Super-resolution reconstruction of seismic section image via multi-scale convolution neural network

Meng-Di Deng¹, Rui-Sheng Jia^{2}, Hong-Mei Sun², and Xing-Li Zhang¹*

¹ College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China.

² Shandong Province Key Laboratory of Wisdom Mine Information Technology, Shandong University of Science and Technology, Qingdao 266590, China.

Abstract. The resolution of seismic section images can directly affect the subsequent interpretation of seismic data. In order to improve the spatial resolution of low-resolution seismic section images, a super-resolution reconstruction method based on multi-scale convolution is proposed. This method designs a multi-scale convolutional neural network to learn high-low resolution image feature pairs, and realizes mapping learning from low-resolution seismic section images to high-resolution seismic section images. This multi-scale convolutional neural network model consists of four convolutional layers and a sub-pixel convolutional layer. Convolution operations are used to learn abundant seismic section image features, and sub-pixel convolution layer is used to reconstruct high-resolution seismic section image. The experimental results show that the proposed method is superior to the comparison method in peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). In the total training time and reconstruction time, our method is about 22% less than the FSRCNN method and about 18% less than the ESPCN method.

1 Introduction

The seismic section image is obtained by two-dimensional mapping of seismic data in a computer. It can achieve the purpose of visualizing seismic data, and can directly display seismic fault or section information, which can more intuitive analyze the region of concern. The spatial resolution of seismic section images can directly affect the subsequent interpretation of seismic data [1-4]. In order to improve the resolution of seismic section images, domestic and foreign scholars have proposed a number of improved methods, which are roughly classified into two categories: (1) processing the signal data itself [5, 6]. (2) encrypting seismic data by increasing the number of exploration channels. However, both methods require a large number of sensors in the acquisition process to ensure the

*Corresponding author: jrs716@163.com

validity of the seismic data, resulting in increased exploration costs. The super-resolution reconstruction technology can reconstruct high-resolution seismic section images with fewer sensors, which can effectively save seismic exploration costs.

At present, the technology of image super-resolution reconstruction has been widely used in medical imaging, satellite remote sensing imaging, face recognition and video surveillance and other fields [7-9]. According to the idea of image reconstruction, it can be divided into three categories: interpolation-based method [10], reconstruction-based method [11] and learning-based method [12, 13]. The interpolation-based method [14-16] combines the weights of adjacent pixels to get the interpolated pixel value. This kind of method has the advantages of simple operation and fast processing speed, but it is easy to lose high frequency information of the image, resulting in image blurring, which is not conducive to the subsequent processing of seismic images. The reconstruction-based method [17-18] introduces the prior information of images into SR reconstruction model as constraint condition. This kind of method better overcomes the shortcomings of the first type of method, but it faces the problems that the prior information is difficult to select and the modeling parameters are difficult to estimate. Learning-based method is currently a hot research topic in the field of super-resolution reconstruction. Freeman et al. proposed a super-resolution reconstruction method based on sample learning [19]. Yang et al. introduced the theory of compressed sensing and proposed a super-resolution reconstruction method based on sparse representation (SCSR) [20, 21]. Timofte et al. proposed a fast super-resolution reconstruction method based on adjusted neighborhood regression(A+) [22]. Schuler et al. proposed a super-resolution forest method (SRF) [23]. Inspired by the deep learning method, Dong et al. applied the deep learning method to image super-resolution reconstruction in 2016, and proposed a super-resolution reconstruction method for single-frame image using convolutional neural networks (SRCNN) [24]. The method uses three convolution layers to map image features from low-resolution space to high-resolution space. It avoids the defect of artificial design features, realizes end-to-end learning, and has better reconstruction effect than other traditional method. However, since the SRCNN is reconstructed after the interpolation operation, that is, the convolution operation is performed on the higher resolution image, this limits the reconstruction performance of the method. In order to solve the above problems, Dong et al. improved the SRCNN method and proposed a fast image super-resolution reconstruction algorithm based on deconvolution (FSRCNN) [25]. The input is Low-resolution images, and the deconvolution layer is used to scale up the images at the end of the network to get high resolution images. Later, Shi et al. proposed a super-resolution reconstruction method based on sub-pixel convolution layer (ESPCN) [26], which rearranges feature images through sub-pixel convolutional layers to obtain high-resolution images.

