

RESEARCH

Open Access



The visual human face super-resolution reconstruction algorithm based on improved deep residual network

Di Fan^{*}, Shuai Fang, Guangcai Wang, Shang Gao and Xiaoxin Liu

Abstract

Deep learning is a hot method in face super-resolution reconstruction in recent years, but it needs to further improve the details of reconstructed images and speed up network training. This paper improves the deep residual network from two aspects of residual unit and network structure and proposes a face super-resolution reconstruction algorithm based on the improved model. We improve the network structure of the residual unit by connecting with a densely connected convolutional layer and removing the BN layer, thereby enhancing the information flow between the inner convolutional layers and eliminate the damage to the spatial information of the image by batch normalization processing. At the same time, we combine the output characteristics of each residual unit on the basis of the global residual structure, so the face feature information is more fully utilized and the model detail recovery ability is also improved. Experiments on Fddb and AFLW face datasets show that the proposed method has better performance in feature description and detail information reconstruction, and higher PSNR and SSIM than other methods.

Keywords: Super-resolution reconstruction, Convolution neural network, Global residual learning, Dense feature fusion

1 Introduction

Face recognition applications continue to heat up, and face-related research has attracted wider attention than before. However, in natural scenes, small scale and blurred face are very common, which brings challenges to face recognition. Face super-resolution reconstruction plays an important role in face recognition in the natural environment. It is of great significance to improve the image quality and the richness of face information.

At present, image super-resolution (SR) reconstruction mainly includes interpolation-based methods [1–3], reconstruction-based methods [4–6], and learning-based methods [7–13]. The method based on interpolation and reconstruction is simple to implement and the computation cost is small, but the reconstruction effect is poor when the scale is large, and it cannot deal with complex image structure [7]. The method based on learning is to

learn the mapping relationship between low-resolution (LR) image and high-resolution (HR) image through a large number of samples training, so as to get the image reconstruction model [8].

Learning-based SR method is the current research hotspot. According to the depth of learning, it can be divided into shallow learning and deep learning. The methods of neighbor embedding (NE) [9, 10] and sparse coding (SC) [11, 12] are representatives of shallow learning. In recent years, with the rapid development of deep learning, convolutional neural network (CNN) [13] has been applied to image super-resolution reconstruction. In 2014, Dong et al. firstly proposed super-resolution CNN (SRCNN) based on convolution neural network, which utilized the strong feature expression ability of convolution neural network to improve the accuracy of reconstructed image [14]. In 2016, Kim et al. employed deeper neural networks to construct very deep convolutional network (VDSR) [15] and recursive convolutional network (DRCN) [16], and achieved better reconstruction results than SRCNN. In the same year, Shi et al.

* Correspondence: skdfandi@126.com

College of Electronic and Information Engineering, Shandong University of Science and Technology, 579 Qianwangang Road, Huangdao Zone, Qingdao, People's Republic of China

proposed the efficient sub-pixel convolutional neural network (ESPCNN) for super-resolution reconstruction, and for the first time considered how to solve the problem of image scale-up within the neural network [17]. In 2017, Tong et al. proposed an image reconstruction algorithm based on DenseNet [18], called SRDenseNet, which greatly increased the information richness in the reconstructed image [19]. The super-resolution reconstruction method based on generative adversarial nets (GANs) [20] has also been developed, and more attention to the visual effect of the image are paid in the reconstruction model [21]. Face super-resolution is the task of image reconstruction in specific scenes. Reconstruction models should pay more attention to the restoration of facial details. Some scholars have proposed several improved methods based on this point of view. In 2015, Zhu et al.'s bi-channel convolution neural network (bi-channel CNN) improves the problem of feature information loss through the cross-layer output of input images [22]. Wang et al. figured out that adding some additional information (such as texture, edge, etc.) to the deep convolution network could improve the quality of the reconstructed image [23]. In 2016, Zhu et al. proposed an iterative two-stage method for face super-resolution reconstruction [24]. In 2018, Sun et al. tried to improve the effect of the reconstructed image by increasing the depth of the network [25]. Yu Chen et al. used geometric prior information extracted from face images to improve reconstruction model performance [26]. Adrian Bulat et al. combined face feature point location and face super-resolution via a combined network, and used a GAN sub-network to enhance the super-resolution effect [27]. Xin Yu et al. propose an upsampling network and a discriminative network to supplement residual images or feature maps with facial attribute information significantly reduced the ambiguity in face super-resolution [28].

Although the convolutional neural network has been much developed in super-resolution reconstruction of the human face, there are still some problems, such as poor recovery ability of human face details and slow optimization speed of the deep network and so on. To solve these problems, a face super-resolution algorithm based on an improved deep residual network (IDRN) is proposed in this paper. IDRN network optimizes the internal structure of the residual unit, and by fusing the output characteristics of different convolution layers, it strengthens the information flow among the layers within the unit, by removing batch normalization layer to ensure the integrity of image spatial information. At the same time, the network adopts the feature-intensive fusion strategy, which fuses the output of each residual unit across layers to provide more detail information for reconstructed images. The proposed method is validated by an experiment carried on FDDB [29] and

AFLW [30] face data sets. The results show that the reconstruction results of the IDRN network are better. The PSNR and SSIM values of the reconstructed images with different scales are higher than that of other several super-resolution reconstruction methods.

