



# A-ESRGAN: Training Real-World Blind Super-Resolution with Attention U-Net Discriminators

Zihao Wei<sup>1,2(✉)</sup>, Yidong Huang<sup>1,2</sup>, Yuang Chen<sup>1,2</sup>, Chenhao Zheng<sup>1,2</sup>, and Jingnan Gao<sup>2</sup>

<sup>1</sup> University of Michigan, Ann Arbor, MI 48109, USA

[zihaowei@umich.edu](mailto:zihaowei@umich.edu)

<sup>2</sup> Shanghai Jiao Tong University, Shanghai 200240, China

[owenhji,cyaa,neymar}@umich.edu](mailto:{owenhji,cyaa,neymar}@umich.edu); [gjn0310@sjtu.edu.cn](mailto:gjn0310@sjtu.edu.cn)

**Abstract.** Generative adversarial networks (GANs) have recently made great progress in blind image super-resolution (SR) with their superiority in learning mappings between manifolds, which benefits the reconstruction of image's textural details. Recent works have largely focused on designing more realistic degradation models, or constructing a more powerful generator structure but neglected the ability of discriminators in improving visual performances. In this paper, we present A-ESRGAN, a GAN model for blind SR tasks featuring an attention U-Net based, multi-scale discriminator that can be seamlessly integrated with other generators. To our knowledge, this is the first work to introduce attention U-Net structure as the discriminator of GAN to solve blind SR problems. And the paper also gives an interpretation of the mechanism behind multi-scale attention U-Net that brings performance breakthrough to the model. Experimental results demonstrate the superiority of our A-ESRGAN over state-of-the-art level performance in terms of quantitative metrics and visual quality. The code can be find in <https://github.com/stroking-fishes-ml-corp/A-ESRGAN>.

**Keywords:** Blind Super Resolution · Generative adversarial networks · attention mechanism · Multi-scale · U-Net

## 1 Introduction and Motivation

Image super-resolution (SR) is a low-level computer vision problem aiming to reconstruct a high-resolution (HR) image from a distorted low-resolution (LR) image. Blind super-resolution, specifically, refers to the idea of restoring LR images suffering from unknown and complex degradation, as opposed to the traditional assumption of ideal bicubic degradation.

By the competition of generator and discriminator, the networks are encouraged to favor solutions that look more like natural images. The state-of-the-art methods using generative adversarial network includes ESRGAN, RealSR, Real-ESRGAN and BSRGAN [6, 20, 22, 25].

Recent works in super-resolution GANs have largely focused on simulating a more complex and realistic degradation process [20] or building a better generator [22], with little work trying to improve the performance of the discriminator. However, the importance of a discriminator can not be ignored since it provides the generator with the direction to generate better images, similar to a loss function. In this work, we construct a new discriminator network structure: **Multi-scale Attention U-Net Discriminator** and incorporate it with the existing RRDB based generator [22] to form our GAN model A-ESRGAN. Our model shows superiority over the state-of-the-art real-ESRGAN model in sharpness and details (see 7b). According to our ablation study, this result owes to the combination of attention mechanism and U-Net structure in our proposed discriminator. The U-Net structure in discriminator can offer per-pixel feedback to the generator [17], which can help the generator to generate more detailed features, such as texture or brushstroke. Meanwhile, the attention layer can not only distinguish the outline of the subject area to maintain the global coherence but strengthen the lines and edges of the image to avoid the blurring effect (this is demonstrated in the attention map analysis section in our paper). Therefore, the combination of U-Net and Attention is very promising. Besides, to increase the perception field of our discriminator, We use two attention U-Net discriminators that have an identical network structure but operate at different image scales as our final discriminator, which is called multi-scale discriminator. Extensive experiments show that our model outperforms most existing GAN models both in quantitative NIQE performance metric and qualitative image perceptual feelings.

In summary, the contributions of our work are:

1. We propose a new multi-scale attention U-Net discriminator network. To the best of our knowledge, it is the first work to adopt attention U-Net structure as a discriminator in the field of generative adversarial networks. This modular discriminator structure can be easily ported to future work.
2. We incorporate our designed discriminator with the existing RRDB based generator to form our generative adversarial network model A-ESRGAN. Experiments show that our model outperforms most state-of-the-art models in image super-resolution tasks.
3. Through detailed analysis and visualization about different layers of our network, we provide convincing reasons why a multi-scale attention U-Net discriminator works better than existing ones in image super-resolution tasks.

