

Article

An Efficient Super-Resolution Network Based on Aggregated Residual Transformations

Yan Liu ¹, Guangrui Zhang ¹, Hai Wang ^{2,*}, Wei Zhao ¹, Min Zhang ² and Hongbo Qin ¹

¹ Key Laboratory of Electronic Equipment Structure Design, Ministry of Education, Xidian University, Xi'an 710071, China; liuy@xidian.edu.cn (Y.L.); grzhang028@163.com (G.Z.); weizhao@xidian.edu.cn (W.Z.); qhb0920qhb@xidian.edu.cn (H.Q.)

² School of Aerospace Science and Technology, Xidian University, Xi'an 710071, China; minzhanghk@gmail.com

* Correspondence: wanghai@mail.xidian.edu.cn; Tel.: +86-029-8820-3115

Received: 30 January 2019; Accepted: 16 March 2019; Published: 20 March 2019



Abstract: In this paper, we propose an efficient multibranch residual network for single image super-resolution. Based on the idea of aggregated transformations, the split-transform-merge strategy is exploited to implement the multibranch architecture in an easy, extensible way. By this means, both the number of parameters and the time complexity are significantly reduced. In addition, to ensure the high-performance of super-resolution reconstruction, the residual block is modified and simplified with reference to the enhanced deep super-resolution network (EDSR) model. Moreover, our developed method possesses advantages of flexibility and extendibility, which are helpful to establish a specific network according to practical demands. Experimental results on both the Diverse 2K (DIV2K) and other standard datasets show that the proposed method can achieve a good performance in comparison with EDSR under the same number of convolution layers.

Keywords: super-resolution; residual network; aggregated transformations

1. Introduction

In recent years, single image super-resolution (SISR) has attracted a lot of attention from researchers in the field of computer vision. SISR aims to reconstruct a high-resolution image I^{HR} from a single low-resolution image I^{LR} [1], and it has been widely used in many fields, such as remote sensing [2], medical imaging [3], and environmental monitoring [4–7]. To our knowledge, the interpolation technique based on sampling theory was the earliest method to solve the super-resolution problem. However, there are serious shortages in predicting details and realistic textures. To address this problem, techniques that learn the mapping relationship between I^{LR} and I^{HR} have been proposed, such as neighbor embedding [8–11] and sparse coding [12–16]. In the last few years, deep learning-based approaches for super-resolution are constantly emerging [16–20]. Dong et al. first applied CNN (convolutional neural networks) into super-resolution [18], with a satisfactory effect in its practical use. Later, Kim et al. designed SRResNet (residual network for super-resolution) [20] based on the well-known residual network ResNet [19]. Benefiting from the jump connection and recursive structure, deeper layers are easy to realize for better performance. To simplify SRResNet, enhanced deep super-resolution network (EDSR) [1] was proposed for super-resolution by Lim et al., which optimizes the architecture of residual blocks by removing unnecessary modules. Although these ResNet-based models can improve the quality of reconstruction due to deeper layers, they all met the same problem: a sharp increase in the number of parameters. Especially in engineering practice, the cost of a large number of residual blocks and parameters has hampered the wider use

of ResNet-based models. Therefore, the question of how to reduce the number of model parameters without reconstruction quality loss has become one of the hottest research issues.

Nowadays, there are various methods reported to reduce the number of parameters [21–24]. Network pruning, SVD (singular value decomposition), and split–transform–merge strategy are three representative methods. In 1990, LeCun et al. first proposed the concept of network pruning, which decreased the model size by cutting off the redundant parameters of the neural network [21]. This method requires a lot of iterative training to ensure network performance. In 2014, Denton et al. proposed the SVD method to reduce the number of weights [22]. In the SVD method, the complex matrix is represented by multiplying smaller and simpler submatrices, which can significantly reduce network parameters. However, with the increase of the matrix scale, the calculation of the singular value becomes complicated and difficult. In recent years, the split–transform–merge strategy attracted more and more attention from researchers. Based on this strategy, the Inception models were developed with less computational complexity and a fewer number of parameters [23]. In the Inception models, the input is split into several low-dimensional embeddings (by 1×1 convolutions), then converted through a set of specialized filters (3×3 , 5×5 , etc.) and finally merged by connection [24]. However, because the hyper-parameters of each branch need to be set properly, it is hard to find a simple design method for the construction of an Inception network. In 2016, Xie et al. proposed the ResNeXt [24] network based on aggregated transformations, which can be regarded as the improvement of the split–transform–merge strategy. However, the ResNeXt was originally designed for image classification, therefore, its structure must be changed and optimized when applying it to super-resolution.