The seismic section image is similar to many natural images, and it has obvious texture features, so the super-resolution reconstruction method can be applied to seismic image processing. However, there are some differences between seismic images and traditional images. For example, the texture features of seismic section images are relatively simple, over-deep convolution neural network easily leads to network degradation and over-fitting phenomenon; The most obvious boundary of seismic images is in the reflection horizon, and the boundary with sinusoidal feature is not as obvious as that of other images. In order to better reconstruct the texture detail information in the seismic section image, the texture detail enhancement should be fully considered when designing the super-resolution reconstruction method. To this end, this paper proposes a seismic section image reconstruction method based on multi-scale convolutional neural network. The method has five layers and uses multi-scale convolution kernels to learn more detailed texture features. At the same time, multiple dimensionality reduction operations are performed in the model, which effectively reduces the complexity of the algorithm.

2 Proposed method

2.1 Reconstruction model

The super-resolution reconstruction process designed in this paper is shown in Figure 1, including the training process and the reconstruction process. In the training process, the high-resolution seismic section images transformed from RGB to YCbCr format, and the Y-channel information is extracted to obtain the high-resolution image components. The low-resolution image component is obtained by down sampling the high-resolution image components. In order to enlarge the number of training samples, the high-resolution image components and the low-resolution image components are segmented separately to obtain the low-resolution image blocks Xl and the high-resolution image blocks Xh. Finally, Xl and Xh paired input into convolutional neural network for training. During the reconstruction process, the input is a low-resolution seismic section image. Convert low-resolution image from RGB format to YCbCr format, and extract its Y channel and Cb, Cr channel information respectively. The Y channel image component is input into the trained convolutional neural network for reconstruction, and the image components of Cb and Cr channels are enlarged by interpolation. Finally, the reconstructed high-resolution seismic section image is obtained through channel fusion and format conversion.

Fig. 1. Super-resolution reconstruction process based on convolutional neural network.

In the reconstruction process of seismic section image, the structure of convolutional neural network directly affects the reconstruction quality of the images. The reconstruction method proposed in this paper establishes a multi-scale convolutional neural network model. The input is low-resolution image, and the output is a high-resolution image. Since down sampling reduces image detail information, there is no pooling layer in the model, and the model is a full convolution structure. The network model can be divided into three parts: feature extraction, nonlinear mapping and reconstruction. The network model structure is shown in Figure 2.

Fig. 2. Neural network model of multi-scale convolution.

1. Feature extraction. The feature extraction stage extracts image blocks from low-resolution images and represents each block as a high dimensional vector. These vectors form a set of feature maps, whose number is equal to the dimension of the vector. In this paper, 64 5×5 convolution kernels are used to extract and characterize the features of low-resolution seismic section image blocks.

2. Nonlinear mapping. In the stage of nonlinear mapping, the idea of multi-scale convolution is used to map low-resolution seismic section features into high-resolution seismic section features. This stage can be called the multi-scale convolution module of the network, which consists of three layers of convolution. Firstly, 1×1 convolution kernels are used to reduce the dimension of feature maps. Secondly, 3×3 and 5×5 convolution kernels are used to perform convolution operations. Then Concat layer is used to join feature maps of convolution layers at different scales. Finally, 1×1 convolution kernels are used to fusion multi-scale features. Multi-scale convolution module can be expressed as:

$$y = \phi(w_3 * F(x) + b_3) \quad (1)$$

Where x and y represent the input and output of the multi-scale convolution module, ϕ represents the activation function, w_3 and b_3 represent the weight and biase of the third-layer convolution in the multi-scale convolution module respectively. “*” denotes the convolution operation. $F(x)$ is a splicing of two-scale nonlinear mapping:

