Research on Remote Sensing Image Restoration from Low Resolution to High Resolution Based on Deep Learning

Cong Gao[1], Yang Zhao[2], Chuanchang Si[1]

[1.]Unit 31401 of PLA, Qingdao, 266000, China
[2.]Unit 71282 of PLA, Baoding, 071000, China

Keywords: Deep learning, Image resolution, Image reconstruction.

Abstract: Image super-resolution reconstruction is an image processing technique that processes and analyzes a low-resolution image or a series of images by computer to recover the desired high-resolution image. In this paper, the deep convolutional network is the main research object, and the improved algorithm is proposed for its deficiency, which improves the reconstruction performance and robustness of the algorithm. Aiming at the shortcomings of poor recovery and poor image detail, combined with the strong learning ability of deep network, deep neural network is introduced based on traditional sparse coding algorithm, and a better SR image method is found. Good experimental results were obtained on the dataset. The experimental results on the dataset show that the improved combination method can effectively improve the image reconstruction quality.

1. Introduction

Digital image and its related processing technology are one of the important contents of information processing technology and are widely used in many fields [1]. However, the degradation of digital images during acquisition and processing is a further development of image processing technology [2]. More and more researchers are working to solve the problem of image degradation, so image super-resolution reconstruction technology came into being. Super-resolution reconstruction technology not only can obtain ideal images with good visual effects, but also has broad application prospects in the fields of medicine, remote sensing, monitoring, military and so on [3].

In recent years, deep learning methods have led the wave of technological development [4]. Its fiery development makes up for the shortcomings of traditional shallow learning methods. On the basis of shallow learning, domestic and foreign scholars have proposed an image SR reconstruction method based on deep learning method, and achieved excellent reconstruction performance in a large number of fields [5,6]. The main idea of the frequency domain algorithm is to improve the resolution of the image by eliminating aliasing in the frequency domain. Related researchers have proposed the earliest concept of image SR reconstruction, they use a single image reconstruction to carry out algorithm research and calculation [7,8]. After that, people have studied the SR reconstruction algorithm in many aspects, and proposed many new algorithms [9]. Although they have improved at a certain point and obtained good experimental simulation results under certain assumptions, they are actually used. The effect does not achieve satisfactory results. Some scholars have made further improvements and extensions to Tsai's model, but they have made a fuss about the scope and speed of the model [10]. There is no breakthrough. Due to the shortcomings of the frequency domain-based image SR reconstruction method, it is difficult to increase the a priori constraint and the flexibility is poor. The research focus of the image SR reconstruction method is slowly shifted from the frequency domain to the airspace.

In this paper, by combining with the deep network, we improve the performance of the traditional sparse coding method. On the other hand, the training convergence of the deep network requires a lot of training data and time. Here we only make the residual image between the low resolution input image and the original image and the recovered Sc SR (Sparse coding Super-Resolution) image. It greatly reduces the training time of our deep network. After the training is completed, we have done a

  DOI: 10.25236/aisct.2019.065

sampling test on the data set. Through the comparison of the obtained test results, it is found that the method obtained after introducing the deep network has better performance on the scale factor than the original ScSR method.

2. Deep learning basic theory and image SR reconstruction related algorithm

2.1 Convolutional neural networks

The convolutional neural network, first proposed by Professor Yann Le Cun and his colleagues, is a neural network model with multiple hidden layers, including input layer, convolutional layer, pooling layer, fully connected layer and output layer. Its network structure is shown in Figure 1.

