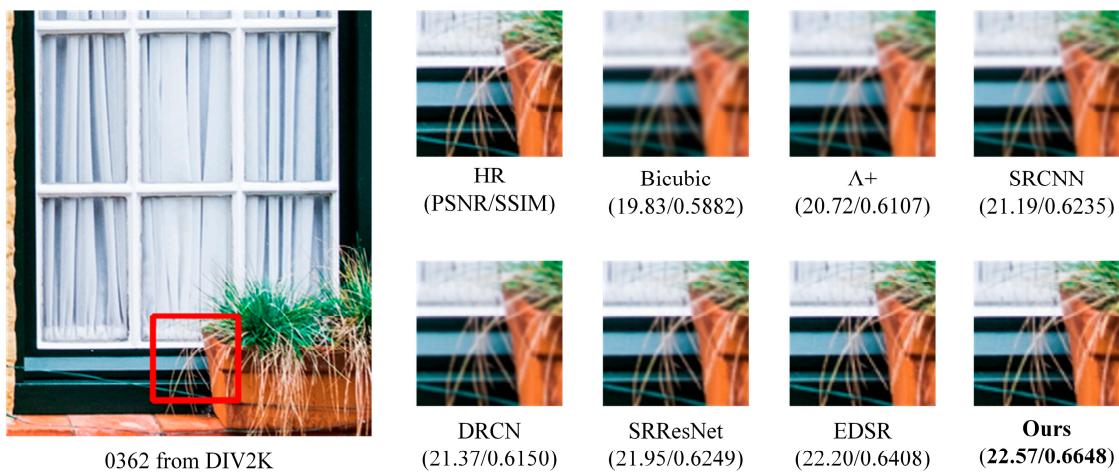


Figure 3. The results of “Image0362” (DIV2K) on $\times 4$ super-resolution. MRSR shows a better performance in the qualitative comparison with other methods.

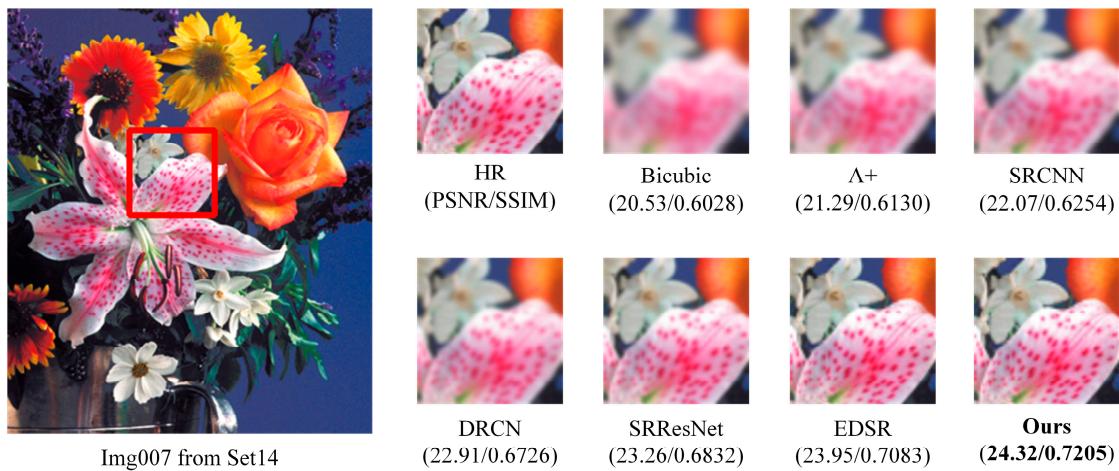


Figure 4. The results of “Image007” (Set14) on $\times 4$ super-resolution. MRSR shows a better performance in the qualitative comparison with other methods.

