Wireless Networks
https://doi.org/10.1007/s11276-020-02496-9

High fidelity single image blind deblur via GAN

Xiaoli Li[1] · Gen Li[1] · Zhenlong Du[1]

Accepted: 3 November 2020

Abstract

To reconstruct the high resolution image makes sense from single image with low resolution. Most conventional methods assume that the blur kernel is known, however the blur kernel within single blurring image is always unknown, then it is necessary to generate the kernel dynamically during deblurring. The paper proposes single image high-fidelity blind deblurring method based on GAN. The degradation of super resolution networks is first used to synthesize high resolution images. A blur kernel discriminator is then trained to analyze the generated high-resolution images and errors that occur when the prediction generator provides incorrect blur kernel. Thus, the blur kernel provided by the generator is closer to the actual image. Finally, through the redefinition and optimization of loss function, the perception loss is replaced by the per-pixel loss to obtain better visual effects. Experiments show that the proposed method can achieve high fidelity deblurring results 1.2% higher than those of traditional methods.

Keywords Image deblur · Blind · Super resolution · Adversarial generation network

1 Introduction

In the latest decades, with the development of deep learning, to recover the high fidelity image from the previous low-resolution acquisition devices has become an active research area. To reconstruct the super-resolution and high fidelity image from single low-resolution image is very urgent in fields of production line, medical imaging, public security, and so on. In the paper, a novel high fidelity blind image deblurring approach is proposed.

The super resolution reconstruction from single image is an inherently ill-posed problem because a low-resolution image can recover many super-resolution images. Traditionally, the low resolution image is interpolated to generate the super resolution image, moreover the generated image is sharpened to be clear.

The pixel can be directly used for producing the super resolution image, this manner readily results the distortion within the reconstructed image. Manifestly, more pixels are used, higher fidelity image can be brought out. With the popularity of deep learning, many learning-based super resolution methods have been studied and achieved persuasive results. However, most of the methods are based on the bicubic degradation blur kernel design, and it is still a challenge for the low-resolution image processing of arbitrary blur kernel.As shown in Fig. 1, convolution kernel is very important to synthesize super resolution images. In addition, most studies choose per-pixel loss function to evaluate the error between the generated image and the real image. However, the use of a per-pixel loss function causes large error. So there is still a gap between the super-resolution image produced and the actual image, since the PSNR metric fundamentally disagrees with the subjective evaluation of human observers.

In the discussed method, First, we propose a framework that can automatically generate super resolution image, generate super-resolution image and corresponding blur kernel. Second, we use the convolution kernel discriminator to identify the convolution kernel for better super resolution estimation. Finally, the perception-loss function is introduced to replace the pixel-by-pixel loss function. This perception-loss function is implemented by VGG

✉ Xiaoli Li
lixl@njtech.edu.cn

Gen Li
ligen1803150216@163.com

Zhenlong Du
duzhl-cad@163.com

[1] School of Computer Science and Technology, Nanjing TECH University, Nanjing 211816, China