## 2 Related Work

### 2.1 GANs-Based Blind SR Methods

GANs-based SR methods have better perceptual results than CNN-based SR methods, because GANs are more competitive in learning mapping between manifolds, which benefits in reconstructing local textures [10]. Recent state of the art methods have raised a perceptual-driven perspective to improve GANs by

better modeling the perceptual loss between images [9, 22]. The ESRGAN [22], as a representative work, proposed a practical perceptual loss function as well as a residual-in-residual block(RRDB) generator network, and produces synthesized HR images with convincing visual quality. Another perspective is to solve the intrinsic problem of blind SR that the LR images used for training are synthesized from HR images in the dataset. Most existing methods are based on bicubic downsampling [3, 8, 16] and traditional degradations [24, 26], while real-world degradations are far more complicated. To produce more photo-realistic results, the real-ESRGAN [20] proposed a practical high-order degradation model and achieved visually impressive results as well as state-of-the-art NIQE [12] performance.

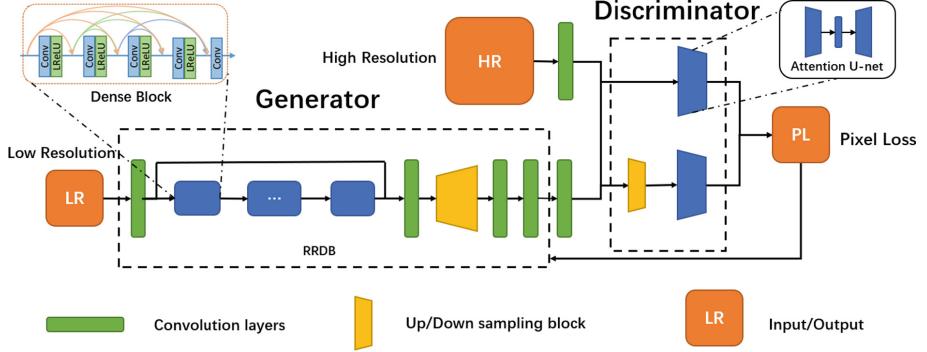
## 2.2 Discriminator Models

Some remarkable attempts have been made to improve the discriminator model [15, 17, 19]. To synthesize photo-realistic HR images, two major challenges are presented: the discriminator needs a large receptive field to differentiate the synthesized image and the ground truth(GT), requiring either deep network or large convolution kernel [19]. Besides, it's difficult for one discriminator to give precise feedback on both global and local features, leading to possible incoherence in the synthesized image such as twisted textures on a building wall [20]. To resolve these issues, Wang et al. [19] proposed a novel multiple discriminator architecture. By using several discriminators taking different scale down-sampled synthesized images as input, the new discriminator can learn from different receptive fields. Another pioneer work [17] introduces U-Net based discriminator architecture into GANs-based blind SR tasks. The U-Net discriminator model can provide per-pixel feedback to the generator while maintaining the global coherence of synthesized images.

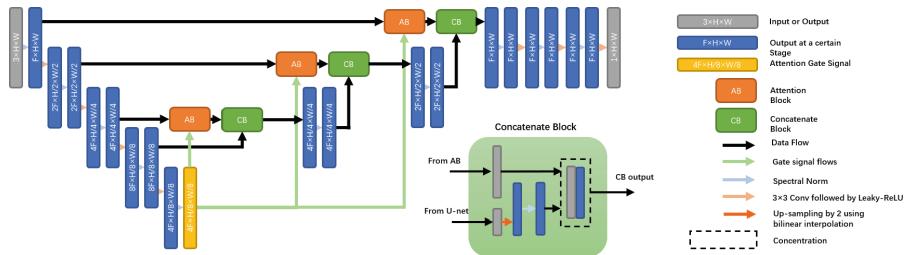
## 3 Method

The overall architecture of A-ESRGAN is shown in Fig. 1, which contains a Generator composed of residual-in-residual dense blocks (RRDBs) [22] and a multi-scale attention U-net discriminator.