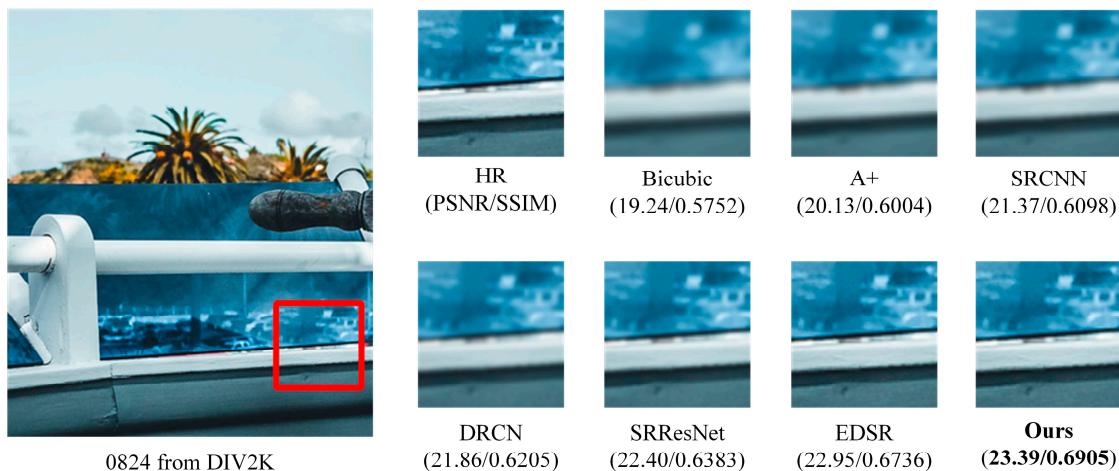


Figure 5. The results of “Image0824” (DIV2K) on $\times 4$ super-resolution. MRSR shows a better performance in the qualitative comparison with other methods.