Published online: 08 January 2021

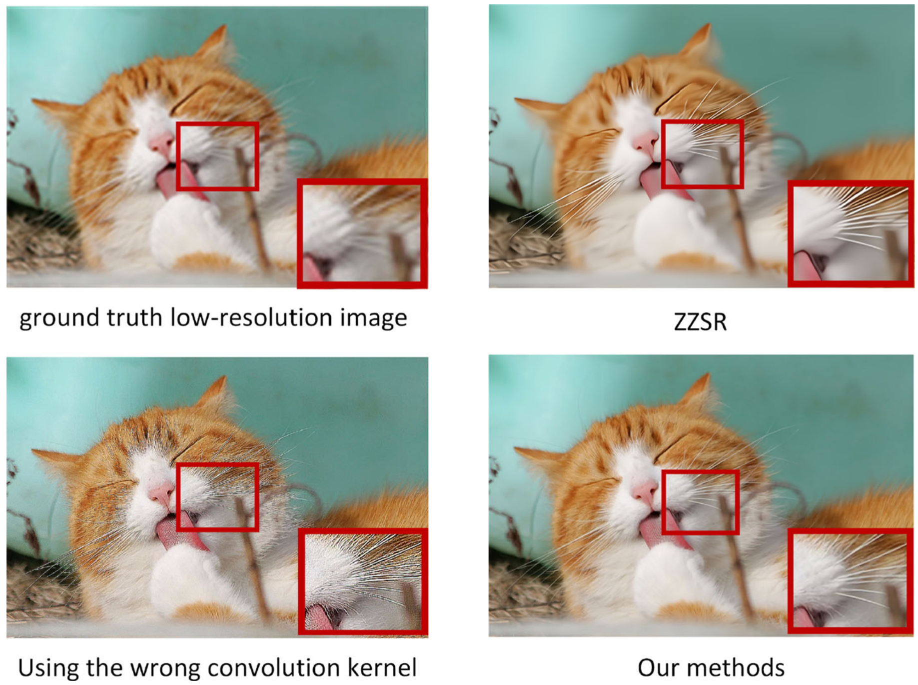


Fig. 1 Compare of synthesized images using different kernel

neural network. Compared with other methods, it retains advanced features and generates super-resolution images with better visual experience. This indicates that in the actual scene of unknown convolution kernel, compared with the method assuming that the convolution kernel is fixed, the generated image is not too smooth or too sharp, and gives the best impression. Just like Fig. 1 in the actual scene, when the assumed degradation model does not match the actual degradation model, the over-sharpened picture is generated, while our method can still produce clear picture for the image with unknown degradation model.

In this paper, the existing blind super resolution image reconstruction algorithms and their applications are analyzed and studied. A reconstruction algorithm of blind super resolution image based on perceptual loss function is proposed. The remainder of the paper is organized as follows, in Sect. 2, relevant studies are discussed. Section 3 related methods are reviewed. Then the presentation and analysis of experimental results are arranged in Sect. 4. The last section is the conclusion and future work. In this paper the algorithm follows a standard process: first, input low-resolution images, and train the generator to generate super-resolution image; second, train the discriminator to identify the artifacts generated image by the wrong convolution kernel; finally, input the generated high-resolution image into the perceptual loss network to achieve the best results based on the evaluation of human eyes.

2 Related works

In recent years, with the significant improvement of neural network to super resolution image generation, deep learning has become a popular research method. Dong et al. [1] applied the deep learning model of convolutional neural network in the field of super-resolution reconstruction. Low resolution and high resolution images are directly mapped by end-to-end learning. On this basis, Zhang et al. [2] combined bicubic interpolation with deep learning and proposed a multi-path learning method to improve the ability of image feature extraction. But it requires a complex transformation, in order to avoid it, Hui et al. [3] applied residual mapping to restore high frequency detail. Most existing CNN-based super-resolution models require high computation cost, which largely limits their application. Rushi Lan et al. [4] proposed a dense and lightweight network called MADNet, which can greatly improve learning efficiency.These methods have achieved impressive results, when the used blur kernel and the selected blur kernel are approximately the same.

Prior knowledge and reconstruction constraints are very important to image super-resolution. The fusion method of input images [5] is very important for obtaining high-quality information at any position of the observation scene. Image priori reflects the characteristics of natural image, which is the knowledge about image degradation model and noise level. The reconstruction constraint means that the output of the subsampled super-resolution image should be approximately equal to the original low-resolution image. If the type of blur degenerate kernel of the super-resolution algorithm is different from the real blur degenerate kernel, the image will be too smooth. In the early academic research, appropriate blur kernel are estimated by repeating patterns between small image blocks. In recent years, some deep learning methods have been applied to blur degenerate kernel. Zhang et al. [6] took blur degenerate kernel and noise as supplementary inputs to the neural network, and used principal component analysis mapping to map the blur degenerate kernel to anisotropic Gaussian blur degradation kernel, so as to adapt the images with different blur degradation kernel. Later on, Zhang et al. [7] proposed a new degradation model to perform blur operations after sampling. The superdecoder was adopted to replace gaussian filter prior, and the superdecoder was inserted as a module to solve the problem of super resolution. Both of the above methods work well, but the parameter information of blur degradation kernel must be known in advance.

Therefore, neither method is suitable for blind super resolution with unknown parameter information. Shocher et al. [8] took advantage of the repeatability of the internal information of a single image and proposed an unsupervised super-resolution method with different blur degradation kernels. The example extracted from the input image trains a small image specific CNN to adapt to images with different blur degradation kernel. Different from the Shocher et al,Gu et al. [9] employs the correlation between super resolution results and kernel mismatch. Gu et al. [9] automatically estimated the blur degradation kernel depending on isotropic conditions. First, the isotropic Gauss kernel was used to check the super-resolution image for blur processing, then the bicubic interpolation was used for down-sampling. The low resolution image is estimated by kernel variance, and another neural network is used to identify whether the generated image bears artifacts. But the kernel space is limited to Gaussian function, which hinders the further expression ability of the model. Cornillere et al. [10] used GAN network to predict the degradation model to achieve the purpose of matching the blur kernel with the generated image. These methods are based on the pixel-by-pixel loss function, which can obtain a high PNSR, but it is not consistent with people's visual experience. So, there is a gap in the generation of visually pleasing images.