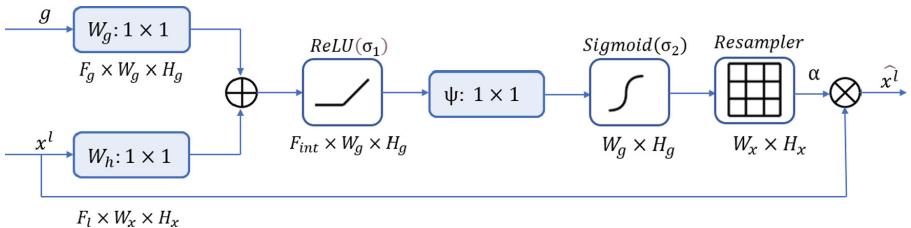
**Attention U-Net Discriminator.** Inspired by [14, 17], we propose the attention U-Net discriminator structure, which is shown in Fig. 2. It is composed of a down-sampling encoding module, an up-sampling decoding module and several attention blocks. The structure of the attention block in Fig. 3 is modified from attention gate, which is used by [14] in 3D medical graphs. By utilizing attention blocks, we want the generator to put more emphasis on important regions during super resolution, which also conforms to how people evaluate a graph as they would focus more on the region where their attention locates. A detailed analysis of how the attention mechanism works is carried out during the experiment. In order to stabilize the training process, we also apply spectral normalization regularization [13] to the attention U-net.



**Fig. 1.** The overall architecture of the A-ESRGAN. The generator of A-ESRGAN is using RRDB and the discriminator of A-ESRGAN is a multi-scale attention U-net structure.



**Fig. 2.** The architecture of a single attention U-Net Discriminator. F, W, H represents output channel number of the first convolution layer, height of the image and width of the image respectively.



**Fig. 3.** The architecture of the attention block (AB). Here  $x^l$  is the input features from the U-Net and  $g$  is the gating signal.  $F_{int}$  is a super parameter denoting the output channels of the one by one convolution in the AB. In the AB,  $x^l$  is scaled by attention coefficient  $\alpha$ .

For a single U-net discriminator, we construct the loss of the discriminator as:

$$L_D(G, D) = \sum_{w=1}^W \sum_{h=1}^H (-E_{x_r}[\log(D(x_r)[w, h])] - E_{x_i}[\log(1 - D(G(x_i))[w, h])]) \quad (1)$$

where  $D(x) = \sigma(C(x))$  is the discriminator's output on the paired data  $x = (x_r, G(x_i))$  after normalization by the sigmoid function;  $[w, h]$  denotes the value of matrix  $D$  on the  $w$ -th row and  $h$ -th column;  $x_r$  is the real (SR) image;  $G(x_i)$  is the generator output based on the corresponding LR image  $x_i$ . It is worth noting that the discriminator will give us pixel-wise loss  $C(x)$ , which means  $C(x)$  is a  $W \times H$  matrix with the same size as the  $G(x_i)$ .

Similarly we can derive the adversarial Generator Loss as:

$$L_G(G, D) = \sum_{w=1}^W \sum_{h=1}^H (-E_{x_i}[\log(D(G(x_i))[w, h])]) \quad (2)$$

**Multi-scale Discriminator.** A-ESRGAN adopts a multiple discriminator architecture that has two identical attention U-Nets as the discriminator. One discriminator  $D_1$  takes an original scale image as input and the other discriminator  $D_2$  takes a  $2 \times$  downsampled image as input. Thus, the overall objective function of the multiscale discriminator is the weighted average of the loss of sub-discriminators:

$$L_{D_{multi}} = \lambda_1 L_D(G, D_1) + \lambda_2 L_D(G, D_2) \quad (3)$$

where  $\lambda_1$  and  $\lambda_2$  are the weight coefficients which denote how much each discriminator contributes to the overall loss.

Likely the overall generator loss is

$$L_{G_{multi}} = \lambda_1 L_G(G, D_1) + \lambda_2 L_G(G, D_2) \quad (4)$$

In the later experiments, we find this setting helps acquire complementary knowledge, which helps generate clearer textures and outlines.

**Improved Generator Loss.** Following previous GAN-based SR methods, we also add  $L_1$  loss and perceptual loss [7] to better tune the generator.