2 Proposed method

2.1 Improved deep residual network

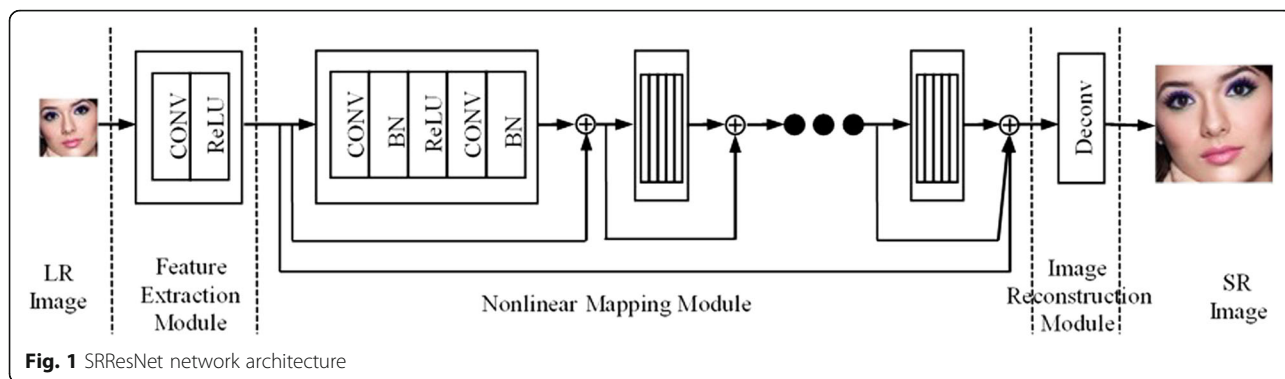
The network structure of image super-resolution reconstruction model (SRResNet) [31] based on depth residual network is shown in Fig. 1. The network consists of a feature extraction module, non-linear mapping module, and image super-resolution reconstruction module. Usually, the non-linear mapping module contains several residual units. The shallow feature of LR image is obtained by the feature extraction module, and the residual information between LR and HR is obtained by non-linear mapping module. Then the residual information and shallow feature are added together and input to super-resolution reconstruction module for image reconstruction. The final output is the SR image.

Kim et al. have proved that deep convolution network can improve the quality of super-resolution reconstructed images [7], but with the increase of network depth, the gradient disappearance or gradient explosion problem is likely to appear in model training. In this paper, on the basis of SRResNet in Fig. 1, an improved deep residual network (IDRN) shown in Fig. 2 is proposed. In the IDRN network, there are two improved aspects: the improved residual unit and the improved residual network. On the one hand, the improved residual units are improved by changing the constituent elements and improving the connection of the network structure, which enhances the performance of local residual learning. On the other hand, the output feature maps of each residual unit are combined, so the model detail recovery ability is also improved, and the network structure changes from global residual to global-local residual, which avoided over-fitting in model training.

2.2 Improvement of residual unit

The traditional residual unit is shown in Fig. 3a, in which the batch normalization layer can accelerate the convergence speed of the network and improve the stability of the training. It can essentially solve the gradient dispersion phenomenon in network training. However, some literatures have pointed out that it can also destroy the spatial information of the image, increase the learning burden of the depth network, and degree the performance of the super-resolution reconstruction network [32, 33].

In this paper, the traditional residual unit is improved. As shown in Fig. 3b, the original two BN layers are removed and the ReLU is added to the output part to improve the non-linear expression ability of the residual unit, which avoid the damage to the spatial information



of the image by BN layer. In addition, by using the “shortcut” connection structure, the input of each convolution layer is induced to its output, and all the combined feature map is passed to the next convolution layer together. Thereby, the densely connected convolutional layer enhanced the information flow between the inner convolutional layers, and the information transferability of the improved residual unit becomes stronger.

2.3 Improvement of residual network

Residual learning can reduce the training burden of a deep network, solve the problem of gradient disappearance and gradient explosion, and ensure the good learning ability of neural network. In order to strengthen the role of residual network, we propose a cross-layer integration strategy. The new network structure, as shown in Fig. 4b, is improved by using shortcut connection to combine the output feature map of each residual unit. And the network structure changed from global residual to global-local residual, which avoided over-fitting in model training and the phenomenon of gradient disappearance in gradient backpropagation is avoided. At the same time, we combine the output characteristics of each residual unit on the basis of the global residual structure, so the output features of each residual unit

are fully utilized to provide more abundant feature information for the reconstruction module and the model detail recovery ability is also improved. The improvement of the residual network is shown in Fig. 4; Fig. 4a is the original residual network structure, and Fig. 4b is the improved residual network structure.

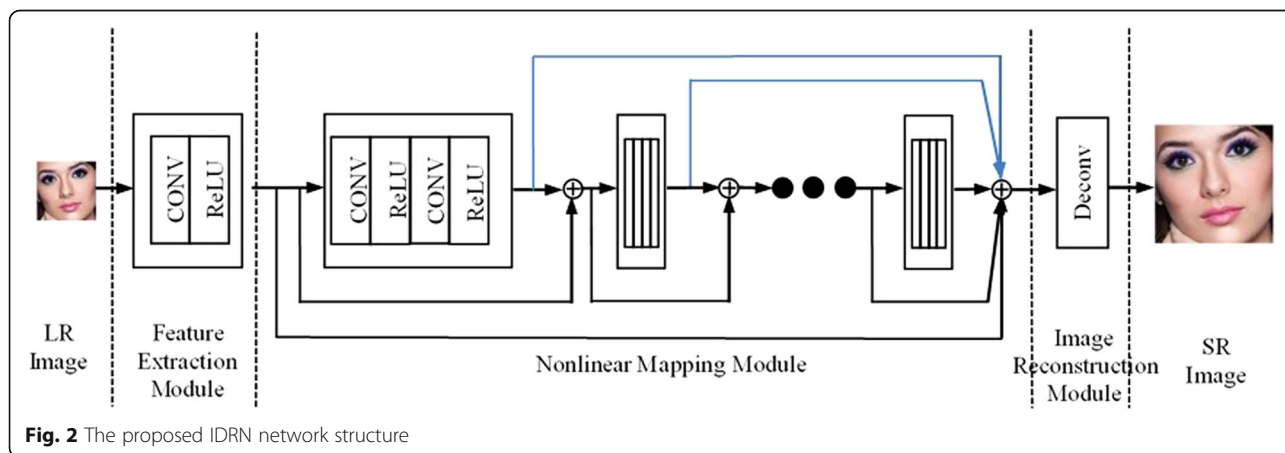
In Fig. 4, X_0 denotes the output characteristics of feature extraction network, i.e., shallow features of low-resolution image; X_i is the output characteristics of the i th residual unit; X_{out} is the sum of the shallow features and all the residual units output. For an IDRNet with N residual units, the X_{out} is:

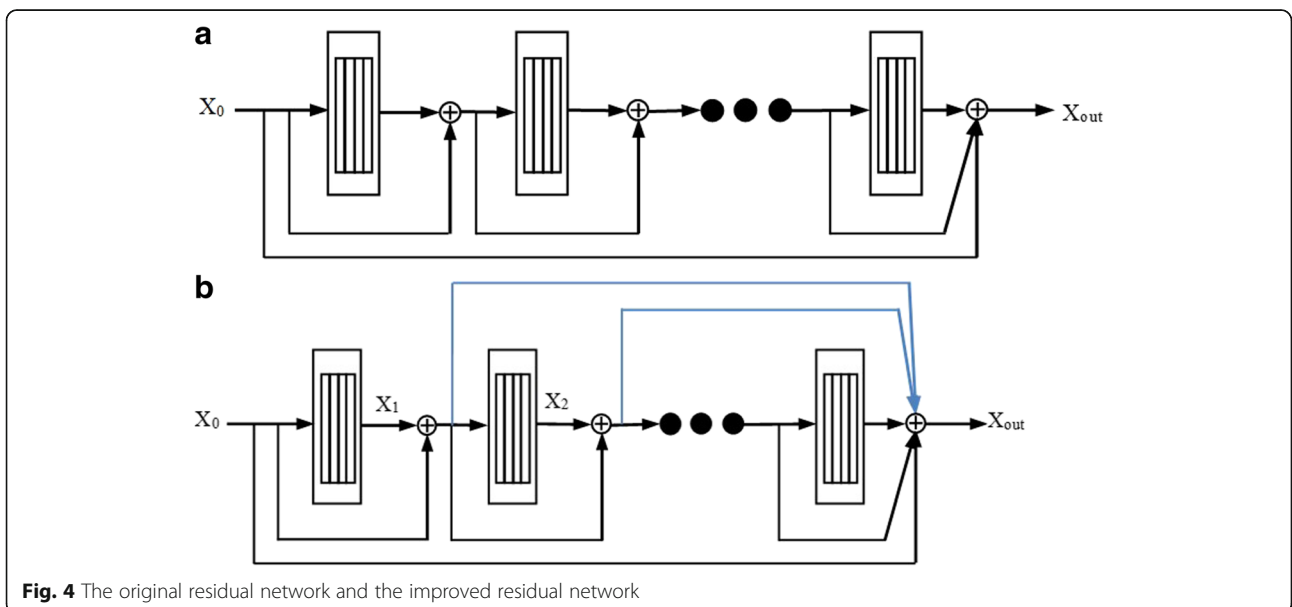
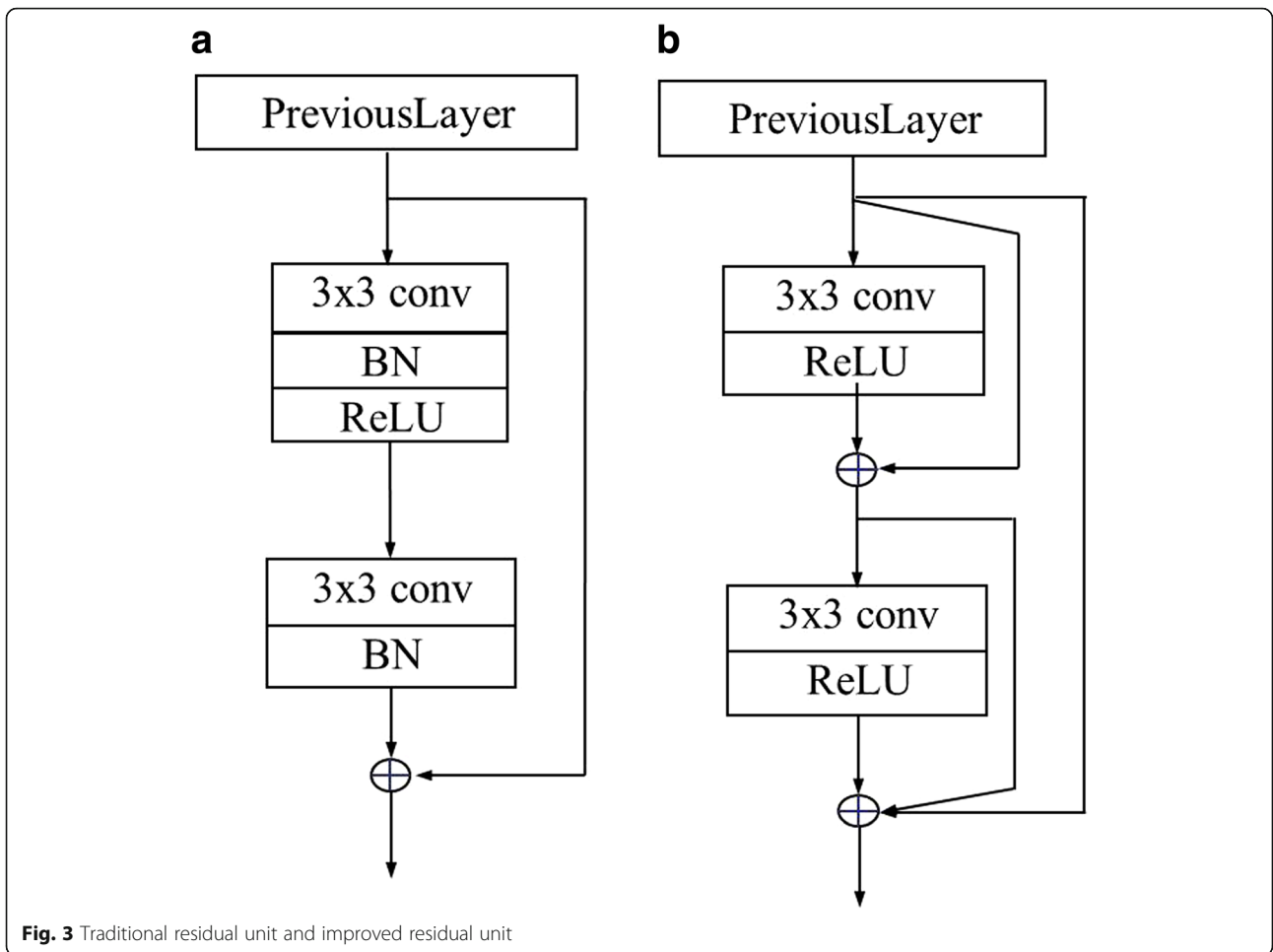
$$X_{out} = \sum_{i=0}^N X_i = X_0 + X_1 + \dots + X_N \tag{1}$$

3 Experiments

3.1 Data preparation and network parameters setting

In this paper, a total of 23500 images from Fddb and AFLW face databases are used as experimental samples, 500 images are randomly selected as test sets, and the rest 23,000 images are as training sets. Because some high-frequency information will be lost in the low-resolution image after interpolation downsampling of





high-resolution image, the original face image is down-sampled by bicubic interpolation (BI), and the face images of 32 *32, 64 *64, 96 *96, and 128 *128 sizes are obtained respectively for super-resolution reconstruction of multi-scale face images.