In this paper, an efficient multibranch residual network for the super-resolution task is proposed. The multibranch architecture is built on the basis of aggregated transformations. In the meantime, we optimize the residual block with reference to EDSR. According to the proposed network structure, two specific models are established and given as examples in this work. Experiments show that our models can achieve a good reconstruction quality with a significant reduction of network parameters.

2. Related Work

Inception: The Inception network is a typical multibranch architecture based on the split–transform–merge strategy. Each branch in the network is carefully designed to gain good performance in terms of speed and accuracy. However, the customized size and number of each filter in the branch make the Inception network hard to implement.

SRResNet: SRResNet is a super-resolution reconstruction network which is inspired by the residual network [20]. Based on the original residual structure, the network removes the active layer after the residual block and obtains a good image reconstruction result in human vision.

EDSR: EDSR is a state-of-the-art super-resolution network which further modifies the residual block structure based on SRResNet [1]. Since BN (batch normalization) layers get rid of the range flexibility from networks and consume a lot of memory, EDSR removes two BN layers in the residual block. Benefiting from the structural modification, EDSR has great improvements in image reconstruction and reduction in the usage of graphics processing unit (GPU) memory.

ResNeXt: Based on the residual block architecture, ResNeXt exploits the split–transform–merge strategy in an easy, extensible way—namely, aggregated residual transformations [24]. This method involves stacking a series of homogeneous, multibranch residual blocks with only a few hyperparameters to set [24]. Branches of ResNeXt each perform their set of convolutions and merge at the end of the block. Compared with ResNet, ResNeXt shows better performance and less computation complexity in the task of image classification.

Grouped convolution: Grouped convolution was first proposed in the AlexNet paper [25] in 2012. The given motivation by the author was to distribute the model over two GPUs to solve the limited hardware resources of a single GPU. Grouped convolution divides the feature maps into multiple GPUs for convoluting and subsequently aggregates the obtained results of multiple GPUs.

3. Methods

EDSR has achieved good results in the super-resolution field, but there is little improvement on the parameter quantity compared with other algorithms. To reduce the number of parameters, the aggregation transformation method is applied to EDSR in this paper. The aggregation transformation method, by which the multibranch architecture of networks can be built in an easy way, is originally presented in ResNeXt. This method can reduce the parameter and time complexity without significantly decreasing the accuracy of image classification.

A simple and obvious way to directly transform EDSR into multibranch architecture is by the aggregation transformation method. However, the original residual block of EDSR with two convolution layers is inconsistent with the aggregation transformation method [24]. This direct transformation would result in a wild and dense model, which not only has no benefit but adds more complexity. To solve this issue, we must redesign the model with multibranch architecture. Three or more convolution layers are required in the residual block of the new model. To simplify the structure of the residual block and enhance the feature extraction capability, we adopted three convolution layers in this work. Compared with the original residual block as shown in Figure 1a, our rebuilt residual block removes the unnecessary rectified linear unit (ReLU) and BN layers with reference to the EDSR structure. This removal operation helps improve the performance of image reconstruction.

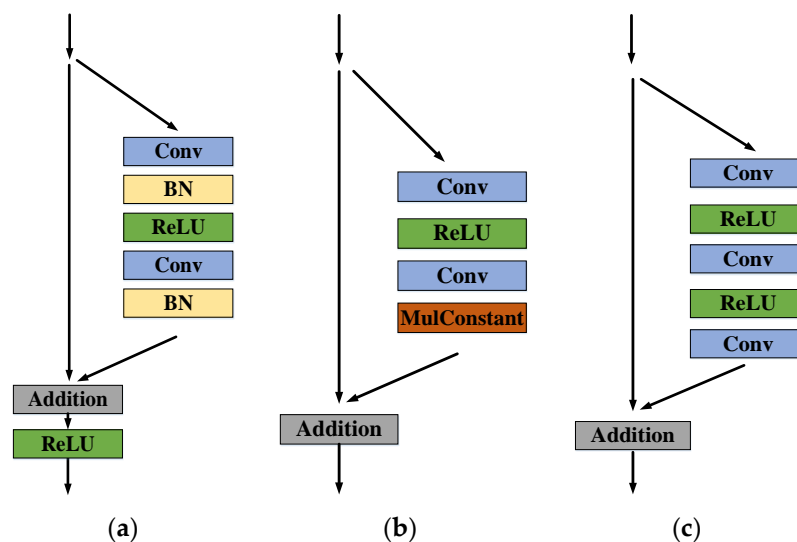


Figure 1. Comparison of residual blocks in the original ResNet, enhanced deep super-resolution network (EDSR), and our model. (a) Original ResNet residual block; (b) EDSR residual block; (c) Our proposed residual block.