In order to achieve a better accordant with human perception, Mahendran et al. [11] understand deep image representations by inverting them. Later on, Johnson et al. [12] introduced perceptual loss based on high level features extracted from pre-trained networks,for the task of style transfer and super resolution. On this basis, in order to make the generated highlights similar not only in low-level pixels, but also in high-level features, Ledig et al. [13] proposed SRGAN. Therefore, SRGAN can generate realistic images.In order to avoid visually unpleasant artifacts, local texture loss is introduced, Sajjadi et al. proposed EnhanceNet [14], which applied a similar approach. Zhang et al. [15] believe that the stronger a feature set is at classification and detection, the stronger it is as a model of perceptual similarity judgments. So the perceptual loss based on deep features fits human visual perception well. Although these algorithms can obtain better perceptual image quality, there is no good objective criterion to evaluate the results in their papers. As proposed in PIRM-SR 2018 [16], measure the performance of the super-resolution methods using pixel-based quality and perception-based quality.

3 Method

In the paper, high fidelity image I^{HR} is reconstructed from the blur image I^{LR} using deep learning, it is consistent with the visual experience, and the reconstruction procedure is blind performed. I^{LR} is degraded from I^{HR}, and the degradation model can be represented as the follow.

$$I^{LR} = \left(k \otimes I^{HR}\right)\downarrow_s + n \tag{1}$$

which $\otimes$ is the convolution operator, k is the image quality degradation kernel, herein the isotropic Gaussian degradation kernel is adopted. $\downarrow s$ is a down-sampling operation with a factor s, n is the additive Gaussian white noise with standard deviation. The flowchart of general image degradation is shown in Fig. 2.

In the paper, a super resolution model is constructed based on GAN, in which the generator is a super resolution neural network considering the convolution degradation kernel. Because the presented model can deal with many types of image degradation, it is suitable for the implementation of blind super resolution. This model can process low resolution images more flexibly and generate more real super resolution images.

The per-pixel loss function commonly used in many literatures [17] don't well evaluate the difference between the real images and the generated images, so it is difficult to generate super-resolution images conforming to human

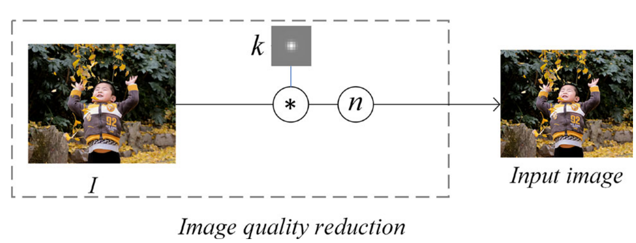


Fig. 2 Flowchart of general image degradation

visual experience. To alleviate this defect, we use the perceived loss network. Perceptual loss network is implemented by VGG16 network, which extracts the advanced perceptual features of the image, so that the generated super-resolution image has a good visual experience. The process of image blind super resolution is shown in Fig. 3.

3.1 The Generator blind super-resolution

The GAN network consists of a convolutional degradation kernel network $\mathcal{F}_k$ and a super-resolution image network generator $\mathcal{F}_g$, as shown in Fig. 4. First, the convolutional degradation kernel is mapped to a low-dimensional implicit presentation layer q_k . By considering the same convolution degradation model at each pixel, we can obtain a set of convolution degradation maps. Next, the mapping ρ is input to the generator together with the low-resolution image I_l to generate a super-resolution image.

In the mapping step of convolution degradation kernel, the neural network $\mathcal{F}_k$ with parameter λ_k is used. Before providing the convolution degradation kernel to the GAN network generator, the low-dimensional implicit representation layer q_k of the convolution degradation kernel needs to be calculated with Eq. (2):

$$q_k = \mathcal{F}_k(k|\lambda_k) \tag{2}$$

The generator in the GAN network predicts the super-resolution image I^* with paramater λg based on the low-resolution image and the degradation mapping ρ set based on each pixel and written as Eq. (3)

$$I^* = \mathcal{F}_g(I_l, \rho|\lambda_g) \tag{3}$$

In the case of a single convolution degradation kernel, it's low-dimensional implicit representation layer q_k is repeated at each pixel location. If each pixel or region uses a different convolution degradation kernel k_i , then apply the above convolution degradation kernel mapping transformation to each kernel and obtain a lower-dimensional implicit representation layer q_{k_i} . Since we provide the degradation information of super-resolution image into low-resolution image in the form of spatial feature map, the convolution degradation kernel can change in different parts of the image. In this way, different parts of the image can be processed in different degradation models, which can produce a better visual experience of the image.