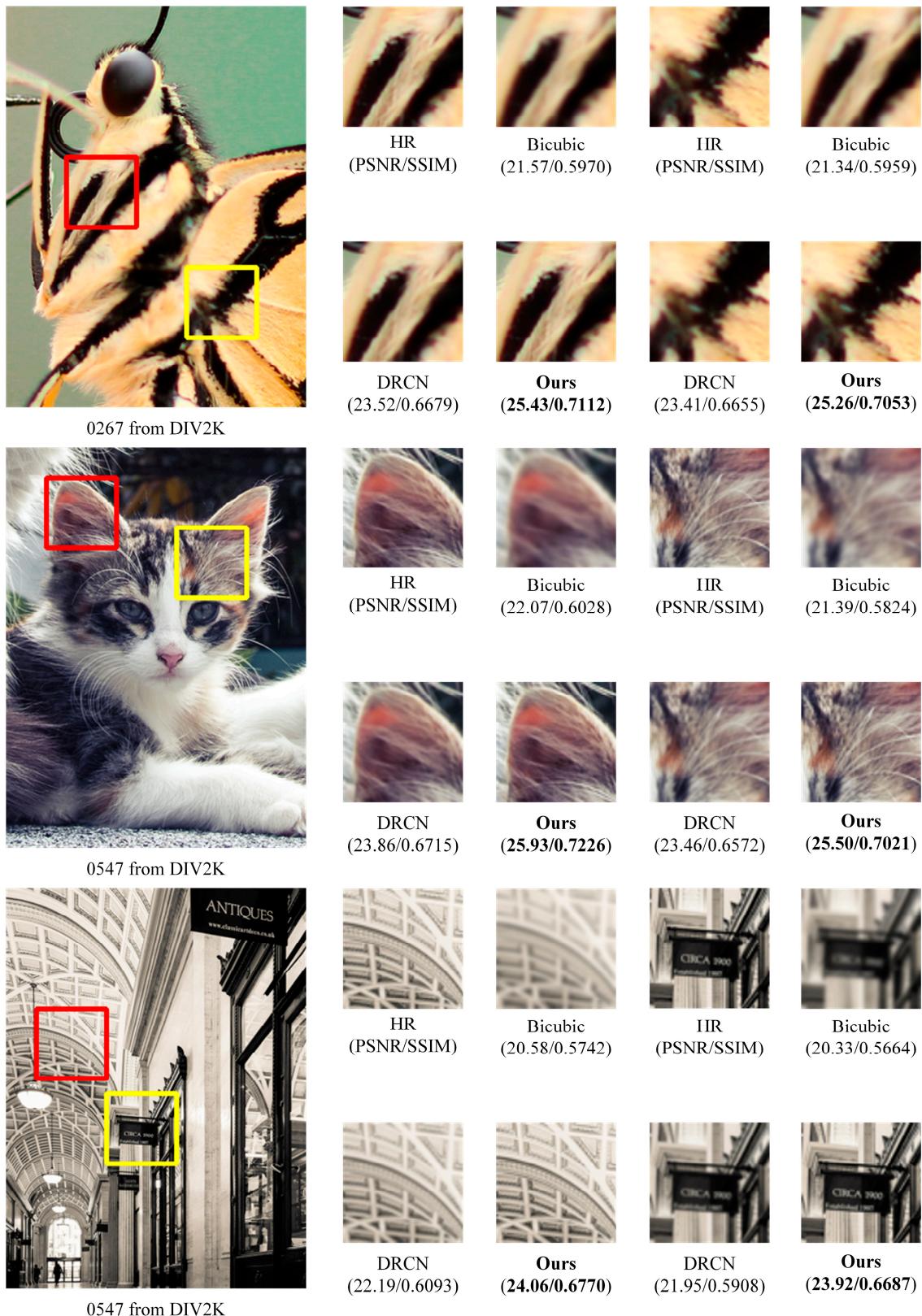


Figure 6. The qualitative super-resolution comparison between our work and DRCN with an upscaling factor of 4. Whether PSNR or SSIM, the SR results have been improved greatly.

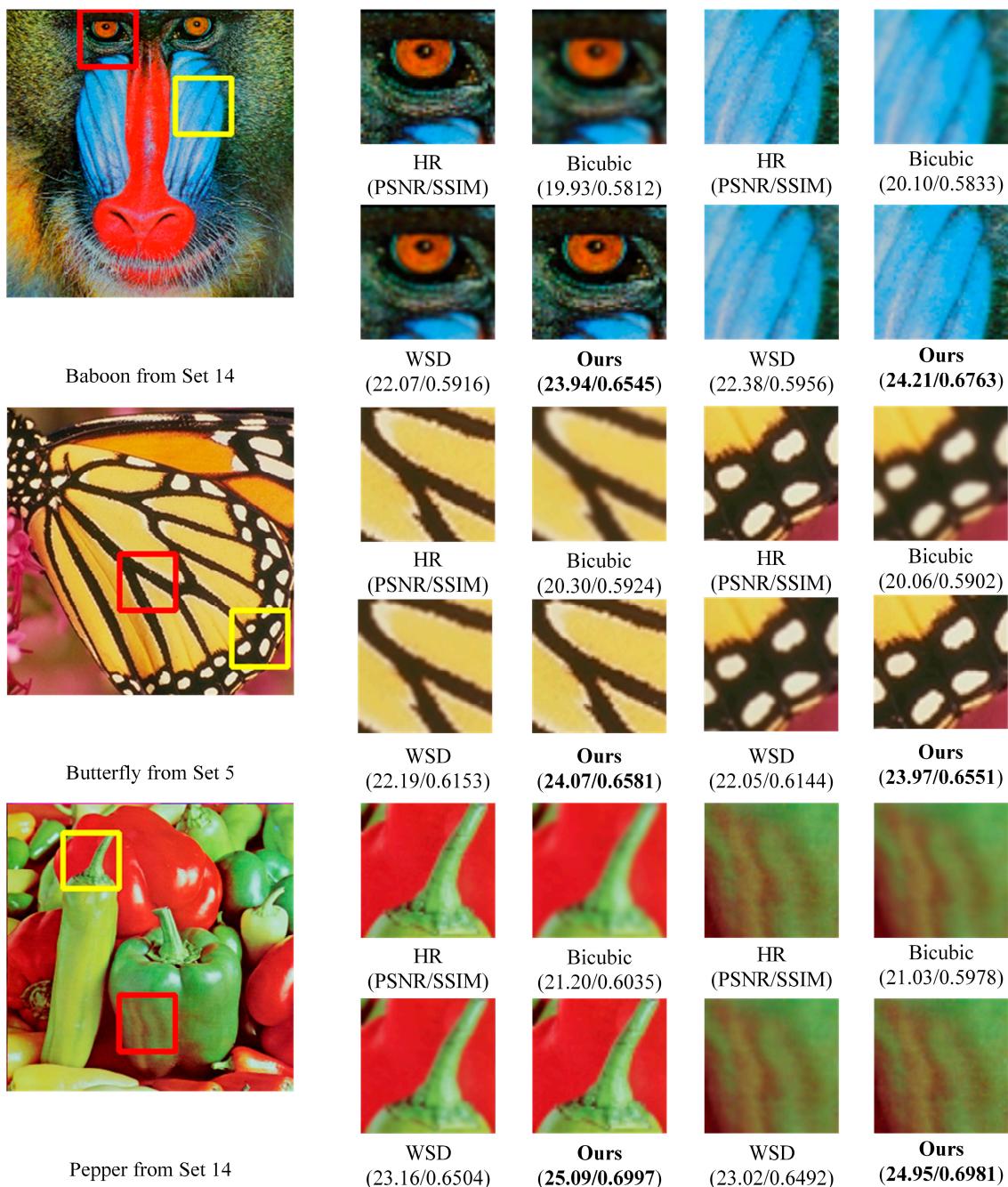


Figure 7. The qualitative super-resolution comparison between our work and WSD with an upscaling factor of 4. Whether PSNR or SSIM, the SR results have been improved greatly.