As shown in Figure 1, the convolutional layer (Conv) was used to perform feature extraction, and ReLU to rectify the network output. The BN layer was used to normalize the features, and Addition represents the additional layer that the network adds as needed.

It is also known from the experiment by Lim et al. [1] that increasing the number of feature maps above a certain level would make the training process numerically unstable. The typical solution is to place a constant scaling layer (also called as MulConstant layer) after the last convolutional layer of each residual block. Owing to the use of aggregation transformations, the number of feature maps per convolution layer can be significantly reduced in comparison with the original EDSR model, therefore, the model proposed in this paper does not require the constant scaling layer. From the results in the following Experiment section, we can see that adding a constant scaling layer could worsen the performance. After removing the constant scaling layer, the architecture of our multibranch network is modeled and shown in Figure 2. The detailed description of ResBlock (residual block) has been given in Figure 1c. Upsample (upsampling structure) can magnify the image to the desired multiple.

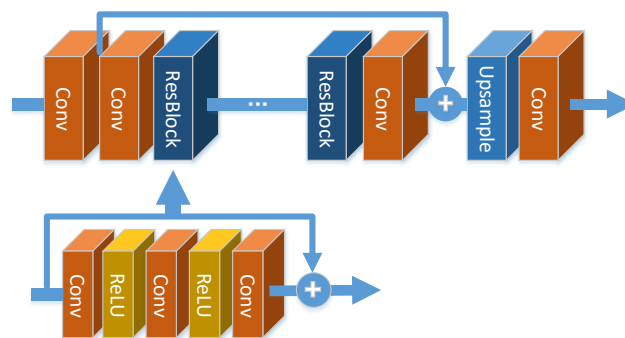


Figure 2. The architecture of the proposed multibranch network.

As shown in Figure 3, we design with different configurations for our multibranch architecture: EDSRSP- 3×3 and EDSRSP- 1×1 . The number represents the size of the first and third convolution kernel. The configuration of the residual block in EDSRSP- 3×3 is as the same as that in EDSR, i.e., 3×3 convolution kernel, 256-d input and 256-d output. It is seen from Table 1 that the number of parameters in EDSRSP- 3×3 is reduced by 1/3 compared with EDSR. To further decrease the parameters, the configuration of EDSRSP- 1×1 is properly adjusted and shown in Figure 3b. The detailed adjustments include using the 1×1 convolution kernel in the first and third layers and the 512-d input and output in the second layer. EDSRSP- 1×1 is similar to the bottleneck structure of ResNet, only with a little modification on the output dimension in the first layer. Due to the use of a 1×1 convolution kernel, the number of parameters in EDSRSP- 1×1 are reduced to 1/4 of those in EDSR.

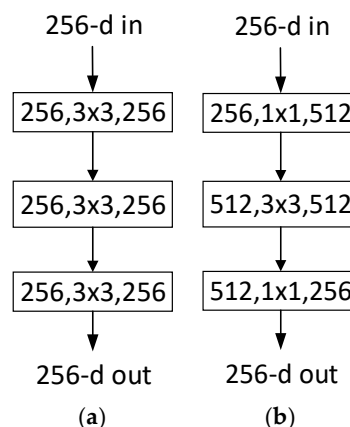


Figure 3. Proposed models. (a) EDSRSP- 3×3 . (b) EDSRSP- 1×1 .

Table 1. Parameters of EDSR and our models.

Model	Number of Residual Blocks	Total Parameters of Residual Blocks
EDSR		
256, 3×3 , 256	32	~32,749 K
256, 3×3 , 256		
EDSRSP- 3×3		
256, 3×3 , 256	21	~25,160 K
256, 3×3 , 256, 32		
256, 3×3 , 256		
EDSRSP- 1×1		
256, 1×1 , 512	21	~7053 K
512, 3×3 , 512, 32		
512, 1×1 , 256		

For the implementation of aggregation transformation, our model has two equivalent structures as shown in Figure 4. The two structures have the same-level reconstruction performance, but the structure based on group convolution (Figure 4b) has the distinct advantages of time complexity and memory usage. Therefore, we use group convolution to realize the aggregation transformation.

