SUGISAKA MASANORI

Research Article

Image Super-Resolution Reconstruction Technology Based on Deep Learning

Lingran An[1], Fengzhi Dai[2], Linghe An[3]

[1]*4th Department of Basic, North China Computing Technology Institute, Beijing, 100083, China*
[2]*College of Electronic Information and Automation, Tianjin University of Science and Technology, China*
[3]*Changchun University of Science and Technology, China*

ARTICLE INFO

ABSTRACT

The traditional super-resolution method has limited ability of feature extraction and feature expression, which cannot meet the requirements of high quality image in practical application. This paper mainly applies the relevant theories of deep learning to image super-resolution reconstruction technology. By comparing three classical network models used for image super-resolution (SR), finally a generative adversarial network (GAN) is selected to implement image super-resolution, which is called SRGAN. SRGAN consists of a generator and a discriminator that uses both perceived loss and counter loss to enhance the realism of the output image in detail. Compared with other algorithms, although the improvement of PSNR and SSIM values of the SGRAN network obtained by the final training is not obvious, the output high-resolution images are the best in the subjective feelings of human eyes, and the reconstruction effect in the image details is far higher than that of other networks.

1. Introduction

Image resolution describes the number of pixels in an image, which is also a measure of the amount of information in the image. It is an important indicator of image detail information presentation ability. In a large number of electronic image applications, high-resolution images are often expected. However, in practice, constrained by many factors, we usually cannot directly obtain the ideal high-resolution image with edge sharpening and non-block blur.

The most direct measure to improve the image resolution is to improve the optical hardware equipment in the imaging system, but this method is expensive and technically difficult Therefore, it is very important to improve the image resolution from the aspect of software and algorithm. In this context, super-resolution reconstruction (SR) technology emerged.

Although the traditional super-resolution technology can obtain high-resolution images to a certain extent, the reconstruction effect is not ideal, and there is still a great room for improvement. With the rise of deep learning, more and more visual fields are trying to apply it. In 2014, Dong et al. used convolutional neural network to achieve image super-resolution reconstruction (SRCNN) [1]. Subsequently, Dong et al. proposed FSRCNN on the basis of SRCNN [2]. This model is the acceleration of SRCNN, achieving a breakthrough in both the speed and quality of reconstruction. Kim et al. combined the idea of residual network and the structure of convolutional neural network to propose the VDSR algorithm, which can deepen the network and shorten the training time at the same time [3]. Later, Kim et al. proposed the DRCN algorithm to apply the recursive neural network structure to the super-resolution processing [4]. Shi et al. proposed a real-time super-resolution reconstruction method based on

Corresponding author's E-mail: daifz@tust.edu.cn, URL: www.tust.edu.cn

convolutional neural network and named the model ESPCN [5]. In 2016, Ledig et al. used the generative adversarial network (GAN) for the super-resolution reconstruction problem (SRGAN), and used the perceptual loss and adversarial loss to improve the authenticity of recovered images [6].

2. Image super-resolution reconstruction

2.1. *Principle*

Image super-resolution technology is to transform low-resolution data into high-resolution data through a certain method on the basis of unchanged image detection system, so as to obtain image observation of high definition images. The formation of low resolution image is often caused by the bad environment, which is often called image degradation process. It can be expressed as:

$$L = DBMH + n \qquad (1)$$

Where H and L represent high and low resolution images, M is the matrix after displacement, B is the fuzzy matrix after degradation, D is the matrix for down-sampling, and n is the additional noise pollution.

The process of low-resolution image imaging is a forward process, while the reverse process is the process of image reconstruction. The lost information can be recovered according to the imaging principle to obtain high-quality images.

2.2. *Classification*

The traditional image super-resolution methods can be divided into three categories: super-resolution technology based on interpolation, reconstruction and learning.

● Interpolation

The algorithm based on interpolation uses the gray value of adjacent pixels to generate the gray value of the pixels to be interpolated so as to realize the super-resolution reconstruction of the image. The classical interpolation methods include neighborhood interpolation, bilinear interpolation and bicubic interpolation.

● Reconstruction

The method based on reconstruction can realize the estimation and reconstruction of high-resolution images by establishing the imaging model of low-resolution images and constructing the prior constraint of high-resolution images. The main research methods include iterative back-projection (IBP), projection onto convex set (POCS) and maximum a posteriori (MAP).

● Learning

The learning-based algorithm can train the images with high and low resolution and master the relationship between them, so as to establish the mapping model. It mainly includes neighborhood embedding (NE), sparse representation (SR), anchored neighborhood regression (ANR), etc. All these methods belong to the field of machine learning.

3. Super-resolution based on deep learning

3.1. *Deep learning theory*

In 2006, Hinton et al. first defined the concept of deep learning [7]. It is clearly expressed in the concept that the artificial neural network model can extract and learn the features of the original data through multi-layer nonlinear transformation operation, so as to carry out reasonable and effective prediction of the data information, and make the training results constantly approach the expected value of the target through layer by layer learning.