Thus, our finally loss function for generator is:

$$L_{G_{Total}} = L_{precep} + \eta L_1 + \lambda L_{G_{multi}} \quad (5)$$

where  $\lambda, \eta$  are the weight coefficients for each loss.

**Overall Objective.** With the above discussions, our full objective is to solve the following task:

$$(D_1, D_2), G = \arg \min_{D_1, D_2} L_{D_{multi}}, \arg \min_G L_{G_{Total}} \quad (6)$$

where  $D_1, D_2$  and  $G$  are trained simultaneously.

## 4 Experiments

### 4.1 Implementation Details

To compare the functionality of multi-scale mechanism, we build two A-ESRGAN models: A-ESRGAN-single and A-ESRGAN-multi. The difference is that A-ESRGAN-single features one single attention U-Net discriminator, while A-ESRGAN-multi features multi-scale network, i.e. two identical attention U-Net discriminator operating at different image scale.

We trained with our A-ESRGAN on DIV2K [1] dataset. For better comparison with Real-ESRGAN, we follow the setting of generating degradation images of Real-ESRGAN [20] and load the pre-trained Real-ESRNET to the generator of both networks. The training HR patch size is 256. We train our models with one NVIDIA A100 and three NVIDIA A40 with a total batch size of 48 by using Adam optimizer.

The A-ESRGAN-Single is trained with a single attention U-Net discriminator for  $400K$  iterations under  $10^{-4}$  rate. The A-ESRGAN-Multi is trained for  $200K$  iterations under  $10^{-4}$  learning rate. For both A-ESRGAN-Single and A-ESRGAN-Multi, the weight for L1loss, perceptual loss and GAN loss are  $\{1, 1, 0.1\}$ . In A-ESRGAN-Multi, the weight for GAN loss of  $D_1$  and  $D_2$  is  $\{1, 1\}$ .

### 4.2 Testsets and Experiment Settings

In prior works, blind image super-resolution tasks are usually tested on synthesized LR images from HR images. However, the human simulated degraded images can hardly reflect the low-resolution image coming from degradation in real world, which usually features complicate combinations of different degradation processes. Besides, there is no real dataset which provides real-world LR images. Therefore, we choose to use real-world images directly as our test dataset and see their performance.

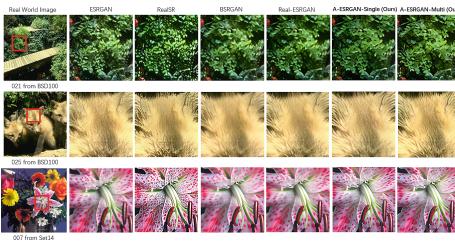
In this paper, we use the real-world images in the seven standard benchmark datasets, Set5 [2], Set14 [23], BSD100 [11], Sun-Hays80 [18], Urban100 [5], OST300 [21] and General100 [4]. These seven datasets contain images from manifold groups, such as portraits, scenery and buildings. We argue that a good general super resolution model should achieve good performance on the overall seven datasets.

### 4.3 Comparing with the State-of-the-Arts

We compare the proposed A-ESRGAN-Single and ESRGAN-Multi with several state-of-the-art(SOTA) generative based methods, i.e. ESRGAN [22], RealSR [6], BSRGAN [25], Real-ESRGAN [20] . Note that the architecture of the generators of ESRGAN, RealSR, BSRGAN and Real-ESRGAN are the same as ours, which can help verify the effectiveness of our designed discriminator.

**Table 1.** The NIQE results of different methods on Set5, Set14, BSD100, Sun Hays80, Urban 100, OST300 and General100 (The lower, the better). The best and second best results are highlighted in red and blue, respectively.