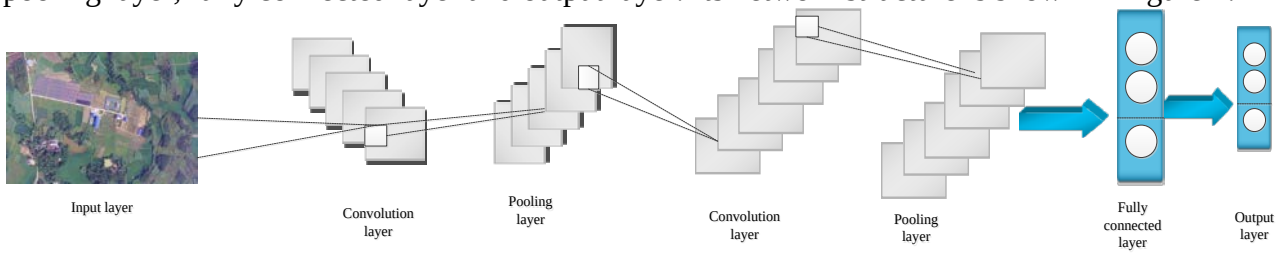

Figure 1 The basic structure of the convolutional neural network

The input layer is the input to the entire neural network. In a neural network in the field of image processing, an input is generally a three-dimensional pixel matrix representing a picture. The length and width of the matrix represent the image size, and the depth represents the color channel of the image. For example, the grayscale depth is 1, and the image depth is 3 in RGB color mode. The image is converted from the input layer to a pixel matrix, and then the input matrix is converted by a convolutional layer with different weights and output to the next layer until the final output layer.

The convolutional layer is the most important component of a neural network. In the convolutional layer, the local features of the input image are extracted by convolution operations, and different sizes of convolution kernels extract features of different scales of the image. Multiple convolution kernels can be set in the convolution layer to extract multiple different features of the image. .

The pooling layer in the convolutional neural network is usually located behind the convolutional layer and forms a convolution structure with the convolution and excitation layers. The pooling layer does not change the depth of the input 3D matrix, but reduces the size of the matrix by performing aggregate statistics on the local image features. This reduces the number of nodes in the fully connected layer and reduces the network parameters.

The excitation layer is usually located after the convolutional layer and forms a convolution structure with the convolution and pooling layers. The excitation layer acts to make a nonlinear transformation of the input data and passes it to the next layer.

2.2 SdSR method

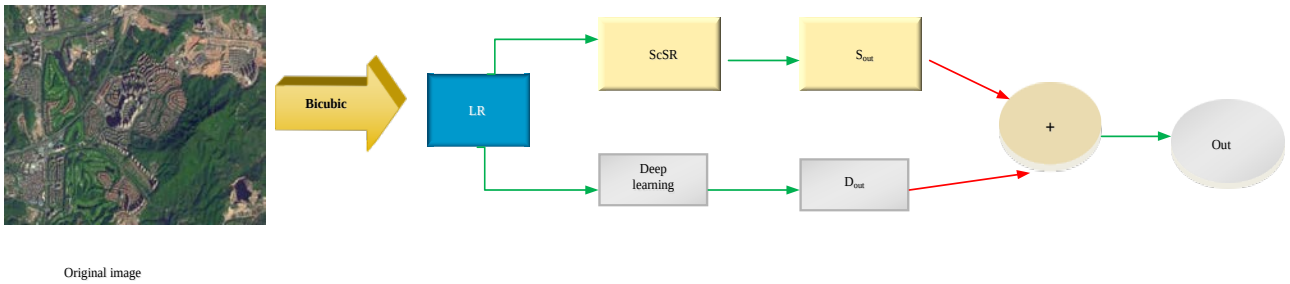

Figure 2 Network structure of the SdSR method

As shown in Figure 2, for SR image reconstruction, we first perform a bicubic interpolation downsampling operation on the original image to obtain low resolution input images at scale factors 2, 3, and 4, respectively. Then, through the learned sparse dictionary and the depth network, the SR is

restored to the input low-resolution image to obtain two restored images. The sparse dictionary portion gets the sparsely encoded restored image, and the deep network portion obtains the residual image between the original image and the sparsely restored image. Finally, the obtained two restored images are subjected to the corresponding position element addition operation to obtain the final output. Below, we detail the composition of the two parts.