The network structure of the generator uses a series of dense compression units. The network predicts a residual image, which is then added to the bicubic upsampling image to produce the output image. The training generator can be expressed as Eq. (4)

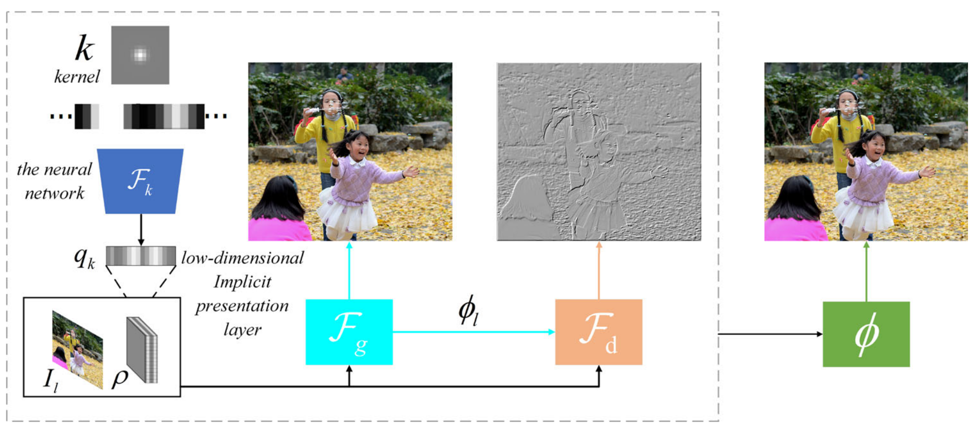


Fig. 3 The process of image super resolution

Fig. 4 Relationship between convolution degradation kernel and generated image

$$\lambda_g^*, \lambda_k^* = \arg\min_{\lambda_g,\lambda_k} {}_{Ip_l,kp_k}\left[\mathcal{L}\left(I, \mathcal{F}_g\left(I_l, \rho|\lambda_g\right)\right)\right] \tag{4}$$

During training, the convolution kernel generator for the whole image is used to generate random convolution degradation kernel. The generated super-resolution image is identified by the discriminator, and the loss function is used for training model. The convolution degradation kernel which is most suitable for low resolution image is selected using the convolution kernel generator discriminator. This process reduces the distortion and artifacts of the predicted super-resolution image.

3.2 Convolution degradation kernel discriminator and optimize

If the convolution degradation kernel is known, the previously described super-resolution network can reconstruct the original super-resolution image. However, the convolution degradation kernel is unknown in blind super-resolution. If improper convolution degradation kernel is selected, artifacts (eg. in Fig. 4) produced in the generated super-resolution image. According to this characteristic, the kernel discriminator is used to estimate the error of super-resolution image generated by generator network. To take advantage of this, we propose the kernel discriminator network to estimate the errors in the generated image I^*.The gap δ_I between the real image and the generated image is formulated as Eq. (5)

$$\delta_I = \mathcal{F}_g\left(I_l, \rho_{\mathrm{GT}}|\lambda_g\right) - \mathcal{F}_g\left(I_l, \rho|\lambda_g\right) \tag{5}$$

where ρ_{GT} is a true degradation map used to generate low resolution image I_l, ρ stands for sampling from the kernel distribution during training.

Because after the proposed generator and convolution degradation kernel discriminator, the convolution kernel can be predicted. The best quality effect is obtained when the correct convolution degradation kernel is provided. On the other hand, the trained convolution degradation kernel discriminator perdicts the error in generating super-resolution images using the generator, thus identifying the best convolution degradation kernel. It mainly identifies the region which artifacts are generated by the use of erroneous convolution degradation kernels. we use the prediction error as the objective function and minimize the objective function by finding the correct convolution degradation kernel.

Formally, convolution degradation kernel optimization can be written as Eq. (6)

$$\rho^* = arg_\rho \min ||F_d(I_l, \phi_l, \rho|\lambda_d)||_1 \tag{6}$$

where ρ is the local adaptive convolution degradation kernel mapping estimated from the low resolution image, which makes the convolution degradation kernel change with spatial variation. In the simple case that there is only one convolution degradation kernel in the whole image, the low dimensional implicit overlay layer of one convolution degradation kernel can be optimized. Firstly, uniform sampling is carried out for the convolution degradation kernel. Then the error of each convolution degradation kernel is evaluated. Finally, The convolution degradation kernel with the minimum error is selected as the initial value, and then it is further optimized by Eq. (7).

$$\rho^* = \rho - \eta \nabla_\rho \theta(I_l, \rho) \tag{7}$$

where $\theta(I_l, \rho)$ is the loss function, η is the weight. Our goal is to use Adam optimizer to minimize the error of convolution degenerate kernel in Eq. (7) and iterate out the convolution kernel which is most suitable for the input image

3.3 Design of perceptive loss function network

As shown in Fig. 5, our system consists of two parts, an image blind super-resolution network f_g and a loss network ϕ for defining several loss functions. The image blind super-resolution network converts the input image x into the output image y by mapping $y = f_g(x)$, and calculates the difference between the output image and the real image. The stochastic gradient descent is used to train the image transformation network to minimize the weighted combination of loss functions Eq. (8):

$$W^* = \arg_w \min E_{x,\{y_i\}}\left[\sum_{i=1} \lambda_i \varepsilon_i (f_W(x), y_i)\right] \tag{8}$$

In order to solve the shortcoming of per-pixel loss, it is necessary to design a loss function to better measure the perception and semantic difference of images. Since the convolutional neural network for image classification preprocessing has learned to encode the perceptual and semantic information to measure in the loss function, we use the pre-processing network of image classification as the perceptual loss function network ϕ.