NIQE	Bicubic	ESRGAN	BSRGAN	RealESRGAN	RealSR	A-ESRGAN-Single(Ours)	A-ESRGAN-Multi(Ours)
Set5	7.8524	5.6712	4.5806	4.8629	<b>3.5064</b>	3.9125	3.8480
Set14	7.5593	5.0363	4.4096	4.4978	3.5413	<b>3.4983</b>	3.5168
BSD100	7.3413	<b>3.1544</b>	3.8172	3.9826	3.6916	3.2948	3.2474
Sun-Hays80	7.6496	3.6639	3.5609	2.9540	3.3109	<b>2.6664</b>	<b>2.5908</b>
Urban100	7.1089	<b>3.1074</b>	4.1996	4.0950	3.9290	3.4728	3.3993
OST300	7.3801	3.4689	3.2931	2.7919	3.0364	<b>2.6778</b>	<b>2.5751</b>
General100	8.0730	4.7717	5.0998	5.7158	<b>4.5357</b>	4.5585	<b>4.3992</b>

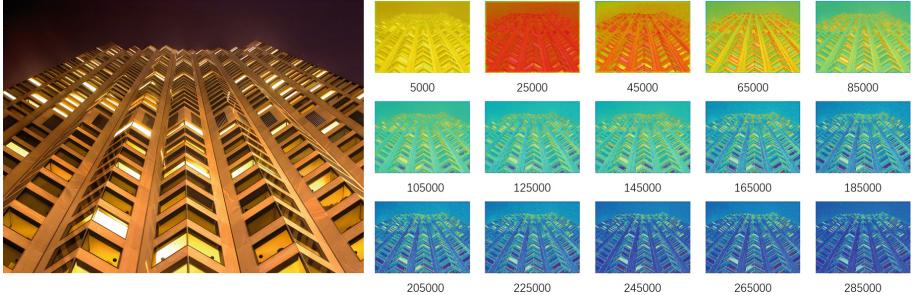


**Fig. 4.** Visual comparison of our method with other  $\times 4$  super resolution methods. Zoom in for the best view.

Since there is no ground-truth for the real-world images of the dataset, so we adopt the no-reference image quality assessment metrics NIQE [12] for quantitative evaluation. NIQE indicates the perceptual quality of the image. A lower NIQE value indicates better perceptual quality. As can be seen from the Table 1, our method outperforms most of the SOTA methods in NIQE metrics. Meanwhile, we can find it is more robust and has stronger generalization ability, since it achieves high score in all kinds of datasets. From visual comparison (some examples are shown in Fig. 4), we observe our methods can recover sharper edges and restore better texture details.

### 4.4 Attention Block Analysis

To verify the effectiveness of attention gate in our discriminator, We visualize the attention weights in the attention layer from test images during our training process. An example is shown in Fig. 5. Initially, the attention weights are uniformly distributed in all locations of the images. As the training process goes on,

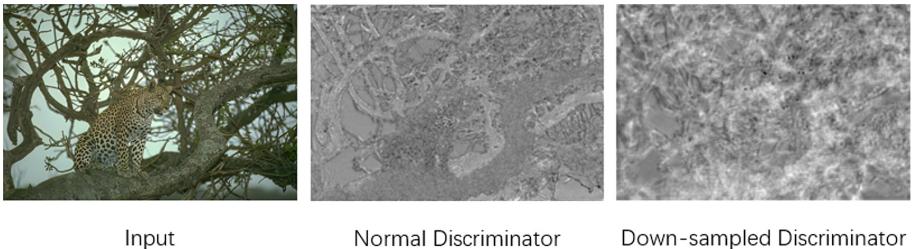


**Fig. 5.** The figure shows the weight in the third attention layer across the training process from iteration 5000 to 285000 at an interval of 20000. The example image is picked from Urban100 [5]. It clearly shows at first the attention is uniformly distributed. Then the attention is gradually updated and begins to focus on the edges. Zoom in for the best view.

we can observe that the attention weight is gradually updated and begin to focus on “particular regions”, which are the edges where color changes abruptly. Meanwhile, by visualizing attention map at different layers, we argue that different attention layers recognize the images at different granularity. The lower attention layers that are coarse-grained and give rough edges of the patches while the upper attention layers are fine-grained and focus on details such as lines and dots.

#### 4.5 Multi-scale Discriminator Analysis

We study the output image generated by the two attention discriminators propose that the two discriminators play different roles in identifying the properties of the images. The normal discriminator, which is also used in the single version, emphasizes more on lines. In contrast, the downsampled inputs with blurred edges force the other discriminator to focus more on larger patches. As shown in Fig. 6, the output image of the normal discriminator judges the edges while the dowsampled discriminator judges thicker blocks, such as textures on the branches of the tree.