Based on Ubuntu 1604 operating system and Nvidia GeForce GTX 1060 (6GB) independent graphics card, an experimental platform is built and TensorFlow open-source framework and python programming language is employed. The super-resolution reconstruction of the human face under different conditions are experimented and compared with other methods. In the experiments, low-resolution images of 32 *32 were input and batch size was set to 16. The whole training process of the network includes 120 epochs. The initial learning rate is set to 0.0001. The learning rate of every 30 epochs decreases to 90% of the original. In order to

reduce the complexity of network training and accelerate the convergence speed of the network, all input images are normalized, and Adam optimization method is adopted. The momentum parameter is set to 0.9. In the training process, mean square error (MSE) is used as the loss function.

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|F(x_i; \theta) - \hat{x}_i\|^2 \tag{2}$$

In Eq. (2), $L(\theta)$ is the loss function of the network; x_i is the LR image; $F(x_i; \theta)$ is the SR image generated by the network; \hat{x}_i is the HR image; and the parameters θ are the weights and bias of all convolution layer. $L(\theta)$ is non-negative, the smaller the value is, the smaller the deviation between the reconstructed SR image and the ideal HR image is, the better the reconstruction result is.

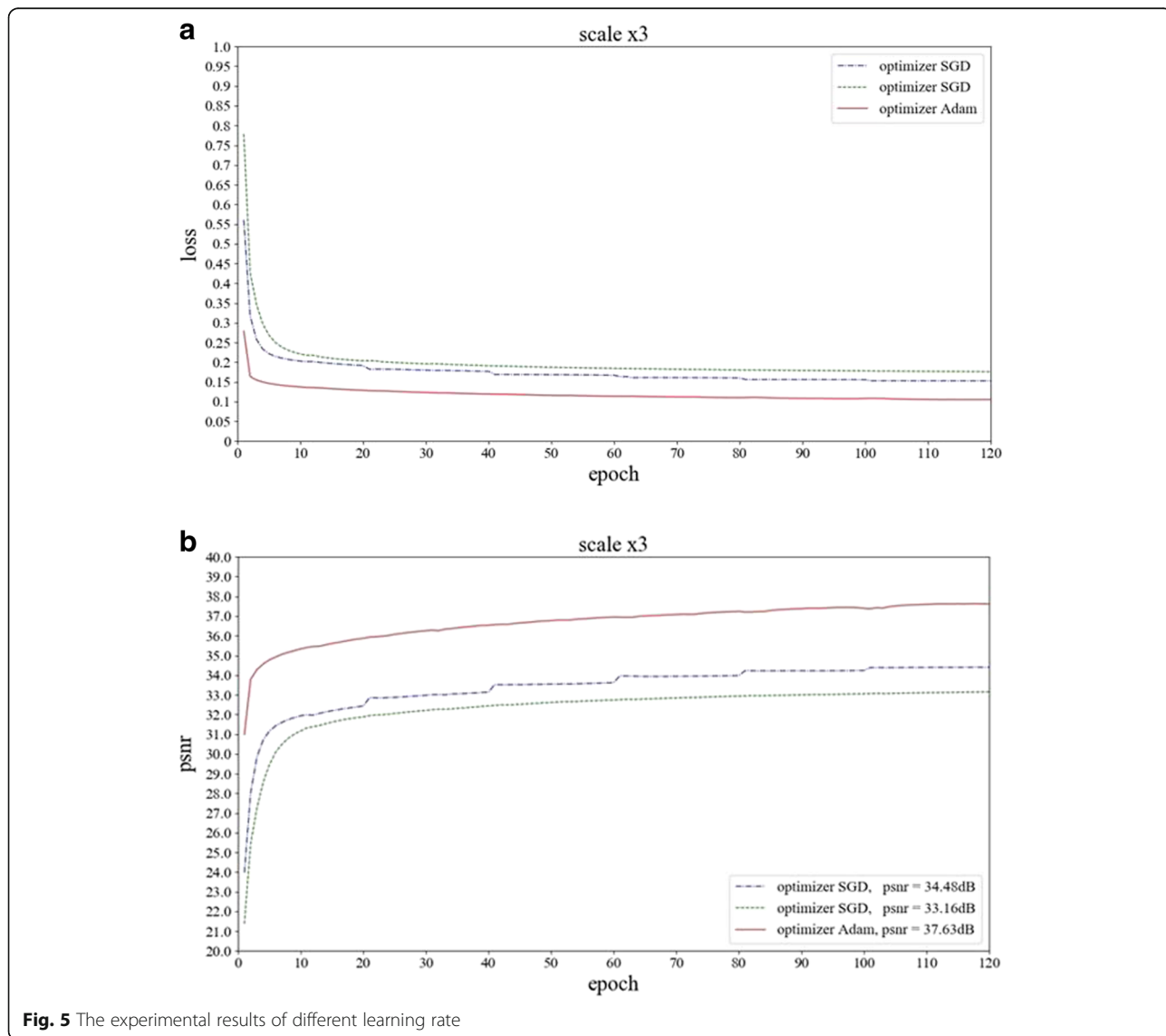


Fig. 5 The experimental results of different learning rate

3.2 Evaluation indexes

In this paper, peak signal-to-noise ratio (PSNR) and structural similarity index method (SSIM) are used as objective quantitative evaluation indexes for face super-resolution reconstruction images. PSNR represents the ratio of variance between signal and noise in an image. The PSNR between two images is

$$PSNR = 10 \cdot \lg\left(\frac{255^2}{MSE}\right) \tag{3}$$

Here, $MSE = \frac{1}{N} \sum_{n=1}^N [\frac{1}{HW} \sum_{j=1}^H \sum_{j=1}^W (\hat{Y}_n(i, j) - Y_n(i, j))^2]$

represents the mean square error between two images. N denotes the number of training samples in each batch; H and W denote the height and width of the image respectively. $\hat{Y}_n(i, j)$ and $Y_n(i, j)$ represent the (i, j) pixel values of

reconstructed image and HR image respectively for n th sample image. The higher the PSNR value between the two images, the higher the fidelity of the reconstructed image and the better the reconstruction effect.