5. Conclusions

In this paper, we present a multi-branch deep residual algorithm for the SR problem. By optimizing the residual network, our model achieves better performance with fewer parameters and filters on all datasets. Coupled with the use of multi-branch networks, the training and convergence problems were partly solved. Due to the proposed supervision function, the reconstructed images show a better performance in the edge details and textures compared with other existing reconstruction methods. Furthermore, we develop a multi-scale SR residual network to achieve superior mapping between the LR and SR images by increasing the reduced-dimensional convolution layers in every two adjacent branch networks. The experiment results prove that the proposed approach achieves state-of-the-art

performance in terms of PSNR and SSIM. In the future, we will continue to improve our algorithms for a superior performance.

Author Contributions: Conceptualization, P.L.; Methodology, P.L.; Software, P.L.; Validation, P.L., Y.H.; Formal Analysis, P.L.; Investigation, P.L.; Resources, P.L.; Data Curation, P.L.; Writing-Original Draft Preparation, P.L.; Writing-Review & Editing, P.L.; Visualization, P.L.; Supervision, Y.H. and Y.L.; Project Administration, P.L.; Funding Acquisition, Y.H.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Irani, M.; Peleg, S. Improving resolution by image registration. *CVGIP Gr. Models Image Process.* **1991**, *53*, 231–239. [[CrossRef](#)]
2. Thornton, M.W.; Atkinson, P.M.; Holland, D.A. Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping. *Int. J. Remote Sens.* **2006**, *27*, 473–491. [[CrossRef](#)]
3. Peled, S.; Yeshurun, Y. Superresolution in MRI: Application to human white matter fiber tract visualization by diffusion tensor imaging. *Magn. Reson. Med.* **2001**, *45*, 29–35. [[CrossRef](#)]
4. Shi, W.; Caballero, J.; Ledig, C.; Zhuang, X.; Bai, W.; Bhatia, K.; de Marvao, A.M.; Dawes, T.; O'Regan, D.; Rueckert, D. Cardiac image super-resolution with global correspondence using multi-atlas patchmatch. *Med. Image Comput. Comput. Assist. Interv.* **2013**, *16*, 9–16. [[PubMed](#)]
5. Gunturk, B.K.; Batur, A.U.; Altunbasak, Y.; Hayes, M.H.; Mersereau, R.M. Eigenface-domain super-resolution for face recognition. *IEEE Trans. Image Process.* **2003**, *12*, 597–606. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, L.; Zhang, H.; Shen, H.; Li, P. A super-resolution reconstruction algorithm for surveillance images. *Signal Process.* **2010**, *90*, 848–859. [[CrossRef](#)]
7. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1646–1654.
8. Kim, J.; Lee, J.K.; Lee, K.M. Deeply-Recursive Convolutional Network for Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1637–1645.
9. Ledig, C.; Wang, Z.; Shi, W.; Theis, L.; Huszar, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the Computer Vision and Pattern Recognition, Venice, Italy, 30 October–1 November 2017; pp. 105–114.
10. Dong, C.; Chen, C.L.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)] [[PubMed](#)]
11. Cun, Y.L.; Boser, B.; Denker, J.S.; Howard, R.E.; Hubbard, W.; Jackel, L.D.; Henderson, D. Handwritten digit recognition with a back-propagation network. *Adv. Neural Inf. Process. Syst.* **1989**, *2*, 396–404.
12. Efrat, N.; Glasner, D.; Apartsin, A.; Nadler, B. Accurate Blur Models vs. Image Priors in Single Image Super-resolution. In Proceedings of the IEEE International Conference on Computer Vision, Harbour, Sydney, 1–8 December 2013; pp. 2832–2839.
13. He, H.; Siu, W.C. Single image super-resolution using Gaussian process regression. In Proceedings of the Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 449–456.
14. Yang, J.; Lin, Z.; Cohen, S. Fast Image Super-Resolution Based on In-Place Example Regression. In Proceedings of the Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 1059–1066.
15. Fernandezgranda, C.; Candes, E.J. Super-resolution via Transform-Invariant Group-Sparse Regularization. In Proceedings of the IEEE International Conference on Computer Vision, Harbour, Sydney, 1–8 December 2013; pp. 3336–3343.