**Fig. 6.** The figure shows Unet output of the two discriminators. The example image is picked from BSD100 [11]. The example shows the normal discriminator(first) would focus on lines in the image while the discriminator that parse the downsampled input will focus on patches. The brighter a pixel is, the more likely it is a real picture.

## 4.6 Ablation Study

**Effectiveness of Attention U-Net Discriminator.** The key factor of A-ESRGAN surpassing the existing models is our designed attention U-Net discriminator. In the ablation study, we compare the results of Real-ESRGAN model with A-ESRGAN-Single model. The only difference between these two networks is that Real-ESRGAN uses a plain U-Net as discriminator, while A-ESRGAN applies an attention U-Net discriminator.

As shown in Table 1, A-ESRGAN-Single achieves better NIQE in all tested datasets. By taking a close look at the result, we could find since plain U-Net uniformly gives weight to each pixel, it can't distinguish between the subject area and background of images. However, as shown in Sect. 4.4, the attention U-Net is able to put more efforts on the edges than on ordinary pixels. We believe this will bring at least two benefits. First, the result image will give sharper and clearer details as shown in 7a. Second, when up-sampling process is based on the main edges of the image, there will be less probability of distortion (like shown in 7b).

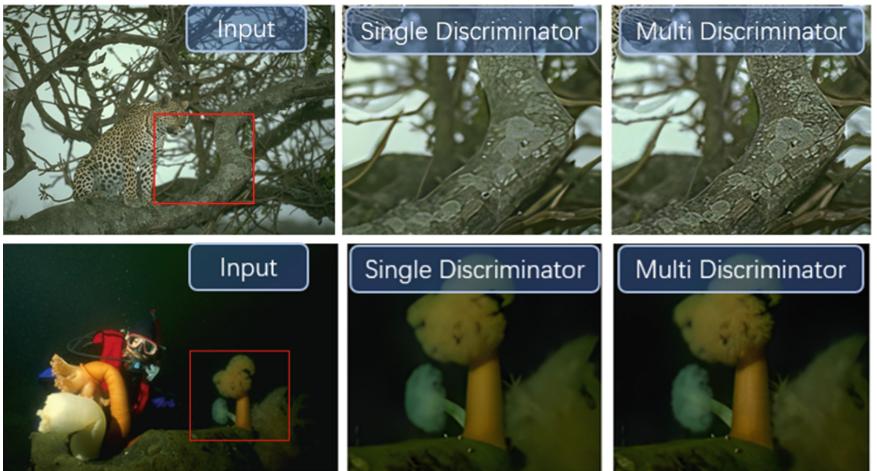


**Fig. 7.** Ablation on the discriminator design.

**Effectiveness of Multi-scale Discriminator.** The multi-scale discriminator enables our model to focus not only on the edges but also on more detailed parts such as textures. In the ablation study, we compare the results of the A-ESRGAN-single and the A-ESRGAN-multi. The latter has the same generator

as the former while it possesses two discriminators, which are a normal one and a downsampled one.

As shown in Table 1, the A-ESRGAN-multi surpasses the performance of A-ESRGAN-single in all dataset except Set14. By analyzing the output images of the two models, we conclude that the A-ESRGAN-multi does much better in showing the texture of items than A-ESRGAN-single. Like the images shown in Fig. 8, the A-ESRGAN-single poorly performs on rebuilding the texture of the branches and the sea creature. In contrast, because the downsampled discriminator focuses on patches, it can rebuild the texture and give sharper edge details.



**Fig. 8.** Ablation on the multi-scale design.

## 5 Conclusions

In this paper, a multi-scale attention U-Net discriminator is proposed to train a deep blind super-resolution model. Based on the new discriminator, we trained a deep blind super-resolution model and compared it with other SOTA generative methods by directly upscaling real images in seven benchmark datasets. Our model outperforms them in both NIQE metrics and visual performance. By systematically analyzing how the attention coefficient changes across time and space during the training process, we give a convincing interpretation of how the attention layer and multi-scale mechanism contribute to the progress in SR problems. We fully believe that other super-resolution models can benefit from our work.