$$F(x) = \text{Concat}(f_1(x), f_2(x)) \quad (2)$$

Where $f_i(x), i=1,2$ represent the output of the nonlinear mapping for each scale respectively

$$f_i(x) = \phi(w_2^i * \phi(w_1^i * x + b_1^i) + b_2^i), i=1,2 \quad (3)$$

Where $w_1^i, b_1^i, i=1,2$ represent the weights and biases of the first layer convolution in this module respectively. $w_2^i, b_2^i, i=1,2$ represent the weights and biases of the second layer convolution in this module respectively. The multi-scale convolution module uses two-scale convolution kernels, that is, 32 3×3 convolution kernels and 8 5×5 convolution kernels. The nonlinear mappings of each scale are composed of two layers of convolution.

The use of multi-scale convolution kernels has two major advantages. First, convolution kernels of different sizes can extract features of different scales of the image in order to extract and learn more abundant image information. Secondly, the convolutional neural network training model is realized by learning the parameters (weights and biases) of the filter, that is, continuously learning the parameters of the filter to achieve an optimal value close to the label. In this paper, multi-scale convolution kernels are used to make the convolution layer have a variety of filters, so that the learning of weights and biases is more diverse, and the useful information of the image can be extracted and learned fully and effectively. At the same time, since the 1×1 convolution kernels are used for dimensionality reduction before multi-scale convolution, the number of parameters of the module is reduced from 32512 to 14656, which reduces the complexity of the algorithm and makes the training speed and reconstruction speed a certain degree of improvement.

3. Reconstruction. Subpixel convolution layer is used to reconstruct high-resolution seismic section image. The sub-pixel convolution operation can be divided into two parts. First, use r^2 (r is the target magnification factor) 3×3 convolution kernels for convolution operation, and then the feature maps of r^2 channels are rearranged by pixels. That is, r^2 feature maps of size $H \times W$ are rearranged into a high-resolution image of $rH \times rW$ size. Subpixel convolution operations can be expressed as:

$$Y = PS(w_{sp} * y + b_{sp}) \quad (4)$$

Where y is the output of the nonlinear mapping phase, w_{sp} and b_{sp} represent weights and offsets of convolution layer, PS represents sub-pixel convolution operation, and Y represents the high-resolution image obtained by super-resolution. Upsampling of the image is achieved by the sub-pixel convolution layer. On the one hand, because the resolution of the input image is low, a smaller filter can be used for the convolution operation, which greatly reduces the complexity of the method. On the other hand, no interpolation method is used means that the network can learn the useful information in the image more clearly, and can better learn the mapping between low-resolution image blocks and high-resolution image blocks. It can improve the quality of reconstructed images.

In this paper, the linear correction unit ReLU [27] (Rectified linear unit) is selected as the activation function, which can greatly improve the convergence speed. The ReLU function has the characteristics of unilateral inhibition and sparse activation, which is closer to the activation state of the brain neurons receiving signals. Using ReLU as an activation function introduces sparsity for convolutional neural networks, which is equivalent to the introduction of unsupervised learning. For a network composed of L layers, let X_l be the input of layer l and Y_{l+1} be the output, under the ReLU activation function, the relationship between input and output is:

$$Y_{l+1} = \max(0, w_l * X_l + b_l) \quad (5)$$

Where w_l , b_l , $l \in (1, L-1)$ are learnable network weights and biases respectively.

The definition of the loss function is extremely important for the performance of the neural network. The training process for generating the model is the optimization process for the loss function. The Mean Square Error (MSE) is closer to human visual perception. Therefore, MSE is chosen as the loss function of the training network. The mathematical expression is:

$$MSE = \frac{1}{n} \sum_{i=1}^n \|Y_i - y_i\|^2 \quad (6)$$

Where n is the number of training samples, and the physical meaning of the MSE is to calculate the difference of centre pixel between the original high-resolution image block Y_i and the high-resolution image block y_i obtained by super-resolution.