SSIM is an index to measure the structural similarity between two images. It mainly involves three aspects: brightness, contrast, and structural characteristics of the images. SSIM is

$$SSIM(x, y) = \frac{(2u_x u_y + c_1)(2\sigma_{x,y} + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{4}$$

Here, x and y represent the reconstructed image and HR image respectively. u_x, u_y, σ_x^2 , and σ_y^2 represent the mean and variance of image x and y respectively. $c_1 = k_1 L$ and $c_2 = k_2 L$ are constant terms that are used to avoid the

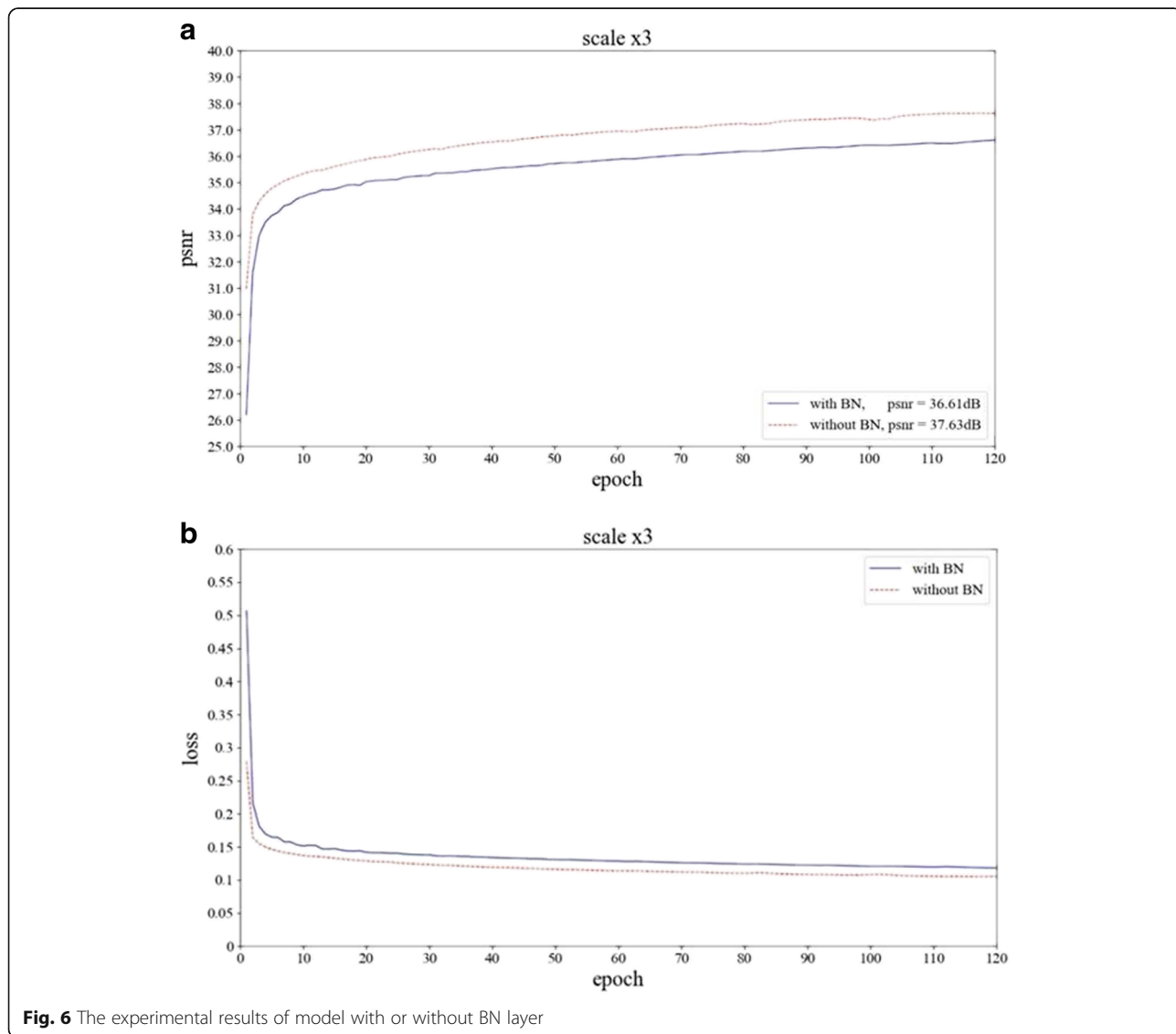


Fig. 6 The experimental results of model with or without BN layer

Table 1 The average PSNR (dB), SSIM, and run time (ms) under different network depth

	DIRN 8	DIRN 12	DIRN 16	DIRN 20
PSNR/dB	32.29	32.13	32.31	32.32
SSIM	0.873	0.891	0.904	0.907
Time/ms	4.4	4.8	5.0	5.7

zero denominator or molecule. γL is the dynamic range of image values. k_1 and k_2 usually are smaller decimal numbers, such as 0.001. The value range of image structure similarity is [0,1]. The closer the SSIM value is to 1, the higher the similarity between the two images is.

4 Results and discussion

4.1 The effect of different learning rates on model reconstruction

In model training, the learning rate will have an important impact on the training speed and reconstruction performance of the model. If the learning rate is too big, the gradient of the network is easy to oscillate near the minimum point, which causes the model to not converge; if the learning rate is too small, the training time of the model will be greatly prolonged.

Here, three schemes are proposed and experimented the effect of learning rate on model reconstruction. The first one is, learning rate is fixed to 0.0001, and SGD algorithm is used to optimize the gradient; the second one is initial learning rate is 0.0001 and the learning rate reduces by 0.1 times every 20 epochs, at the same time, SGD algorithm is used also; the third one is, initial learning rate is 0.0001, using Adam algorithm is adopted

for gradient optimization, which can adjust the learning rate adaptively according to the network training situation. The experimental results are shown in Fig. 5, where a shows the loss changing with the epoch and b shows that of the PSNR.

In Fig. 5, green stands for the first scheme, blue for the second, and red for the third. From Fig. 5, we can see the blue loss line that shows the learning rate gradually decreases during network training for there is stepped descent every 20 epoch, which can make the model converge further; the green loss line smoothly and slowly declines, indicating that the model is hovering around the local extremum and the network is no longer convergent after 20 epoch because the learning rate is constant value; the red loss line indicates that the model keeps a relatively suitable and smooth step for optimization due to the Adam algorithm can adjust the learning rate adaptively according to the training situation of the network. The good performance of the third scheme also can be seen from Fig. 5b. Therefore, Adam algorithm with initial learning rate of 0.0001 is chosen for our model training.