The perceptual loss function network ϕ measures the difference in content between images. x is a low-resolution input, content target y_c is a ground-truth image.

The perceptual loss functions is used to measure the high-level perceptual difference between images and written as Eq. (9). In the experiment, the loss function neural network is a 16-layer VGG network trained on ImageNet. Instead of encouraging the pixels of the output image $y = f_w(x)$ to exactly match the pixels of the target image Y , we encourage them to have similar feature representations calculated by the loss network ϕ . $\phi_j(x)$ is the activation of layer j of network ϕ when processing image x,

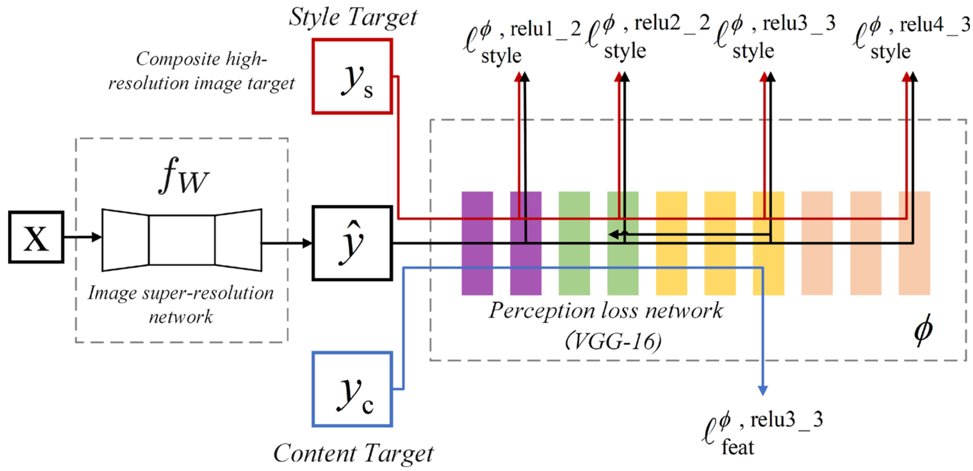


Fig. 5 The image super resolution network with perceptive loss function is added

If j is a convolution layer, $\phi_j(x)$ will be a feature map of $C_j \times H_j \times W_j$ shape.

$$\varepsilon_{\text{feat}}^{\phi,j}(\hat{y}, y) = \frac{1}{C_j H_j W_j} ||\phi_j(\hat{y}) - \phi_j(y)||_2^2 \tag{9}$$

When the output image deviates from the ground-truth image in content, the loss function will punish the output image. Allow semantic knowledge to be transferred from preprocessing loss network to super-resolution network.

$$\text{Perceptualindex} = \frac{1}{2}((10 - Ma) + NIQE) \tag{10}$$

where $Ma(\cdot)$ is the quality score measure, and $NIQE(\cdot)$ means the quality score by the natural image quality evaluator (NIQE) metric.we measure the performance of the super-resolution methods by Eq. (10).

4 Blind super-resolution experiment and analysis

The computer used in this experiment is configured with CPU 4.2GHz, RAM 16GB and graphics card NVIDIA GeForce GTX 1080TI. The tested data set is the real data set collected from the internet, and then the convolution kernel is randomly selected from the convolution degradation kernel space for degradation operation. First, We evaluated our method in detail and compared it with the existing blind super-resolution methods. Figure 6 shows the results of x4 blind super resolution. Compared with other methods, the model trained for feature reconstruction has done very well in reconstructing sharp edges and details. Second, we designed two cases, the generator with convolution degenerate kernel and the generator without convolution degenerate kernel, which proved that convolution degenerate kernel is very important in image super-resolution. Finally, Compared with other blind super-resolution methods by questionnaire and the perceptual quality of the super-resolution image by perceptual index, we quantitatively evaluated our method.

4.1 Convolution degradation kernel space representation

The image degradation process is described as a blurring operation followed by downsampling. Experience and theoretical analysis show that the influence of an accurate blur kernel is much greater than that of a complex image prior. Specifically, when the kernel is assumed to be smoother than the real kernel, the reconstruction image will be excessively smooth. Most super-resolution methods actually support this situation. On the other hand, when the assumed kernel is larger than the real kernel, the reconstruction image will be excessively sharpened. The most common choice is isotropic Gaussian blur kernel parameterized by standard deviation or kernel width.

In order to deal with various situations, pulse, disk and bicubic convolution degradation kernels related to common scaling operations are selected in this paper and shown as Fig. 7.