2.2 Algorithm flow

The original training samples are pre-treated to generate training samples $\{D^i, d^i\}$, D^i is a high-resolution image block, and d^i is its corresponding low-resolution image block. The training samples are input into the multi-scale convolution neural network for training. The super-resolution reconstruction algorithm is shown in Table 1. Where m is the number of training samples contained in each batch, i is the i -th input image in each batch, L is the total number of layers of the convolutional neural network, Z^l represents the output of layer $l-1$, PS represents the sub-pixel convolution operation, δ^l is the error of the layer l , α is the learning rate and σ is the activation function. After the training is completed, a convolutional neural network composed of W and b of each hidden layer and output layer is obtained.

Table 1. Super-resolution reconstruction algorithm

Algorithm: Super-resolution reconstruction algorithm
Input: Training sample pair $\{D^i, d^i\}$
Output: convolutional neural network consisting of W and b of each hidden layer and output layer
Step:
1. initialize W and b of each hidden layer and output layer
2. Repeat
3. For $i \leftarrow 1$ to m
4. Input D^i
5. For $l \leftarrow 1$ to $L-2$
6. $D^{i,l+1} \leftarrow \sigma(Z^{i,l+1}) \leftarrow \sigma(W^l * D^{i,l} + b^l)$
7. End
8. $D^{i,L} \leftarrow \sigma(Z^{i,L}) \leftarrow PS(W^{L-1} * D^{i,L-1} + b^{L-1})$
9. $\delta^{i,L} \leftarrow \nabla_a C \leftarrow \sigma'(Z^{i,L})$
10. For $l \leftarrow L-2$ to 1
11. $\delta^{i,l} \leftarrow (W^{i,l+1})^T \delta^{i,l+1} \leftarrow \sigma'(Z^{i,l})$
12. End
13. End
14. For $l \leftarrow L-1$ to 1
15. $W^l \leftarrow W^l - \frac{\alpha}{m} \sum_i \delta^{i,l} (D^{i,l-1})^T$
16. $b^l \leftarrow b^l - \frac{\alpha}{m} \sum_i \delta^{i,l}$
17. End
18 Update Batch
19. Until reaching the stop condition

3 Experiment

3.1 Data set

In order to study the texture features of seismic section images better, 114 seismic section images are selected in this experiment, of which 96 are used as original training samples and 18 are used as test samples. Since the seismic section image does not have an official data set, the experimental data set is obtained from the network database. The data set image is shown in Figure 3.

Fig. 3. Seismic section image data set

3.2 Evaluation index

The metrics widely used to evaluate the image super-resolution performance are Peak Signal-to-Noise Ratio (PSNR) [28] and Structural Similarity (SSIM) [28]. PSNR and SSIM were chosen for quantitative evaluation in this paper. The formula for calculating PSNR is as follows:

$$PSNR = 10 \lg \frac{255^2}{\frac{1}{W * H} \sum_{i=1}^W \sum_{j=1}^H (I_{i,j}^{HR} - I_{i,j}^{SR})^2} \quad (7)$$

Where $W * H$ is the size of the image, $I_{i,j}^{HR}$ is the original image, and $I_{i,j}^{SR}$ is the high-resolution images obtained by super-resolution. PSNR is used to quantitatively calculate the error between the reconstructed image and the original image. The higher the PSNR value, the smaller the distortion of the reconstructed image. The formula for calculation SSIM is:

$$SSIM(x, y) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1} * \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} * \frac{\sigma_{xy} + C_3}{\sigma_x^2 + \sigma_y^2 + C_3} \quad (8)$$

Where u, σ are the pixel mean and variance of the two images, respectively, and C_1, C_2 and C_3 represent constants to keep the denominator zero. SSIM illustrates the degree of similarity between the reconstructed image structure and the original image structure. The closer the value of SSIM is to 1, the better the quality of the reconstructed image generated.