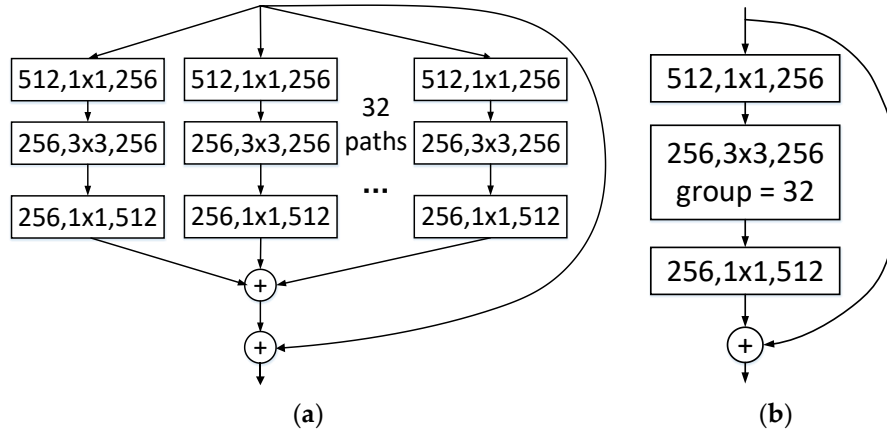


Figure 4. Equivalent building blocks of EDSRSP-1x1. (a) Aggregated residual transformations. (b) A block equivalent to (a), implemented as grouped convolutions.

4. Experiment

4.1. Datasets

For our experiment, the newly proposed Diverse 2K (DIV2K) dataset [26] is used due to its high-quality (2K) resolution for the image reconstruction tasks. The DIV2K dataset consists of 800 training images, 100 validation images, and 100 test images. Since the test dataset ground truth has not been published, the performance comparison was made on the validation dataset. We also compared the performance on three standard benchmark datasets: Set5 [9], Set14 [12], and B100 [27].

4.2. PSNR and SSIM Criteria

Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are the two most-used indicators in the field of super-resolution reconstruction, which can measure the similarity between the reconstructed image and the original high-resolution image [28,29]. The mathematical expression of PSNR is as follows:

$$\text{PSNR} = 10 \log \left[\frac{(2^n - 1)^2}{\text{MSE}} \right], \quad (1)$$

where n is the number of bits per pixel, and mean square error (MSE) is defined as shown below:

$$\text{MSE} = \frac{1}{MN} \sum \sum [f(i, j) - f'(i, j)]^2, \quad (2)$$

where $f(i, j)$ and $f'(i, j)$ represent the original and reconstructed images, respectively. Both of them are of size $M \times N$, and (i, j) stands for the pixel coordinate. The larger the value of PSNR, the better effect of image reconstruction.

SSIM is another popular criteria to compare the reconstructed image x and the original high-definition image y . The formula of SSIM is shown as follows:

$$\text{SSIM}(x, y) = \frac{(2u_x u_y + c_1)(2\sigma_{xy} + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (3)$$

where u_x, u_y are the mean value of x, y . σ_x^2, σ_y^2 are the variance of x, y . σ_{xy} is the covariance of x and y . $c_1 = (k_1 L)^2$ and $c_2 = (k_2 L)^2$ are constants to maintain formula validity, avoiding the denominator being zero. L represents the dynamic range of the pixel value. $k_1 = 0.01$ and $k_2 = 0.03$ by default. The larger the value of SSIM, the better similarity of the two images.

4.3. Training Details

For training, we use and adjust the training parameters given in Lim et al. [1]. Neither the pre-training model nor the geometric self-ensemble strategy is used in this training. The chop size is set to 4.0×10^4 and patch sizes of $3/4$ were set to 96. We also learnt from the code published by the EDSR paper and trained the models by using NVIDIA Titan Xp GPUs. According to the official baseline model, the used EDSR model is retrained with no modifications other than those mentioned above. It takes seven days to train EDSR compared with three days for our models.

4.4. Comparison between the Cases with and without MulConstant Layer

To analyze the effect of the MulConstant Layer in our designed residual block, we performed experiments on the EDSRSP- $1 \times 1 \times 4$ model and the EDSRSP- $3 \times 3 \times 2$ model. The three experiments correspond to three different cases: (1) without the MulConstant layer; (2) MulConstant layer with the factor set to 0.1; (3) MulConstant layer with the factor set to 0.01. From the experimental results as shown in Figure 5, we can see that removing the MulConstant layer in our model results in better performance.