4.2 Effect of BN layer on reconstruction

By batch normalization (BN) layer, the stability of network training can be guaranteed and the speed of model training can be accelerated. However, it has been pointed out that the normalization operation can destroy the image spatial information and affect the performance of the super-resolution reconstruction model. Therefore, two reconstruction models based on residual network are constructed and work under the same experimental conditions to compare the effects of BN. The experimental results are shown in Fig. 6.

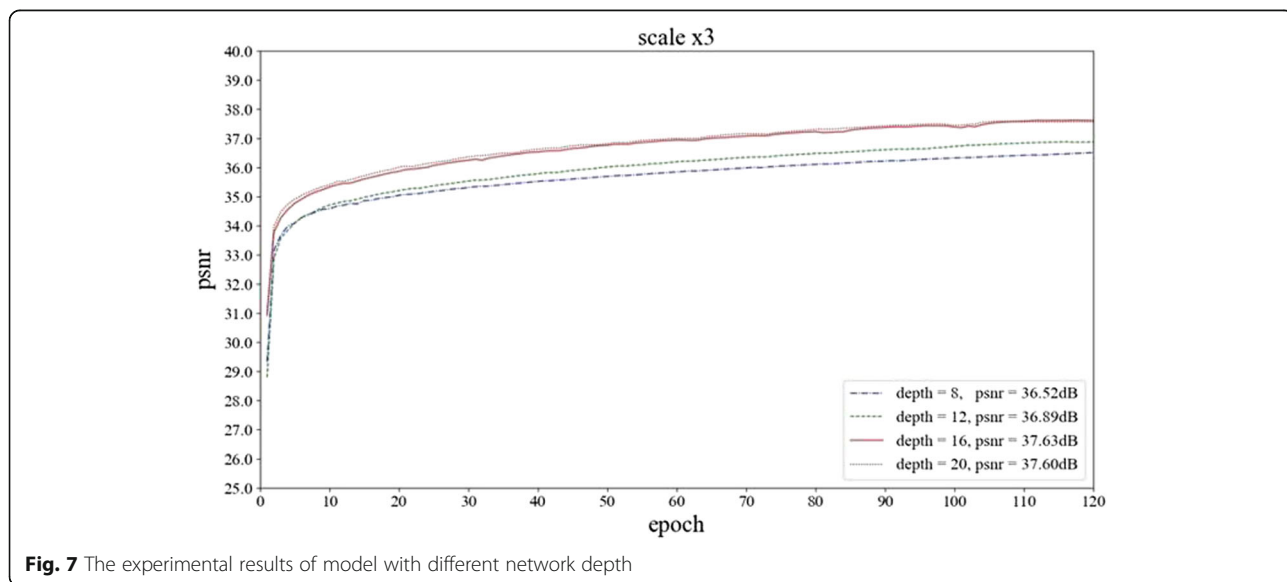


Fig. 7 The experimental results of model with different network depth

Table 2 The comparison for IDRN under different network width

	IDRN-64	IDRN-128	IDRN-192
PSNR/dB	32.29	32.49	32.55
SSIM	0.904	0.923	0.926
Time/ms	4.35	6.29	7.96

In Fig. 6, the blue line stands for the result with BN and the red one for without BN. Figure 6a shows that both the initial value and the following values of PSNR in the reconstruction model without BN layer is higher than that of the model with BN layer in the whole processing, which indicates that the performance of the reconstruction model will be adversely downgraded by BN layer. Figure 6b shows that the model without the BN layer has better convergence effect in the generally.

4.3 Contrast experiments under different network depths

Generally, as the network depth (i.e. the number of residual units) increases, the better the reconstructed image can be obtained, but it also increases the complexity and time consumption of the network, so it is necessary to find a balance among multiple mutually constrained factors. In this paper, the reconstruction under the models at depths of 8, 12, 16, and 20 are experimented. In the experiments, the face images in the test set are reconstructed with single-scale super-resolution (magnification scale is 4). PSNR, SSIM, and time cost are used as evaluation indexes. The experimental results are shown in Table 1. By comparison, the average PSNR and SSIM values of the reconstruction models are improved with the increase of network depth while computing time increase slightly. From the PSNR and SSIM indexes, the model DIRN 20 achieves the best reconstruction effect. Considering PSNR, SSIM, and time-consuming, the depth of the network model in this paper is selected as 16.

Figure 7 shows the PSNRs obtained under reconstruction models with different network depths. It shows that the PSNR is higher when network depth is bigger; however, when the network depth is 16, the PSNR under network depth 16 is much closer to that of 20. Considering the speed and the amount of calculation, 16 for network depth is optimal.

4.4 Contrast experiments under different network widths

In order to discuss the influence of network width (i.e., the number of convolution cores) on the performance of reconstruction models, the reconstruction models with different network widths are experimented and compared under the conditions of magnification scale 4 and network depth 8. The network widths selected to experiment are 64, 128, and 192 (corresponding models were named as "IDRN-64," "IDRN-128," and "IDRN-192") respectively. The experiment results are shown in Table 2. It shows that the performance of the reconstructed model DIRN-192 is the best, and the PSNR and SSIM values are the highest. Compared with IDRN-128, the PSNR value of IDRN-192 increases slightly, but its model complexity is much higher and time consumption is greater. Therefore, considering the three indexes, the network width of IDRN model is 128, which PSNR and SSIM are relatively high, and the time consumption is moderate.

4.5 Contrastive experiments with different reconstruction methods

To further discuss the reconstruction effect of the proposed algorithm, the IDRN with network depth of 16 and width of 128 is experimented and compared with several main image super-resolution reconstruction algorithms (such as Bicubic, SRCNN, VDSR, SRResNet) under different magnification (that is scale = 2, 3, 4). The experimental results are shown in Table 3. It shows that with the increase of image magnification, the difficulty of image super-resolution reconstruction much increases. However, under the same scale, the PSNR and SSIM of the proposed algorithm are higher than that of the other four super-resolution reconstruction algorithms.