3.3 Experimental results and analysis

The experimental environment consisted of hardware devices and software configurations. The computer used for the test was configured as Intel Core i5-3230M CPU @ 2.6 GHz with 12 GB of memory. The operating system on the experimental platform is 64-bit Windows 7, Matlab 2012b and caffe [29] deep learning framework.

Fig. 4. Method reconstruction result graph. (a) Original image. (b) Low resolution image. (c) Manually magnified image. (d) Reconstructed image by super-resolution reconstruction.

The experiment uses the LR image blocks and the corresponding labels as the input of the network. The upscale factor is 3, the training process uses the SGD optimizer, the momentum is set to 0.9, the learning rate is set to 0.0001, and the training iteration is 600,000 times. The experimental result diagram of this experiment is shown in Figure 4. A test image is randomly selected in the test set with a resolution of 489×390 , as shown in Figure 4(a). After down sampling, the low-resolution image is obtained with a resolution of 163×130 , as shown in Figure 4(b). Without changing its resolution, the Figure 4(b) is manually enlarged to obtain the Figure 4(c). It can be seen that the detail of the image is blurred. It will be difficult to carry out subsequent seismic data processing. In this paper, the super-resolution reconstruction method is used to reconstruct the low-resolution image, as shown in Figure 4(d). The high-resolution seismic section image with a resolution of 489×390 is obtained, which can reconstruct the detailed texture information of the image more accurately.

Randomly select 6 test images in the test set, and low-resolution image LR were obtained after 3 times down sampling. The reconstructed high-resolution images were obtained by Bicubic, ScSR, FSRCNN, ESPCN and our method, respectively. Figure 5 shows the experimental results of three of these images. The partial area in the test image is selected for enlargement. It can be seen that the edges of the reconstructed image by traditional bicubic are blurred and the texture details are not obvious. The artifacts appear on the edge of the image obtained by ScSR and A+. The FSRCNN, ESPCN and our method have achieved good reconstruction results.

Fig. 5. Reconstruction results of different method

In this paper, PSNR and SSIM are used to objectively evaluate the test results of six images. Table 2 gives the results of objective evaluation. It can be seen from the table that except for the test image 4, the PSNR and SSIM obtained by our method are higher than the comparison method, and have a greater degree of improvement than Bicubic, ScSR, and A+.

Table 2. Comparison of different method evaluation indicators

Image	Index	Bicubic	ScSR	A+	FSRCNN	ESPCN	Our Method
Test 1	PSNR	19.2501	19.7342	19.8441	20.7561	21.0691	21.3940
	SSIM	0.7252	0.7584	0.7801	0.8069	0.8314	0.8421
Test 2	PSNR	23.0494	23.5795	23.6651	23.9587	23.9973	24.1106
	SSRM	0.7971	0.8080	0.8009	0.8029	0.8077	0.8288
Test 3	PSNR	19.9570	20.3196	20.4241	20.8833	20.9495	21.0630
	SSIM	0.8489	0.8498	0.8505	0.8508	0.8583	0.8604
Test 4	PSNR	24.5830	24.9370	25.0115	25.2375	25.4764	25.4560
	SSIM	0.7894	0.8549	0.8599	0.8651	0.8674	0.8653
Test 5	PSNR	23.9598	24.7736	24.8761	25.4564	25.7060	25.7909
	SSIM	0.8523	0.8617	0.8671	0.8994	0.8841	0.8957
Test 6	PSNR	24.6501	25.1342	25.3175	25.9561	26.0691	26.3447
	SSIM	0.7742	0.7784	0.7921	0.8069	0.8114	0.8170

In order to compare the effective improvement of proposed method with FSRCNN and ESPCN, a series of training and learning processes were carried out. During the training process, a model is saved every 3000 iterations. According to the three different network learning models obtained by the FSRCNN method, ESPCN method and the proposed method under the same number of iterations, the above test images are super-resolution processed to obtain corresponding PSNR and SSIM values. Taking test image 3 as an example, the network model obtained by training different iterations is tested and recorded to obtain the trend of PSNR and SSIM, as shown in Figures 6.