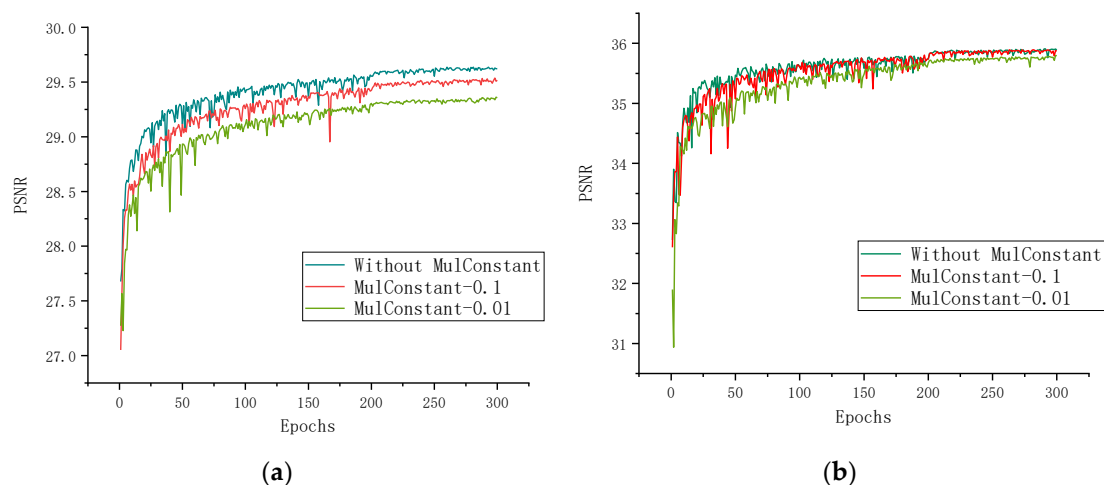


Figure 5. (a) Peak signal-to-noise ratio (PSNR) validation of EDSRSP- $1 \times 1 \times 4$ models. (b) PSNR validation of EDSRSP- $3 \times 3 \times 2$ models.

4.5. Evaluation on DIV2K Dataset

For the performance evaluation, a comparison between the retrained EDSR model and our model is made and shown in Figure 6. The detailed evaluation method is given and described in Lim et al. [1]. Using PSNR and SSIM criteria, the evaluation is conducted on 10 images of the DIV2K validation set. Concretely, we use full RGB channels and ignore the (6 + scale) pixels from the border. The small difference between EDSR and our models could verify the performance of the proposed method.