With the advancement of research, a variety of powerful deep learning models have emerged, among which the typical ones are convolutional neural network, recurrent neural network, deep belief network and generative adversarial network. Among them, the convolutional neural network was the first to be introduced into the super-resolution technology and has achieved great success. In recent years, the super-resolution method based on the generative adversarial network structure has made the reconstructed images improve in terms of details.

The convolutional neural network consists of multiple layers stacked one on top of the other. It takes the raw data that's coming in, and it extracts the high-level information from it, and abstracts it, and that's the feedforward operation of convolution. The error obtained by comparison is fed back to the front layer continuously until the error is reduced to the minimum, so that the model converges to achieve the purpose of training. Its network structure generally includes input layer, convolution layer, pooling layer, full connection layer.

The structure of generative adversarial network is composed of a generator and a discriminator. The generator is used to synthesize the network data, and the discriminator is used to judge whether the network data generated by the experiment is feasible and effective to approach the real value.

3.2. *Three super-resolution models*

In this section, three deep neural network models are introduced for image super-resolution reconstruction, namely SRCNN, ESPCN and SRGAN.

SRCNN uses only three layers of network to achieve image super resolution. It uses mean square error (MSE) as the loss function, which is beneficial to obtain a higher PSNR. The network structure of SRCNN is shown in Fig.1.
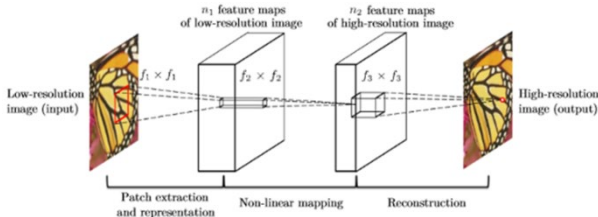

Fig.1. The network structure of SRCNN

ESPCN is an efficient method to extract features directly from low-resolution image sizes and calculate high-resolution images. The core concept of the network is the sub-pixel convolutional layer, which greatly reduces the computation, saves time and improves the speed of the experimental process. The activation function is tanh function, and the loss function is MSE. The network structure of ESPCN is shown in Fig.2.
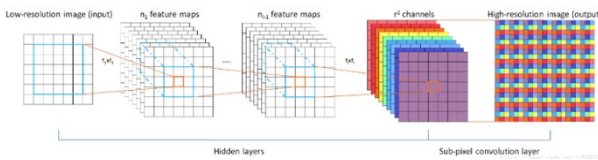

Fig.2. The network structure of ESPCN

SRGAN applies the generative adversarial network to solve the super-resolution problem. The network is composed of a generator network and a discriminator network, which can improve the sense of reality of the generated image by both the loss of perception and the loss of resistance. The network structure of SRGAN is shown in Fig.3.
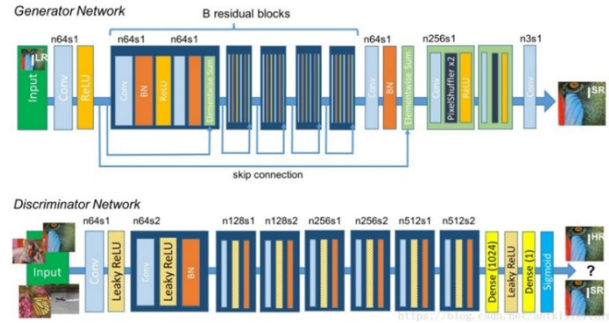

Fig.3. The network structure of SRGAN

3.3. *Model selection*

Among the above four models, the network model structure of SRCNN is simple and the information obtained is very limited. ESPCN can only deal with images with a small magnification degree, and when the image is enlarged to a greater extent, the resulting image will be blurred. SRGAN can still generate the details in the image in the case of 4 times or more magnification, which can be close to the real effect visually. Therefore, we chose SRGAN as the base network.

4. Model training and testing

4.1. *Experimental environment*

Experiments in this paper are carried out on Tensorflow, which is an open source deep learning framework, and all algorithms are implemented by Python. The environment configuration used in the network model training and testing experiments is shown in Table 1.

Table 1 Experimental environment

Configuration	Attribute
System	16.04-Ubuntu
CPU	Intel Xeon E5-2640 v4
Memory	32GB DDR4
GPU	NVIDIA TITAN Xp

4.2. *The data set*

3000 high-quality images are selected from ImageNet and rotated by 0°, 90°, 180° and 270°, respectively, and then flipped vertically. In this way, the training set is expanded eight times to 24,000 images. The interpolation function is used to get the corresponding low-resolution image data set by downsampling all the high-resolution images four

times. Set5, Set14 and BSD100 data sets commonly used in the field of super-resolution are adopted as test sets.

4.3 *Evaluation index of experimental results*

In evaluating the effect of image super-resolution, the evaluation criteria mainly lie in the gap in data between the reconstructed image and the expected real image. Common indicators include Structural Similarity (SSIM) and Peak Signal to Noise Ratio (PSNR). The value range of SSIM is [0,1], and the closer it is to 1, the better. PSNR is in dB, and the bigger the value, the better.