In addition, we also randomly select three images in the test set and use different methods to reconstruct a super-resolution image with scale = 4. The results are shown in Fig. 8. Compared with other four reconstruction methods, the method of this paper restores the details of the eyebrows, eyes, face, and other parts better, the image is clearer and has better visual effects.

5 Conclusions

In view of the problems of insufficient detail recovery and slow network optimization existing in face super-resolution reconstruction, this paper improves the deep residual network from the residual unit and network

Table 3 The comparison of different SR algorithms

Scale	Bicubic		SRCNN		VDSR		SRResNet		IDRN	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
2	32.76	0.931	36.11	0.962	37.42	0.977	38.97	0.985	39.64	0.986
3	30.35	0.893	32.41	0.927	35.38	0.960	35.52	0.965	36.44	0.972
4	27.93	0.851	29.01	0.874	32.80	0.915	32.54	0.931	33.23	0.930

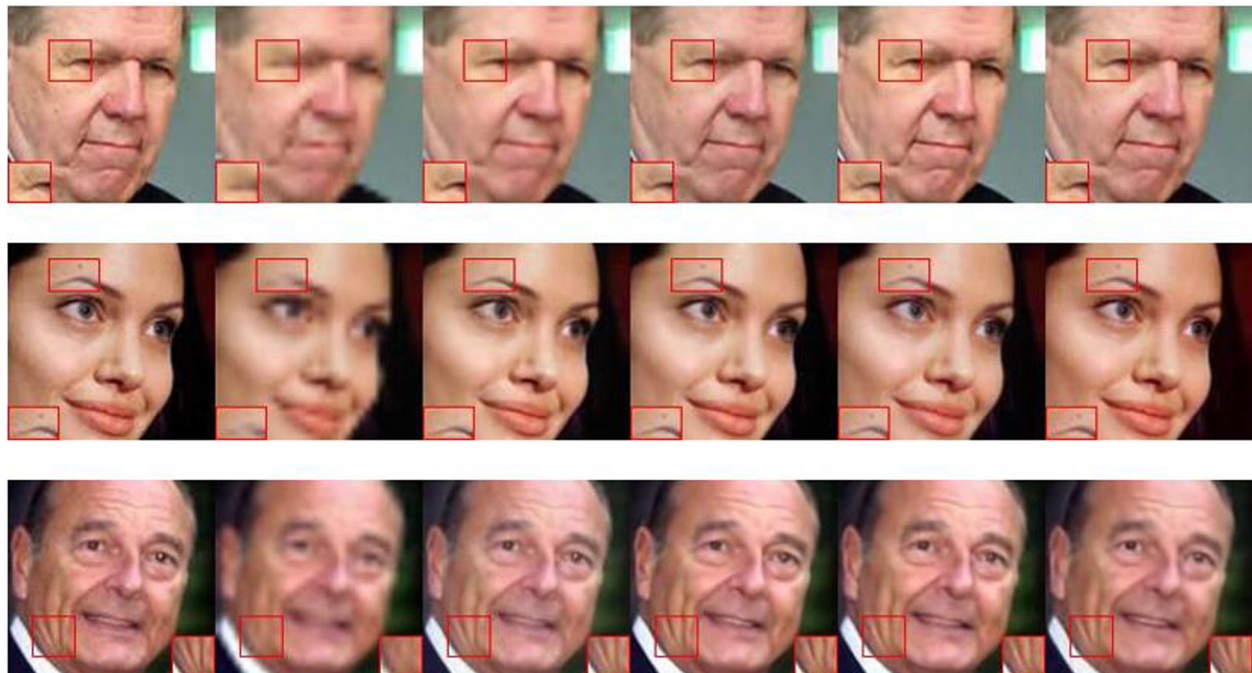


Fig. 8 Reconstructed images by different methods (scale = 4)

structure. The improvements are firstly to increase the information flow between different convolutional layers within the unit, eliminates the damage of batch normalization processing to image spatial information. And secondly, apart from global residual learning, local features are intensively fused together to speed up network optimization processing and enhance the network's ability to describe detailed features. This paper discusses the influence of network depth and width on the reconstruction effect, so as to determine the network parameters considering both the quality and time-consuming, and also compares the reconstruction indexes and reconstructed images of the similar methods. The results show that compared with the current mainstream super-resolution reconstruction algorithms, the proposed algorithm has stronger feature description ability, better image detail recovery, and better image quality.

Abbreviations

BN: Batch normalization; CNN: Convolutional neural network; DRCN: Recursive convolutional network; ESPCNN: Efficient sub-pixel convolutional neural network; GANs: Generative adversarial nets; HR: How resolution; IDRN: Improved deep residual network; LR: Low resolution; MSE: Mean square error; NE: Neighbor embedding; PSNR: Peak signal-to-noise ratio; ReLU: Rectified linear unit; SC: Sparse coding; SR: Superresolution; SRCNN: Super resolution CNN; SSIM: Structural similarity index method; VDCN: Very deep convolutional network

Acknowledgements

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions. I would like to acknowledge all our team members.

Authors' contributions

All authors took part in the discussion of the work described in this paper. All authors read and approved the final manuscript.

Authors' information

Di Fan, Doctor of Control theory and control engineering, Associate Professor. She graduated from Shandong University of Science and Technology in 2010 and worked in Shandong University of Science and Technology. Her research interests include digital signal and image processing, machine vision, and deep learning.

Shuai Fang, Master student of electronics and communication engineering of Shandong University of Science and Technology. His research field is deep learning, target detection, and super resolution construction.

Guangcai Wang, Master student of communication and information systems of Shandong University of Science and Technology. Her research direction is image processing and analysis, and face attribute.

Shang Gao, Master student of electronics and communication engineering of Shandong University of Science and Technology. Her research direction is deep learning and machine vision.

Xiaoxin Liu, Master student of electronics and communication engineering of Shandong University of Science and Technology. Her research direction is deep learning and machine vision.

Funding

This research was partially aided by subproject of project 863(2015AA016404-4), Shandong Province Young Scientist Foundation (BS2012DX034), China Postdoctoral Science Foundation (2012M521361), Shandong Province Natural Science Foundation (ZR2012EEM021), Project of Shandong Province Higher Educational Science and Technology Program (J13LN17), SDUST Research Fund (2010KYTD101), and Project of South Africa/China Research Collaboration in Science and Technology (2012DFG71060).