4.2 Compare with existing methods

For the input image, the original training image is first downsampled by an amplification factor N to become a low-resolution image. For quantitative evaluation, the traditional metrics to evaluate super-resolution are PSNR and SSIM, but these are quite different from human evaluation of visual quality. PSNR and SSIM depend on the low-level difference between pixels, PSNR is equivalent to a per-pixel loss, and the assumption of SSIM is additive Gaussian noise. Therefore, the goal of the experiment is not to obtain the best PSNR or SSIM results, but to show the

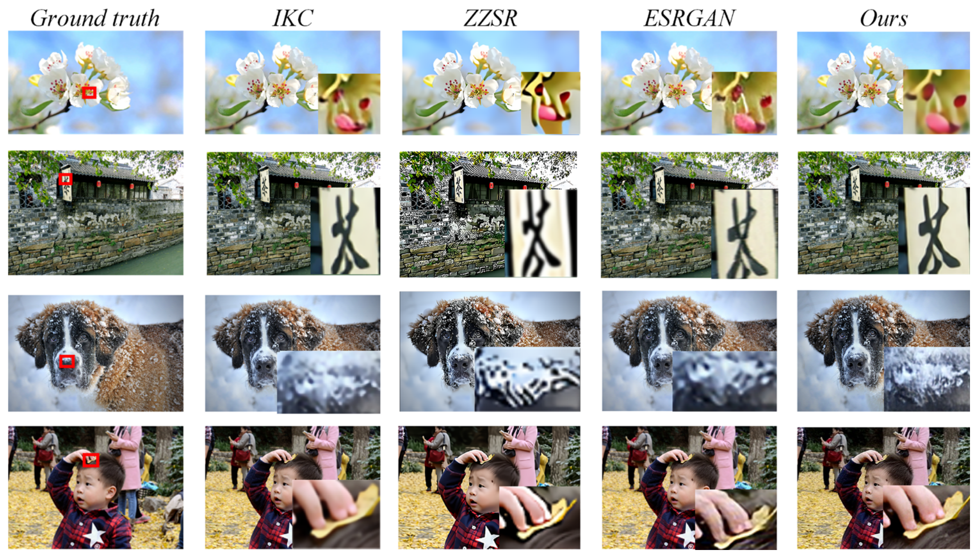


Fig. 6 Comparison with other methods

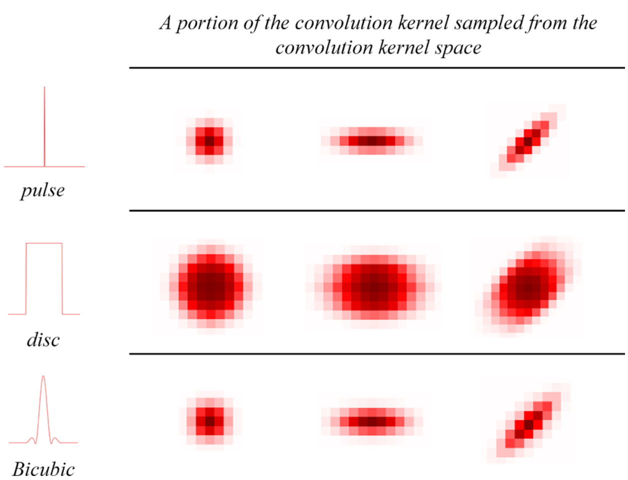


Fig. 7 Convolution kernels

quality difference between models trained with per-pixel and feature reconstruction loss. Using Adam iterator, bathsize is set to 200, 50 iterations are performed, the learning rate is set to 0.001, and the activation function uses relu function. On NVIDIA GTX 1080TI, convolution kernel grid search initialization takes about 200 seconds and optimization process takes 8 minutes. See Fig. 8 for more comparisons

In petal images of Fig. 8, a disk convolution kernel is used, so our details are clearer, and other methods produce artifacts on the image to make the contour excessively smooth. In the second row of ancient architecture images, pulse convolution kernel is used for down-sampling, which can well reconstruct the details of buildings, such as bricks and tiles on buildings. It can be clearly seen that our method is the most natural to reconstruct the words on the building signboard, with the best results based on human visual evaluation. In the animal image of the third line, The proposed method better reconstructed the details of the dog's nose. Obviously our method reconstructed the snowflakes more clearly on the nose of dog. In the last line, our method reconstructs the details of children's hands well and makes the imaging effect more perfect in general.