In the reconstruction module, due to the single-scale nature of the sparse dictionary, we adopt single-scale reconstruction, reconstruct the feature map after up-sampling recovery and output the final output. The reconstruction module consists of seven convolutional layers with ReLU as a nonlinear mapping. The first four convolutional layers consist of 64 filters with a convolution kernel size of 33. The feature maps obtained from the previous layer of each layer are inputs, and a new 64-channel feature map is generated. The fifth layer is a dimensionality reduction layer composed of 16 filters with a convolution kernel size of 11, and outputs a 16-channel feature map. The sixth layer restores and reconstructs the feature map, and finally feeds the obtained feature map to the seventh layer composed of a filter having a convolution kernel size of 11 for weighting operation and outputs the final SR image.

Our training is divided into two parts. One is sparse dictionary training with scale factors of 2, 3, and 4, and the other is deep network training.

3. Experimental results and analysis

3.1 Training and test data sets

1) Training data set

In this experiment, the training data we took 1600 images randomly extracted from the Image Net data set. Image Net datasets have greatly contributed to the development of deep learning. Image Net is an image dataset built on the World Net hierarchy (currently limited to nouns), with each node in the hierarchy consisting of hundreds or thousands of images. Currently, the Image Net dataset has more than 14 million images covering more than 20,000 categories, with an average of more than 500 images per node. More than one million of them have clear category labels and annotations of object locations in the image. Image Net dataset is a dataset widely used in deep learning image processing, such as image classification, positioning, detection, reconstruction and so on.

2) Test data set

For the test data, we mainly use the data set "Set5" and do the sampling test on the data sets "Urban100", "BSD100" and "Image Net". Here we only tested on three common scale factors 2, 3, and 4.

3.2 Experimental results on the Set5 data set

The experimental results are shown in Table 1.

Table 1 Test results on the Set5 data set

PSNR	Baby	Bird	Butterfly	Head	Woman
ScSR(2)	41.01	41.21	34.01	36.87	37.81
ScSR(2)	40.29	40.29	33.76	35.98	37.83
ScSR(3)	39.77	39.67	32.87	25.67	33.01
ScSR(3)	38.99	39.01	31.76	25.21	33.02
ScSR(4)	38.29	38.29	30.42	34.19	31.20
ScSR(4)	36.87	37.65	28.98	34.01	31.11

Table 1 shows the experimental results of the two methods on the Set5 data set under scale factors 2, 3, and 4 respectively, and the blue font indicates the best performance on the current scale factor. From the data in the table, it can be intuitively found that the method we use on the data set Set5 and the traditional ScSR method have advantages and disadvantages on the Set5 data set.

3.3 Experimental results on the BSD100 dataset

The experimental results are shown in Table 2. Because there are many images in the BSD100 dataset, we extract the first ten images in the BSD100 dataset as test images. For the convenience of representation, we simply represent 10 images as B1-B10.

Table 2 Sample test results on the B100 data set

PSNR	B1	B2	B3	B4	B5
ScSR(2)	35.01	31.06	37.01	31.01	31.90
SdSR(2)	34.98	30.57	36.87	30.98	30.37
ScSR(3)	33.87	30.01	36.03	30.61	29.76
SdSR(3)	32.65	29.65	33.65	29.76	28.65
ScSR(4)	31.87	29.41	32.08	29.01	27.89
SdSR(4)	30.41	28.38	32.10	28.90	26.98
PSNR	B6	B7	B8	B9	B10
ScSR(2)	31.01	30.21	26.01	33.02	40.34
SdSR(2)	30.78	30.01	25.98	33.01	40.01
ScSR(3)	30.01	29.76	24.76	32.98	39.87
SdSR(3)	29.76	28.54	23.09	31.98	38.05
ScSR(4)	28.61	28.01	22.56	30.67	37.66
SdSR(4)	27.54	27.74	21.69	29.76	36.69

Table 2 shows the experimental results of the two methods on the B100 dataset under scale factors 2, 3, and 4, respectively. The blue font indicates the best performance on the current scale factor. From the data in the table, we can intuitively see that the SdSR method we use is superior to the traditional ScSR method in the BSD100 data set.