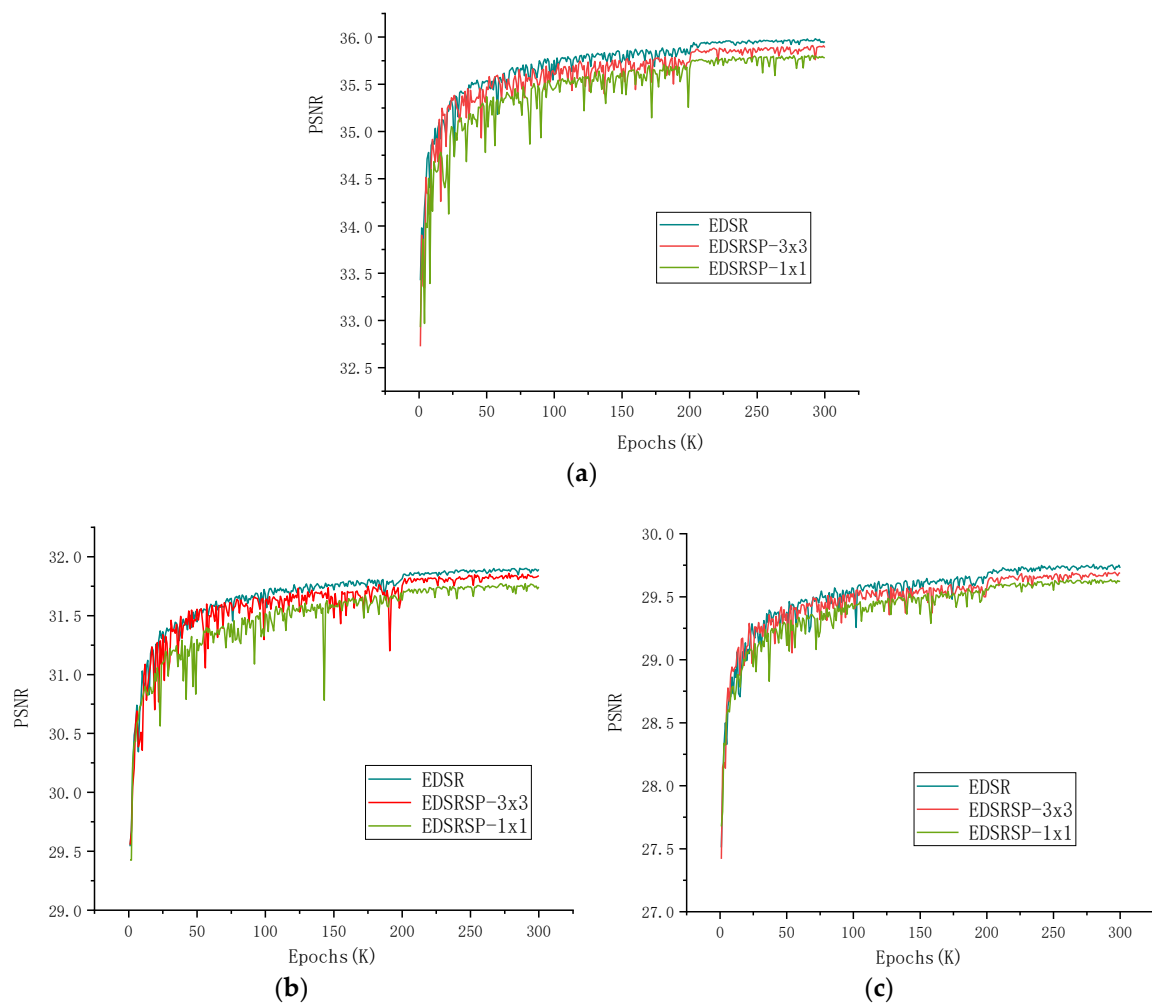


Figure 6. (a) Validation PSNR of EDSR $\times 2$ model and proposed $\times 2$ models. (b) Validation PSNR of EDSR $\times 3$ model and proposed $\times 3$ models. (c) Validation PSNR of EDSR $\times 4$ model and proposed $\times 4$ models.

Table 2 gives PSNR and SSIM scores of EDSR and our models on the DIV2K validation set, where the results are consistent with those in Figure 6. In addition, visual comparisons of the super-resolution images are shown in Figure 7. It can be seen, intuitively, that our models show high quality regardless of details or textures.

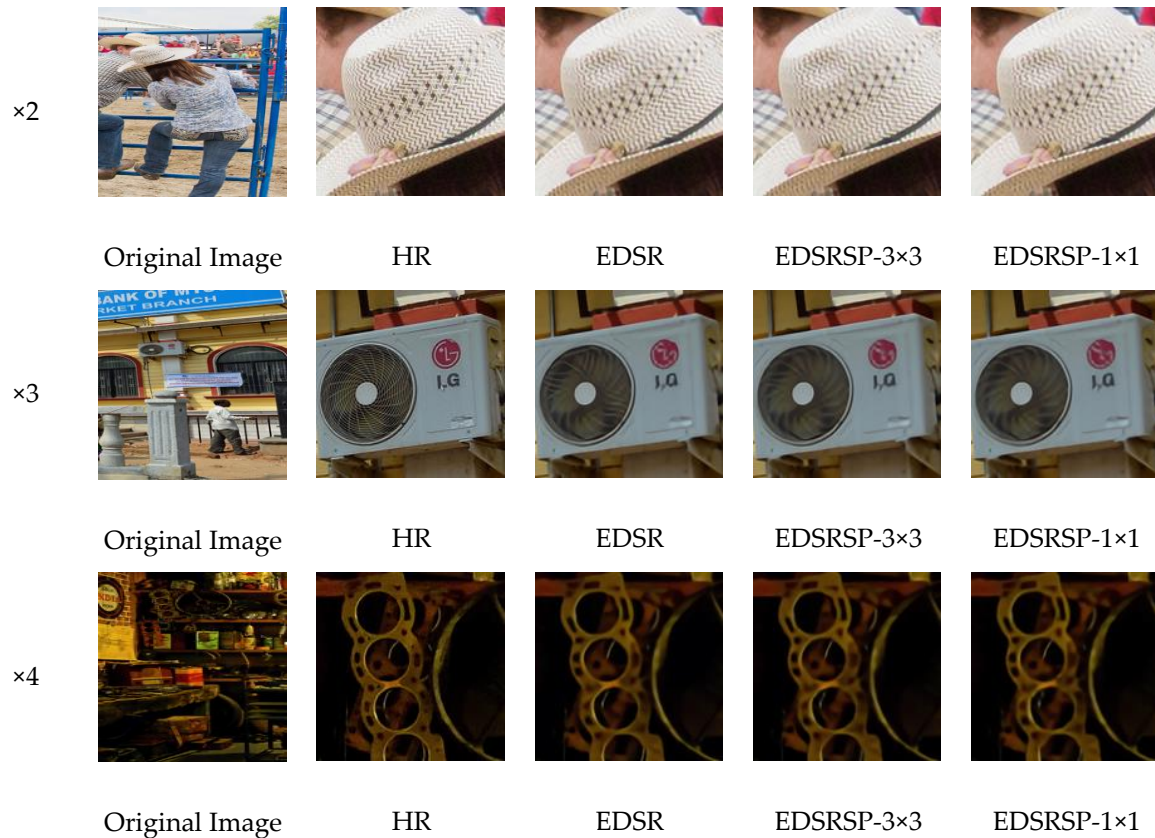
Table 2. Performance comparison between architectures on the DIV2K validation set (PSNR (dB)/SSIM).