Fig. 6. Iteration number and PSNR/SSIM relationship curve. (a) The PSNR value of test image 3 varies with the number of iterations. (b) The SSIM value of test image 3 varies with the number of iterations.

It can be seen from Figure 6 that the results obtained by the proposed method are better than FSRCNN and ESPCN in both PSNR and SSIM. In the training process, the convergence speed of the proposed method is faster. The experimental results show that the proposed method only needs to iterate 20w times to achieve better reconstruction results. In

order to avoid the contingency of the experiment. Train the network and reconstruct the test set image at 2, 3, and 4 times magnification. Table 3 shows the average PSNR and SSIM for the different methods on the test set. It can be found that the PSNR and SSIM of the proposed method are still higher than the comparison method under different magnification ratios, which fully demonstrates that the high-resolution seismic section image reconstructed by our method is closer to the original high-resolution seismic section image, and the reconstruction effect is better.

Table 3. Average PSNR and SSIM for different methods on the test set

Scale	Index	Bicubic	ScSR	A+	FSRCNN	ESPCN	Our Method
×2	PSNR	26.8067	27.3523	27.4810	27.7965	27.7976	27.9421
	SSIM	0.8753	0.8791	0.8802	0.8864	0.8857	0.8901
×3	PSNR	22.7570	23.2196	23.3205	23.7833	23.8495	24.0030
	SSIM	0.8441	0.8577	0.8591	0.8628	0.8663	0.8704
×4	PSNR	19.0494	19.5775	19.6825	19.9814	20.0017	20.1106
	SSIM	0.7771	0.7981	0.8023	0.8107	0.8111	0.8288

In order to verify the timeliness of the model, five test images were randomly selected in the test set. Bicubic, ScSR, A+, FSRCNN, ESPCN and our method were used to reconstruct and record the reconstruction time. The experimental results are shown in Table 4. It can be seen from the table that the Bicubic method has the shortest time to reconstruct an image, but the above analysis shows that the reconstruction effect is poor. Compared with other comparison methods, the reconstruction speed of our method is improved to some extent, which fully demonstrates that the proposed method achieves better time performance while maintaining high reconstruction accuracy.

Table 4. Time(s) for reconstruction of images by different methods

Image	Bicubic	ScSR	A+	FSRCNN	ESPCN	Our Method
Test 1	0.0079	38.9170	0.5214	0.2639	0.2542	0.2016
Test 2	0.0028	21.1752	0.4149	0.2031	0.2006	0.1605
Test 3	0.0071	35.1492	0.5010	0.1801	0.1726	0.1542
Test 4	0.0047	24.2485	0.4816	0.2095	0.1984	0.1605
Test 5	0.0034	22.1486	0.4018	0.2236	0.2114	0.1699
Average	0.0052	28.9277	0.4641	0.2160	0.2074	0.1693

4 Conclusion

At present, seismic section images are mainly improved by seismic data denoising and improved seismic data acquisition methods. The emergence of CNN theory has brought new ideas to improve the resolution of seismic section images. In the process of seismic data exploration, less sensors can be used to collect seismic data, and the super-resolution reconstruction can be used to recover the detailed features by convolutional neural network, which solves the problem of high acquisition cost of seismic survey. In order to improve the reconstruction quality of seismic section images, we propose a multi-scale convolutional neural network method to learn the feature information of seismic section images at different scales through multi-scale convolution kernels. At the same time, the

1×1 convolution kernels are used to reduce the dimension, which effectively reduces the complexity of the method. The experimental results in this paper show that under the training set and test set of seismic section images, the proposed network structure has better reconstruction effect and higher efficiency than the other representative methods.

Acknowledgement: The authors are grateful for collaborative funding support from the Natural Science Foundation of Shandong Province, China (ZR2018MEE008).