4.3 Detailed evaluation

Compared with the existing image super-resolution methods, and the proposed method is tested with the same data set. These methods include IKC [9], ESRGAN [18], ZZSR [8]. Super-resolution images were collected from DIV2K [19]. The training data set is enhanced by random horizontal inversion and rotation, and several parameters of Gaussian are sampled for each basic convolution kernel. After down-sampling, under the blind super-resolution and non-blind super-resolution settings, different generators are used to upscale the image. Reference results are provided for convolution-free kernel generators. PSNR and the similarity of learning perception image blocks from [19] are used as error metrics. Higher PSNR is better, while lower LPIPS is better. See Table 1 for more information

From these results, it can be known that the quality of the generated pictures are the worst in all cases without using any blur degradation kernel information compared with other results. Under the conditions of blind super-resolution and non-blind super-resolution, information about degradation can improve the quality of generated pictures, and PSNR and LPIPS values show obvious improvement. When we operate with convolution kernel discriminator under blind setting, the generated super-resolution image is closest to the original real image. In the complex disc convolution kernel degradation types, compared with other types, the quality of the final generated

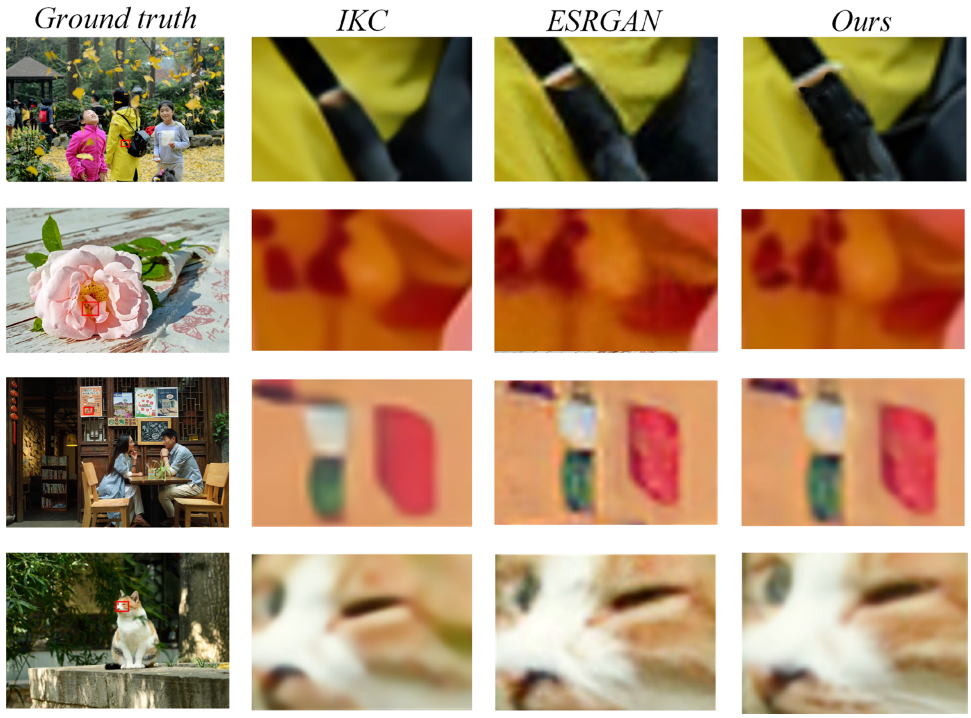


Fig. 8 Comparison with other methods

Table 1 Detailed SR evaluation

degradation	No convolution kernels		With convolution kernels		Blind super resolution	
	PNSR	LPIPS	PNSR	LPIPS	PNSR	LPIPS
pulse	31.43	0.145	32.68	0.106	33.12	0.123
bicubic	32.63	0.161	33.34	0.101	34.63	0.118
disk	30.98	0.169	32.15	0.125	31.68	0.139

image is better than that of the non-blind super-resolution image.

Table 2 shows the performance of the considered super-resolution methods for the Set5, Set14, and BSD100 datasets. The lower the perceptual index is, the better the perceptual quality is. Our model shows slightly lower PSNR results, while the perceptual quality is significantly improved. These results show that our model achieves proper balance between the distortion and perception aspects.

We use questionnaires to evaluate our methods quantitatively. Study participants were shown input images, The input was shown at the top and marked as A , and the two super resolution images were placed at the bottom and marked as B and C . Participants were then asked to select the result that resembles the input the most: Which of the two images on the bottom B or C better represents image A? If both are equally good then select Both , and if neither represent A then select Neither . The answer options were B , C , Both , and Neither. The results are shown in Fig. 9.