Dataset	Scale	EDSR	EDSRSP-3 \times 3	EDSRSP-1 \times 1
DIV2K	$\times 2$	35.80/0.9676	35.71/0.9673	35.60/0.9670
	$\times 3$	32.17/0.9345	32.06/0.9337	31.99/0.9331
	$\times 4$	30.07/0.9057	29.97/0.9050	29.88/0.9045

We also performed the running time test on the pictures in Figure 7. The experimental results are shown in Table 3. As can be seen from the data in the table, the proposed model has a faster running time than EDSR.

Table 3. Running time (s) comparison between EDSR and proposed models.

Scale	EDSR	EDSRSP-3×3	EDSRSP-1×1
×2	12.562	9.966	6.472
×3	7.700	6.348	4.665
×4	4.426	3.363	2.442

**Figure 7.** Super-resolution reconstruction results on the DIV2K dataset.

4.6. Evaluation on Other Datasets

More experiments were implemented on the standard datasets of B100, Set5, and Set14. For comparison, we measured PSNR and SSIM on the y-channel, ignoring the same number of pixels as the boundary scaling. The MATLAB code was provided by the EDSR paper for this evaluation. As can be seen from Table 4, our models can achieve the same level performance as EDSR with fewer parameters, in theory.

Table 4. Public benchmark test results (PSNR (dB)/SSIM).

Dataset	Scale	EDSR	EDSRSP-3×3	EDSRSP-1×1
Set5	×2	38.08/0.960	38.04/0.9599	37.99/0.9598
	×3	34.59/0.9275	34.48/0.9267	34.40/0.9261
	×4	32.36/0.8950	32.21/0.8937	32.15/0.8926
Set14	×2	33.71/0.9185	33.65/0.9180	33.58/0.9169
	×3	30.35/0.8435	30.32/0.8428	30.24/0.8412
	×4	28.60/0.7831	28.57/0.7821	28.51/0.7809
B100	×2	32.30/0.9009	32.24/0.9004	32.20/0.8995
	×3	29.20/0.8080	29.16/0.8067	29.12/0.8055
	×4	27.64/0.7390	27.60/0.7378	27.57/0.7366

It can be seen from the experimental results that under the premise of ensuring the reconstruction quality, the proposed models have obvious advantages in time complexity and space complexity. This also means a reduction in the demand for hardware resources in practical applications, which makes our models easier to implement in real conditions.

5. Conclusions

In this paper, we propose an efficient super-resolution network based on aggregated residual transformations. Based on the proposed network, two specific models were designed and built in this work. Each of the two models has its own advantages regarding the reconstruction performance and the number of parameters. Experiments on both the DIV2K and other standard datasets were implemented to evaluate the performance of our network. The experiment results proved that our method is effective and easy to implement. Compared with EDSR, the number of parameters is significantly reduced with the same-level performance.

Author Contributions: Conceptualization, G.Z.; Methodology, G.Z.; Investigation, G.Z. and H.W.; Writing—Original Draft, G.Z. and W.Z.; Writing—Review & Editing, G.Z. and Y.L.; Project Administration, G.Z. and H.W.; Supervision, H.W. and Y.L.; Software, M.Z. and H.Q.