16. Chang, H.; Yeung, D.Y.; Xiong, Y. Super-resolution through neighbor embedding. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; Volume 271, pp. 275–282.
17. Pan, Q.; Liang, Y.; Zhang, L.; Wang, S. Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. In Proceedings of the Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2216–2223.
18. Li, X.; Tao, D.; Gao, X.; Zhang, K. Multi-scale dictionary for single image super-resolution. In Proceedings of the Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 1114–1121.
19. Gao, X.; Zhang, K.; Tao, D.; Li, X. Image super-resolution with sparse neighbor embedding. *IEEE Trans. Image Process.* **2012**, *21*, 3194–3205. [PubMed]
20. Zhu, Y.; Zhang, Y.; Yuille, A.L. Single Image Super-resolution Using Deformable Patches. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 2917–2924.
21. Timofte, R.; Smet, V.D.; Gool, L.V. A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution. In Proceedings of the Asian Conference on Computer Vision, Singapore, 1–5 November 2014; pp. 111–126.
22. Dai, D.; Timofte, R.; Van Gool, L. Jointly Optimized Regressors for Image Super-resolution. *Comput. Graph. Forum* **2015**, *34*, 95–104. [CrossRef]
23. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
24. Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 1132–1140.
25. Tai, Y.W.; Liu, S.; Brown, M.S.; Lin, S. Super resolution using edge prior and single image detail synthesis. In Proceedings of the Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 2400–2407.
26. Timofte, R.; De, V.; Gool, L.V. Anchored Neighborhood Regression for Fast Example-Based Super-Resolution. In Proceedings of the IEEE International Conference on Computer Vision, Harbour, Sydney, 1–8 December 2013; pp. 1920–1927.
27. Allebach, J.; Wong, P.W. Edge-directed interpolation. In Proceedings of the International Conference on Image Processing, Lausanne, Switzerland, 19 September 1996; Volume 703, pp. 707–710.
28. Li, X.; Orchard, M.T. New edge-directed interpolation. *IEEE Trans. Image Process.* **2001**, *10*, 1521–1527. [PubMed]
29. Glasner, D.; Bagon, S.; Irani, M. Super-resolution from a single image. In Proceedings of the IEEE International Conference on Computer Vision, Lausanne, Switzerland, 19 September 1996; pp. 349–356.
30. Kim, K.I.; Kwon, Y. Single-Image Super-Resolution Using Sparse Regression and Natural Image Prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1127–1133. [PubMed]
31. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
32. Yang, J.; Wang, Z.; Lin, Z.; Cohen, S. Couple Dictionary Training for Image Super-resolution. *IEEE Trans. Image Process.* **2012**, *21*, 3467–3478. [CrossRef] [PubMed]
33. Shi, W.; Caballero, J.; Huszar, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1874–1883.
34. Johnson, J.; Alahi, A.; Li, F.F. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 694–711.
35. Bruna, J.; Sprechmann, P.; Lecun, Y. Super-Resolution with Deep Convolutional Sufficient Statistics. *arXiv* **2015**, arXiv:1511.05666.

36. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
37. Bevilacqua, M.; Roumy, A.; Guillemot, C.; Morel, A. Low-Complexity Single Image Super-Resolution Based on Nonnegative Neighbor Embedding. In Proceedings of the 23rd British Machine Vision Conference (BMVC), Surrey, UK, 3–7 September 2012.
38. Zeyde, R.; Elad, M.; Protter, M. On single image scale-up using sparse-representations. In Proceedings of the International Conference on Curves and Surfaces, Avignon, France, 24–30 June 2010; pp. 711–730.
39. Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In Proceedings of the Eighth IEEE International Conference on Computer Vision, Vancouver, BC, Canada, 7–14 July 2001; Volume 412, pp. 416–423.
40. Huang, J.B.; Singh, A.; Ahuja, N. Single image super-resolution from transformed self-exemplars. In Proceedings of the Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5197–5206.
41. Timofte, R.; Agustsson, E.; Gool, L.V.; Yang, M.H.; Zhang, L.; Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Ntire 2017 Challenge on Single Image Super-Resolution: Methods and Results. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 1110–1121.
42. Cruz, C.; Mehta, R.; Katkovnik, V.; Egiazarian, K. Single Image Super-Resolution based on Wiener Filter in Similarity Domain. *IEEE Trans. Image Process.* **2017**, *27*, 1376–1389. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).