## References

1. Agustsson, E., Timofte, R.: NTIRE 2017 challenge on single image super-resolution: dataset and study. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1122–1131 (2017). <https://doi.org/10.1109/CVPRW.2017.150>
2. 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 British Machine Vision Conference, pp. 1–10. BMVA Press (2012). <https://doi.org/10.5244/C.26.135>
3. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. CoRR abs/1501.00092 (2015). [arxiv.org/abs/1501.00092](http://arxiv.org/abs/1501.00092)
4. Dong, C., Loy, C.C., Tang, X.: Accelerating the super-resolution convolutional neural network. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 391–407. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46475-6\\_25](https://doi.org/10.1007/978-3-319-46475-6_25)
5. Huang, J.B., Singh, A., Ahuja, N.: Single image super-resolution from transformed self-exemplars. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5197–5206 (2015)
6. Ji, X., Cao, Y., Tai, Y., Wang, C., Li, J., Huang, F.: Real-world super-resolution via Kernel estimation and noise injection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2020)
7. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. CoRR abs/1603.08155 (2016). [arxiv.org/abs/1603.08155](http://arxiv.org/abs/1603.08155)
8. Lai, W.S., Huang, J.B., Ahuja, N., Yang, M.H.: Deep Laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 624–632 (2017)
9. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. CoRR abs/1609.04802 (2016). [arxiv.org/abs/1609.04802](http://arxiv.org/abs/1609.04802)
10. Li, C., Wand, M.: Combining Markov random fields and convolutional neural networks for image synthesis. CoRR abs/1601.04589 (2016). [arxiv.org/abs/1601.04589](http://arxiv.org/abs/1601.04589)
11. 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 Eighth IEEE International Conference on Computer Vision, ICCV 2001, vol. 2, pp. 416–423. IEEE (2001)
12. Mittal, A., Fellow, I.E.E.E., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. IEEE Sig. Process. Lett. **20**(3), 209–212 (2013)
13. Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y.: Spectral normalization for generative adversarial networks. In: International Conference on Learning Representations (2018). [www.openreview.net/forum?id=B1QRgziT](http://www.openreview.net/forum?id=B1QRgziT)
14. Oktay, O., et al.: Attention U-Net: learning where to look for the pancreas (2018)
15. Park, S.-J., Son, H., Cho, S., Hong, K.-S., Lee, S.: SRFeat: single image super-resolution with feature discrimination. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11220, pp. 455–471. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01270-0\\_27](https://doi.org/10.1007/978-3-030-01270-0_27)
16. Sajjadi, M.S., Scholkopf, B., Hirsch, M.: EnhanceNet: single image super-resolution through automated texture synthesis. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4491–4500 (2017)
17. Schonfeld, E., Schiele, B., Khoreva, A.: A U-Net based discriminator for generative adversarial networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020)

18. Sun, L., Hays, J.: Super-resolution from internet-scale scene matching. In: 2012 IEEE International Conference on Computational Photography (ICCP), pp. 1–12. IEEE (2012)
19. Wang, T.C., Liu, M.Y., Zhu, J.Y., Tao, A., Kautz, J., Catanzaro, B.: High-resolution image synthesis and semantic manipulation with conditional GANs. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
20. Wang, X., Xie, L., Dong, C., Shan, Y.: Real-ESRGAN: training real-world blind super-resolution with pure synthetic data. In: 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), pp. 1905–1914 (2021). <https://doi.org/10.1109/ICCVW54120.2021.00217>
21. Wang, X., Yu, K., Dong, C., Loy, C.C.: Recovering realistic texture in image super-resolution by deep spatial feature transform. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
22. Wang, X., et al.: ESRGAN: enhanced super-resolution generative adversarial networks. In: Leal-Taixé, L., Roth, S. (eds.) ECCV 2018. LNCS, vol. 11133, pp. 63–79. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-11021-5\\_5](https://doi.org/10.1007/978-3-030-11021-5_5)
23. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **19**(11), 2861–2873 (2010)
24. Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R.: Plug-and-play image restoration with deep denoiser prior. *IEEE Trans. Pattern Anal. Mach. Intell. PP*, 1 (2021)
25. Zhang, K., Liang, J., Van Gool, L., Timofte, R.: Designing a practical degradation model for deep blind image super-resolution. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 4791–4800 (2021)
26. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 294–310. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01234-2\\_18](https://doi.org/10.1007/978-3-030-01234-2_18)