Table 2 Performance of the SR methods

Set5	RMSE	PSNR	SSIM	Perceptual Index
ESRGAN	9.14	30.89	0.92	3.42
IKC	8.02	30.59	0.89	4.56
ZZSR	8.71	30.31	0.91	5.17
Ours	9.58	31.54	0.88	2.97

Set14	RMSE	PSNR	SSIM	Perceptual Index
ESRGAN	14.47	26.11	0.69	2.88
IKC	12.65	27.27	0.74	3.55
ZZSR	11.64	28.26	0.77	5.51
Ours	11.58	27.9	0.7459	2.81

BSD100	RMSE	PSNR	SSIM	Perceptual Index
ESRGAN	16.33	25.18	0.64	2.71
IKC	14.12	26.32	0.69	5.18
ZZSR	13.07	27.11	0.71	3.43
Ours	11.64	27.87	0.75	2.67

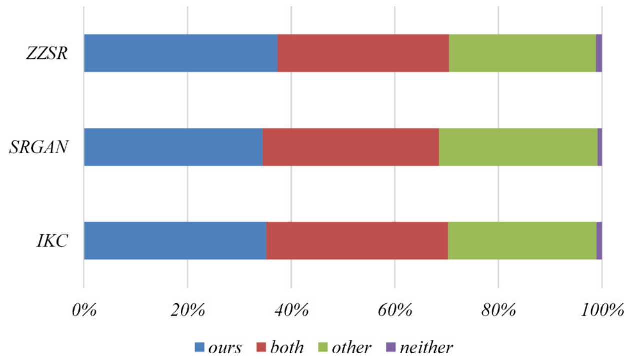


Fig. 9 Percentage comparison with other methods

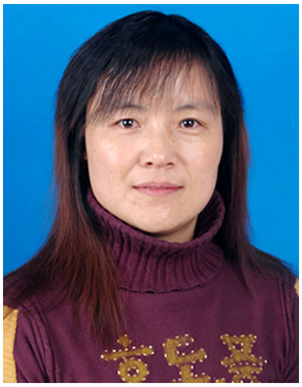
5 Conclusion

In this paper, a framework that can execute blind super resolution in a fully automated manner is established based on GAN, and optimize it using a perceptual loss network. The convolution kernel discriminator network analyzes the degradation model of the image, so as to select a more suitable convolution kernel, which is more in accordance with the actual situation. The experimental results show that the method used in this paper can produce visually plausible pictures, which is superior to other super resolution methods. The following will further study how to expand the space of convolution kernel and design a more compact network structure to reconstruct higher quality super resolution images.

Acknowledgements The paper is supported by National Natural Science Foundation of China (61672279), and Open Foundation of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, China (2016491411).

References

1. Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, *38*(2), 295–307.
2. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., & Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 286–301).
3. Hui, Z., Wang, X., & Gao, X. (2018). Fast and accurate single image super-resolution via information distillation network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 723–731).
4. Lan, R., Sun, L., Liu, Z., Lu, H., Pang, C., & Luo, X. (2020). Madnet: A fast and lightweight network for single-image super resolution. *IEEE Transactions on Cybernetics*. https://doi.org/10.1109/TCYB.2020.2970104.
5. Zhao, W., Huimin, L., & Wang, D. (2017). Multisensor image fusion and enhancement in spectral total variation domain. *IEEE Transactions on Multimedia*, *20*(4), 866–879.
6. Zhang, K., Zuo, W., & Zhang, L. (2018). Learning a single convolutional super-resolution network for multiple degradations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3262–3271).
7. Zhang, K., Zuo, W., & Zhang, L. (2019). Deep plug-and-play super-resolution for arbitrary blur kernels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1671–16810.
8. Shocher, A, Cohen, N., & Irani, M. l. (2018). "zero-shot" super-resolution using deep internal learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3118–3126).
9. Gu, J., Lu, H., Zuo, W., & Dong, C. (2019). Blind super-resolution with iterative kernel correction. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1604–1613).
10. Cornillere, V., Djelouah, A., Yifan, W., Sorkine-Hornung, O., & Schroers, C. (2019). Blind image super-resolution with spatially variant degradations. *ACM Transactions on Graphics (TOG)*, *38*(6), 1–13.
11. Mahendran, A., & Vedaldi, A. (2015). Understanding deep image representations by inverting them. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp 5188–5196).
12. Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision* (pp. 694–711). Springer.
13. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., & Wang, Z. et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp 4681–4690).
14. Sajjadi, M. S. M., Scholkopf, B., & Hirsch, M. (2017). Enhancenet: Single image super-resolution through automated texture synthesis. In *Proceedings of the IEEE International Conference on Computer Vision* (pp 4491–4500).
15. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 586–595).
16. Blau, Y., & Michaeli, T. (2018). The perception-distortion tradeoff. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6228–6237).
17. Xu, J., Chae, Y., Stenger, B., & Datta, A. (2018). Dense bynet: Residual dense network for image super resolution. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (pp. 71–75). IEEE.
18. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., Loy, C. C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 0-0).
19. Agustsson, E., Timofte, R. (2017). Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 126–135).



Xiaoli Li Who received the Ph.D. degree from Hohai University. Her research is mainly in the fields of Water information processing and Computer vision.

Zhenlong Du Who received the Ph.D. degree from Zhejiang University. His research is mainly in the fields of Computer vision, Virtual/Augmented reality and Quantum computing.

Gen Li Who is a master student. His research is mainly in the field of Computer vision.