3.4 Experimental results on the Ur100 dataset

The experimental results are shown in Table 3. Due to the large number of images in the Ur100 dataset, we extracted the first five images in the Ur100 dataset as test images. For ease of representation, here we simply represent 5 images as U1-U5.

Table 3 Sample test results on the Ur100 data set

PSNR	U1	U2	U3	U4	U5
ScSR(2)	31.01	24.01	28.65	29.01	25.76
SdSR(2)	30.98	23.87	28.45	28.92	25.20
ScSR(3)	30.56	23.02	27.97	27.88	24.01
SdSR(3)	29.86	22.97	27.90	26.44	23.56
ScSR(4)	29.01	22.81	26.01	25.39	22.87
SdSR(4)	28.54	21.53	25.20	24.01	22.04

Table 3 shows the experimental results of the two methods on the Ur100 dataset under scale factors 2, 3, and 4, respectively. The blue font indicates the best performance on the current scale factor. From the data in the table, we can intuitively obtain that the SdSR method we use is superior to the traditional ScSR method in the Ur100 data set.

4. Conclusion

This paper studies an SR method for sparse coding combined with deep networks. Based on the image restoration of the original sparsely encoded SR, the residual image of the original image and the sparsely restored image is obtained by adding the depth network part. Finally, the two parts are restored and the corresponding position elements are added to obtain the final weight. A large number of experimental results show that compared with the original sparse coding SR method, our method

not only achieves a higher PSNR index, but also improves the details of texture and edge structure. In the future, we should try to combine more different networks with traditional methods, adjust the network structure, and optimize network parameters in order to achieve better recovery results.

References

[1] Strack R, Deep learning advances super-resolution imaging, *Nature Methods*, vol.15, pp. 403-403, 2018.

[2] Alkus U, Ermeydan E S, Sahin A B, et al, Enhancing the image resolution in a single-pixel sub-THz imaging system based on compressed sensing, *Optical Engineering*, vol.57, pp. 1, 2018.

[3] Cheng J, Cai C, Luo J, Reconstruction of high-resolution early-photon tomography based on the first derivative of temporal point spread function, *Journal of Biomedical Optics*, vol.23, pp. 1, 2018.

[4] Huang H, Yang J, Huang H, et al, Deep Learning for Super-Resolution Channel Estimation and DOA Estimation Based Massive MIMO System, *IEEE Transactions on Vehicular Technology*, vol.67, pp. 8549-8560, 2018.

[5] Kim K, Wu D, Kuang G, et al, Penalized PET Reconstruction Using Deep Learning Prior and Local Linear Fitting, *IEEE Transactions on Medical Imaging*, vol.37, pp. 1478-1487, 2018.

[6] Shen D, Wang L, Guest Editorial Special Issue on Deep Learning in Medical Imaging, *IEEE Transactions on Biomedical Engineering*, vol.65, pp. 1898-1899, 2018.

[7] Zhou F, Bai X, High-Resolution Sparse Subband Imaging Based on Bayesian Learning With Hierarchical Priors, *IEEE Transactions on Geoscience and Remote Sensing*, vol.56, pp. 4568-4580, 2018.

[8] Shen C, Gonzalez Y, Chen L, et al, Intelligent Parameter Tuning in Optimization-Based Iterative CT Reconstruction via Deep Reinforcement Learning, *IEEE Transactions on Medical Imaging*, vol.37, pp. 1430, 2018.

[9] Dawar N, Ostadabbas S, Kehtarnavaz N, Data Augmentation in Deep Learning-Based Fusion of Depth and Inertial Sensing for Action Recognition, *IEEE Sensors Letters*, vol.3, pp. 1-4, 2019.

[10] Han L, Yang G, Dai H, et al, Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data, *Plant Methods*, vol.15, pp. 10, 2019.