References

1. O'Malley, S., Kakadiaris, I, *IEEE Conf. Computer Vision and Pattern Recognition* (2004)
2. Admasu, F., Back, S. Toennies, K. J. *Geophysics* **71:6**, 49 (2006)
3. Tingdahl, K., De Rooij, M., *Geophysical Prospecting. J.* **53:4**, 533 (2005)
4. Pedersen, S. I., Randen, T., Sonneland, L., Steen, O., Soc. Explorational Geophysicists 72nd Annu. Meeting (2005)
5. Wu, J., Ruan, Q., *The 18th International Conference on Pattern Recognition*, **3**, 718 (2006)
6. Zhang, X., Ye, H., Zhang H., Multi-Channel Wavelet-Based Diffusion Method for Denoising DTI Images. *BMEI*, 178 (2018)
7. Nasrollahi, K., Moeslund, T. B., Super-resolution: A comprehensive survey. *Mach. Vis. Appl*, **25:6**, 1423 (2014)
8. Mandal, S., Bhavsar, A., Sao, A. K. Noise adaptive super-resolution from single image via non-local mean and sparse representation. *Signal Process*, **132**, 134 (2017)
9. Liu, D., Wang, Z., Wen, B., Yang, J., Han, W., Huang, T. S., *IEEE Trans. Image Process* **25:7**, 3194 (2016)
10. Batz, M., Eichenseer, A., Seiler, J., Jonscher, M., Kaup, A. *IEEE Conf. Comput. Vis. Pattern Recog* 58 (2015)
11. Kim, K. I., Kwon, Y., *IEEE Trans. Pattern Anal. Mach. Intell*, **32:6**, 1127 (2010)
12. Schulter, S., Leistner, C., Bischof, H., *IEEE Conf. Comput. Vis. Pattern Recog*, 3791 (2015)
13. Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., et al. Photo-realistic single image super-resolution using a generative adversarial network. *Conf. Comput. Vis. Pattern Recog.*, 4681 (2020)
14. Keys, R. G. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Processing* **29:6**, 1153 (1981)
15. Zhou, F., Yang, W., Liao, Q. Interpolation-based image super-resolution using multisurface fitting. *IEEE Trans. Image Process* **21(7)**, 3312 (2020)
16. Hung, K.-W., Siu, W.-C., *IEEE Trans. Image Process*, **21:3**, 1061 (2012)
17. Ren, C., He, X., Teng, Q., Wu, Y., Nguyen, T. Q., *IEEE Trans. Image Process*, **25:5**, 2168-2183 (2016)
18. Zhu, J., Zhou, C., Fan, D., Zhou, J. A. *Journal of Nanomaterials* (2014)
19. Freeman, W. T., Jones, T. R., Pasztor, E. C., *Computer Graphics & Applications IEEE*, **22:2**, 56 (2002)

20. Yang, J., Wright, J., Huang, T. S., Ma, Y., *IEEE Trans. Image Process*, **19:11**, 2861 (2010)
21. Yang, J., Wang, Z., Lin, Z., Cohen, S., Huang, T., *IEEE Trans. Image Process*, **21:11**, 3467 (2012)
22. Timofte, R., De Smet, V., Van Gool, L., *Proc. Asian Conf. Comput.*, 111 (2014)
23. Schuler, S., Leistner, C., Bischof, H., *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3791 (2015)
24. Dong, C., Loy, C. C., He, K., Tang, X., *IEEE Trans. Pattern Anal. Mach. Intell*, **38:2**, 295 (2016)
25. Dong, C., Loy, C. C., Tang, X., *Proc. ECCV*, 391 (2016)
26. Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., *IEEE Conf. Comput. Vis. Pattern Recog.*, 1874 (2016)
27. Ye, Y., Shan, J., Bruzzone, L., Shen, L., *IEEE Trans. Geosci. Remote Sens*, **55:5**, 2941 (2017)
28. Gong, C., Liu, T., Tao, D., Fu, K., Tu, E., Yang, J., *IEEE Trans. Neural Netw. Learn. Syst*, **26:10**, 2261 (2015)
29. Jia, Y. Caffe: Convolutional architecture for fast feature embedding. *ACM Multimedia* 2, **4** (2014)