4.4 *Network training parameter setting*

The adaptive moment estimation (Adam) method is selected as the optimization strategy for network training to update the learning rate adaptively. During the experiment, the learning rate is initialized to 0.001. At the beginning of the training model, the epoch is preset as 300, and then the optimal number of iterations is selected according to the trend of loss and result curve during the training.

4.5 *Experimental results and analysis*

We tested the trained SRGAN network model on the Set5, Set14 and BSD100 test sets, and compared the performance of SRGAN with Bicubic, SRCNN and ESPCN, and summarized the results in Table 2, where the bold part is the maximum value.

Table 2 The performance comparison

		Bicubic	*SRCNN*	*ESPCN*	*SRGAN*
Set5	*PSNR*	29.26	30.07	30.76	**30.95**
	SSIM	0.8594	0.8627	0.8784	**0.8836**
Set14	*PSNR*	26.74	27.18	27.66	**27.78**
	SSIM	0.7648	0.8627	0.8784	**0.8792**
BSD100	*PSNR*	26.13	26.68	27.02	**27.12**
	SSIM	0.7045	0.7219	0.7442	**0.7449**

As can be seen from the above table, the PSNR and SSIM values of SRGAN are slightly higher than those of other algorithms. Although the improvement of objective indicators is not obvious, but the high-resolution images generated by SRGAN are the most realistic compared with other methods. A detailed comparison of a set of output images is shown in Fig.4.
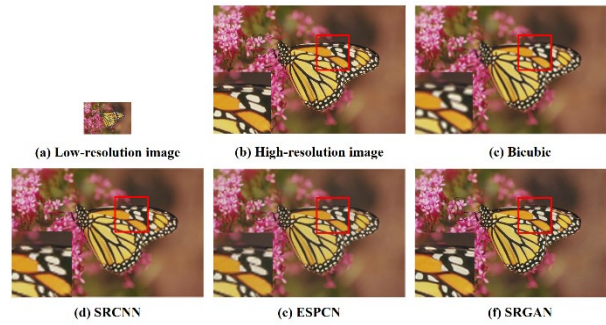

Fig.4. Comparison of reconstruction results

5 Conclusion

In this paper, an image super-resolution reconstruction network based on GAN model is trained, and the trained network is tested on the open data set and compared with three classical super-resolution algorithms. It can be seen from the experimental results that although the PSNR and SSIM values of this network are not improved very much, the reconstructed image is closer to the real image.

References

1. Dong C, Loy C C, He K M, et al, Learning a Deep Convolutional Network for Image Super-Resolution, *European Conference on Computer Vision*, 2014: pp 184-199.
2. Dong C, Loy C C, Tang X O, et al, Accelerating the Super-Resolution Convolutional Neural Network, *European Conference on Computer Vision*, 2016: pp 391-407.
3. Kim J，Lee J K，Lee K M, Accurate Image Super-Resolution Using Very Deep Convolutional Networks, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: pp 1646-1654.
4. Kim J, Lee J K, Lee K M, Deeply-Recursive Convolutional Network for Image Super-Resolution, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: pp 1637-1645.
5. Shi W Z, Caballero J, Huszar F, et al, Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: pp 1874-1883.

6. Ledig C, Theis L, Huszar F, et al, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: pp:105-114.
7. Hinton G E, Osindero S, Teh Y W, A fast learning algorithm for deep belief nets, *Neural Computation*, 2006, 18(7): pp 1527-1554.

Authors Introduction

Ms. Lingran An
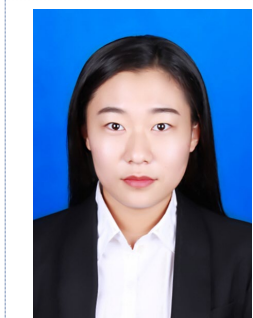
She received her Master of Engineering degree from Tianjin University of Science and Technology, China in 2020. Her main research areas are deep learning, image processing, artificial intelligence and robotics. During her study, she has published many conference papers. She started to work in North China Computing Technology Institute, Beijing in July 2020, where she is currently an assistant software test engineer of the 4th Department of Basic of the Institute.

Dr. Fengzhi Dai

He received an M.E. and Doctor of Engineering (PhD) from the Beijing Institute of Technology, China in 1998 and Oita University, Japan in 2004 respectively. His main research interests are artificial intelligence, pattern recognition and robotics. He worked in National Institute of Technology, Matsue College, Japan from 2003 to 2009. Since October 2009, he has been the staff in Tianjin University of Science and Technology, China, where he is currently an associate Professor of the College of Electronic Information and Automation.

Mr. Linghe An

He is a sophomore in the College of Science, Changchun University of Science and Technology, China, majoring in information and computing science. His major subject areas are mathematics and computer science and technology.

167