Availability of data and materials

Please contact the corresponding author for data requests.

Competing interests

The authors declare that they have no competing interests.

Received: 24 March 2019 Accepted: 14 June 2019

Published online: 19 July 2019

References

- P. Meisen, Y. Xiaoli, T. Jingtian, Research on Interpolation Methods in Medical Image Processing. *J. Med. Syst.* **36**(2), 777–807 (2012)
- D. Rajan, S. Chaudhuri, in *Super-Resolution Imaging*. Generalized Interpolation for Super-Resolution (Springer, Boston, 2002), pp. 45–72
- H. Aftab, A.B. Mansoor, M. Asim, in *2008 IEEE International Multitopic Conference*. A new single image interpolation technique for super resolution (IEEE, 2008), pp. 592–596
- X. Gu, J.X. Du, X.F. Wang, Leaf Recognition Based on the Combination of Wavelet Transform and Gaussian Interpolation. *Lect. Notes Comput. Sci* **3644**, 253–262 (2005)
- Y. Pang, Y. Mao, A Super-resolution Image Reconstruction Algorithm on the Theory of Projection Onto Convex Sets (POCS). *Comput. Eng. Appl.* **41**(4), 69–71 (2005)
- Z. Shu-Ping, Super-resolution sequential image reconstruction based on SURF registration method and POCS. *J. Comput. Appl.* **39**(10), 46–53 (2012)
- X. Ma, S. Dai, Study on Improvement of Image Super-Resolution Algorithm Based on Residual Network. *Softw. Guid.* **4**, 1672–7800 (2018)
- S. Chen, X. Xie, Y. Yang, Q. Lian, Image Super-Resolution Algorithm Based on Multi-Scale Convolution Neural Network. *J. Signal Process.* **9**, 1003–0530 (2018)
- H. CHANG, D.-Y. YEUNG, Y. XIONG, in *CVPR 2004: Processings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Super-resolution through neighbor embedding (IEEE Computer Society, Washington, DC, 2004), pp. 275–282
- R. Timofte, V. De Smet, L. Van Gool, in *Proceedings of the IEEE international conference on computer vision*. Anchored neighborhood regression for fast example-based super-resolution (2013), pp. 1920–1927
- J. Yang, J. Wright, T. Huang, Y. Ma, Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **19**(11), 2861–2873 (2010)
- X. Lu, H. Yuan, P. Yan, Y. Yuan, X. Li, Geometry constrained sparse coding for single image super-resolution. *Proc. IEEE Conf. Comput. Vis. Pattern Recognition (CVPR)* **157**(10), 1648–1655 (2012)
- Y. Lecun, L. Bottou, Y. Bengio, et al., Gradient-Based Learning Applied to Document Recognition [J]. *Proc. IEEE* **86**(11), 2278–2324 (1998)
- C. Dong, C.C. Loy, K. He, et al., in *European conference on computer vision*. Learning a deep convolutional network for image super-resolution[C] (Springer, Cham, 2014), pp. 184–199
- J. Kim, J. Kwon Lee, K. Mu Lee, in *Proceedings of the IEEE conference on computer vision and pattern recognition*. Accurate image super-resolution using very deep convolutional networks (2016), pp. 1646–1654
- J. Kim, K. Lee, M. Lee, in *Proceedings-29th IEEE Conference on Computer Vision and Pattern Recognition*. Deeply-recursive convolutional network for image super-resolution (2016), pp. 1637–1645
- W. Shi, J. Caballerom, F. Huszar, et al., *Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network* (2016), pp. 1874–1883
- G. Huang, Z. Liu, V.D.M. Laurens, et al., in *IEEE Conference on Computer Vision and Pattern Recognition*. Densely Connected Convolutional Networks (2017), pp. 2261–2269
- T. Tong, G. Li, et al., in *IEEE International Conference on Computer Vision (ICCV)*. Image Super-Resolution Using Dense Skip Connections (2017), pp. 4809–4817
- I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., in *Advances in neural information processing systems*. Generative adversarial nets (2014), pp. 2672–2680
- C. Ledig, Z. Wang, W. Shi, et al., *Photo-realistic single image super-resolution using a generative adversarial network* (2016), pp. 105–114
- E. Zhou, H. Fan, Z. Cao, et al., in *Twenty-Ninth AAAI Conference on Artificial Intelligence*. Learning face hallucination in the wild, vol 7 (2015), pp. 3871–3877
- Z. Wang, D. Liu, J. Yang, et al., in *2015 IEEE International Conference on Computer Vision (ICCV)*. Deep Networks for Image Super-Resolution with Sparse Prior (2015), pp. 370–378
- S. Zhu, S. Liu, C.C. Loy, et al., in *Computer Vision - ECCV 2016*. Deep Cascaded Bi-Network for Face Hallucination (Springer International Publishing, Cham, 2016), pp. 614–630
- Y. Sun, H. Song, K. Zhang, F. Yan, Face super-resolution via very deep convolutional neural network. *J. Comput. Appl.* **38**(04), 1141–1145 (2018)
- Y. Chen, Y. Tai, X. Liu, et al., in *Computer Vision and Pattern Recognition*. FSRNet: End-to-End Learning Face Super-Resolution with Facial Priors (2018)
- A. Bulat, G. Tzimiropoulos, in *Computer Vision and Pattern Recognition*. Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs (2018)
- X. Yu, B. Fernando, R. Hartley, et al., in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes (IEEE, 2018)
- C. Ledig, L. Theis, F. Huszar, et al., in *Computer Vision and Pattern Recognition*. Photo-realistic single image super-resolution using a generative adversarial network (2016), pp. 4681–4690
- Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *Comput. Sci.*, 448-456,2015
- Z. Yang, K. Zhang, Liang, et al., in *International Conference on Multimedia Modeling*. Single image super-resolution with a parameter economic residual-like convolutional neural network (Springer, Cham, 2017), pp. 353–364
- V. Jain, E.G. Learned-Miller, *Fddb. A benchmark for face detection in unconstrained settings*. UMass Amherst Technical Report (2010)
- M. Köstinger, P. Wohlhart, P.M. Roth, et al., in *IEEE International Conference on Computer Vision Workshops, ICCV 2011 Workshops*. Annotated Facial Landmarks in the Wild: A large-scale, real-world database for facial landmark localization (IEEE, Barcelona, 2011), pp. 6–13

6 Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)