Funding: This research is supported by the China Postdoctoral Science Foundation No. 2018M633471 and the AeroSpace T.T. and C. Innovation Program.

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

References

1. Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 June 2017.
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. Greenspan. Super-resolution in medical imaging. *Comput. J.* **2008**, *52*, 43–63. [[CrossRef](#)]
4. Matuszewski, J.; Sikorska-Lukasiewicz, K. Neural network application for emitter identification. In Proceedings of the 18th International Radar Symposium (IRS), Prague, Czech Republic, 28–30 June 2017.
5. Dudczyk, J.; Kawalec, A. Adaptive forming of the beam pattern of microstrip antenna with the use of an artificial neural network. *Int. J. Antenn. Propag.* **2012**, *2012*, 13. [[CrossRef](#)]
6. Dudczyk, J. A method of feature selection in the aspect of specific identification of radar signals. *Bull. Pol. Acad. Sci.-Tech.* **2017**, *65*, 113–119. [[CrossRef](#)]
7. Pietrow, D.; Matuszewski, J. Objects detection and recognition system using artificial neural networks and drones. In Proceedings of the 2017 Signal Processing Symposium (SPSymo), Jachranka Village, Poland, 12–14 September 2017.
8. Chang, H.; Yeung, D.-Y.; Xiong, Y. Super-resolution through neighbor embedding. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Washington, DC, USA, 27 June–2 July 2004.
9. Bevilacqua, M.; Roumy, A.; Guillemot, C.; Alberi-Morel, M.L. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In Proceedings of the 23rd British Machine Vision Conference Location (BMVC), Guildford, UK, 3–7 September 2012.
10. Gao, X.; Zhang, K.; Tao, D. Image super-resolution with sparse neighbor embedding. *IEEE Trans. Image Process* **2012**, *21*, 3194–3205. [[PubMed](#)]
11. Roweis, S.T.; Saul, L.K.J.S. Nonlinear dimensionality reduction by locally linear embedding. *Science* **2000**, *290*, 2323–2326. [[CrossRef](#)] [[PubMed](#)]
12. Zeyde, R.; Elad, M.; Protter, M. On single image scale-up using sparse-representations. In Proceedings of the International Conference on Curves and Surfaces (ICCS), Avignon, France, 24–30 June 2010.
13. Yang, J.; Wang, Z.; Lin, Z.; Cohen, S.; Huang, T. Coupled dictionary training for image super-resolution. *IEEE T. Image Process* **2012**, *21*, 3467–3478. [[CrossRef](#)] [[PubMed](#)]

14. Timofte, R.; De Smet, V.; Van Gool, L. A+: Adjusted anchored neighborhood regression for fast super-resolution. In Proceedings of the Asian Conference on Computer Vision (ACCV), Singapore, 1–2 November 2014.
15. Yang, J.; Wright, J.; Huang, T.S. Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **2010**, *19*, 2861–2873. [[CrossRef](#)] [[PubMed](#)]
16. Kim, J.; Kwon Lee, J.; Mu Lee, K. Deeply-recursive convolutional network for image super-resolution. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 27–30 June 2016.
17. Kim, J.; Kwon Lee, J.; Mu Lee, K. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 27–30 June 2016.
18. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In Proceedings of the European Conference on Computer Vision (ECCV), Zurich, Switzerland, 6–12 September 2014.
19. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 27–30 June 2016.
20. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 June 2017.
21. Hassibi, B.; Stork, D.G. Second order derivatives for network pruning: Optimal brain surgeon. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Denver, CO, USA, 1 April 1993.
22. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311. [[CrossRef](#)]
23. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), San Francisco, CA, USA, 4–9 February 2017.
24. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 June 2017.
25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Harrahs and Harveys, Lake Tahoe, NV, USA, 3–8 December 2012.
26. Timofte, R.; Agustsson, E.; Van Gool, L.; 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 Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 June 2017.
27. 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 IEEE International Conference on Computer Vision (ICCV), Vancouver, BC, Canada, 9–12 July 2001.
28. Huynh-Thu, Q.; Ghanbari, M. Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **2008**, *44*, 800–801. [[CrossRef](#)]
29. Wang, Z.; Bovik, A.C.; Sheikh, H.R. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Signal Process.* **2004**, *13*, 600–612. [[CrossRef](#)]

