Research on Image Super-Resolution Reconstruction Based on Deep Learning

Lingran An [1,3], Fengzhi Dai [1,2,3]*, Yasheng Yuan [1]

[1] *College of Electronic Information and Automation,*
Tianjin University of Science and Technology, China;

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China;*

[3] *Advanced Structural Integrity International Joint Research Centre,*
Tianjin University of Science and Technology, Tianjin, 300222, China
*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

This paper mainly applies the relevant theories of deep learning to image super-resolution reconstruction technology. By comparing four 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. The data sets used by the training network are partly from the network and partly from the artificial. Compared with other network models, the final trained SRGAN network is above average in PSNR and SSIM values. Although it is not optimal, 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.

Keywords: Super-resolution, deep learning, neural network, Generative Adversarial 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.

The idea of image super-resolution was first proposed by Harris[1] and Goodman[2] in the 1960s, who adopted interpolation to improve the spatial resolution of a single image. 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)[3]. Subsequently, Dong et al. proposed FSRCNN on the basis of SRCNN[4]. This model is the acceleration of SRCNN, achieving a breakthrough in both the speed and quality of reconstruction. Kim et al. proposed the DRCN algorithm to apply the recursive neural network structure to the super-resolution processing[5]. Wenzhe et al. proposed a real-time super-resolution reconstruction method based on convolutional neural network and named the model

ESPCN[6]. 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[7].

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:

$$L = DBMH + n \tag{1}$$

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

A simple neural network model is shown in Fig.1. The input layer is on the left, the hidden layer is in the middle, and the output layer is on the right.

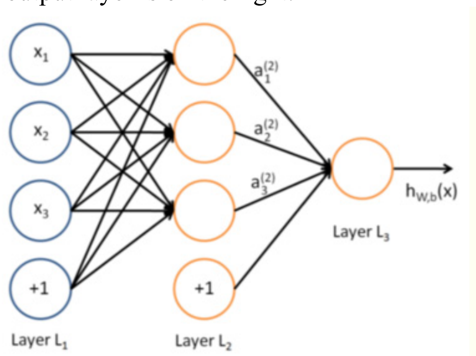

Fig.1. Simple neural network model schematic diagram

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. Four super-resolution models

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

- SRCNN

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.2.

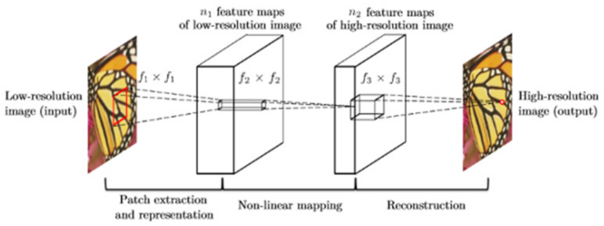

Fig.2. The network structure of SRCNN

- ESPCN

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.3.

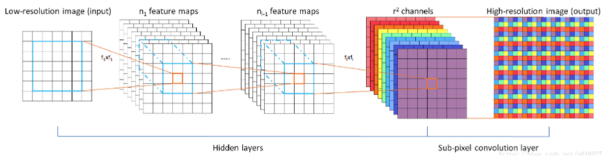

Fig.3. The network structure of ESPCN

- DRCN

DRCN applies recursive neural network in image super-resolution processing, and at the same time uses the idea of residual learning (Skip-Connection) to deepen the network structure (16 recursion), increase the network receptive field, and improve the performance. The network structure of DRCN is shown in Fig.4.

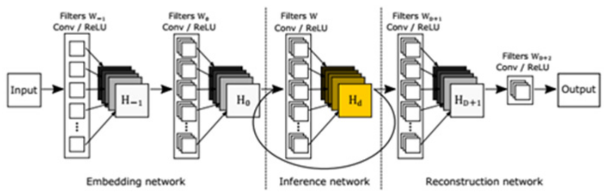

Fig.4. The network structure of DRCN

- SRGAN

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.5.

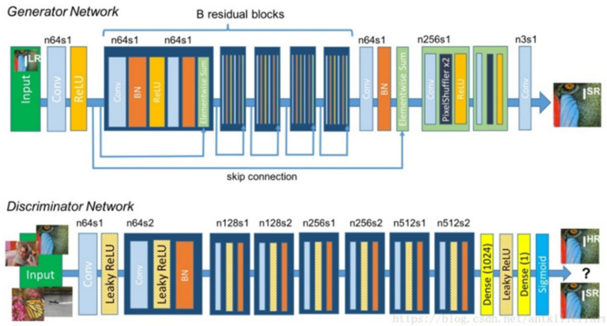

Fig.5. The network structure of SRGAN

3.3. Model selection

Among the above four models, the network model structure of SRCNN is simple and the information obtained is very limited. DRCN network has complex structure and large computation. ESPCN can only deal with images with a small magnification degree. SRGAN can still generate the details in the image in the case of 4 times or more magnification. Therefore, we chose SRGAN as the base network.

4. Model training and testing

4.1. Experimental environment

In this paper, the image super-resolution reconstruction network model based on the generative adversarial network is implemented on the TensorFlow deep learning framework, which provides a Python programming interface to conveniently implement the method proposed in this paper. The environment used in this paper is configured as Ubuntu 16.04 operating system, 32GB DDR4 memory, graphics card GPU model TITAN XP, CPU model e5-2640v4.

4.2. The data set

The data set used by the training network consists of two parts: one part is 8156 high-resolution images from the RAISE data set, and the other part is artificially collected, 844 high-resolution images are randomly shot by camera. A total of 9000 images were collected from the two data sets to form the high-resolution training set, and then the corresponding 9000 low-resolution training sets were obtained by sampling down 4 times. The data sets used by the test trained network are Set5, Set14, and BSD100 test sets.

4.3. Model training process

The training process of SRGAN model is as follows:

1. Train the SRResnet with 1000000 iterations.
2. Train the SRGAN with the weights from the generator of SRResnet for 500000 iterations using the MSE loss.
3. Train the SRGAN with the weights from the generator and discriminator of SRGAN (MSE loss) for 200000 iterations using the VGG54 perceptual loss.

4.4. 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.

4.5. 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 SRCNN, ESPCN and DRCN, and summarized the results in Table 1. Where the bold part is the maximum value.

Table 1 The performance comparison

		SRCNN	*ESPCN*	*DRCN*	*SRGAN*
Set5	*PSNR*	30.07	30.76	**31.52**	30.95
	SSIM	0.8627	0.8784	**0.8938**	0.8836
Set14	*PSNR*	27.18	27.66	**28.02**	27.78
	SSIM	0.8627	0.8784	**0.8938**	0.8792
BSD100	*PSNR*	26.68	27.02	**27.21**	27.12
	SSIM	0.7219	0.7442	**0.7493**	0.7449

As can be seen from the above table, SRGAN's PSNR and SSIM values are only in the upper middle level, 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.6.

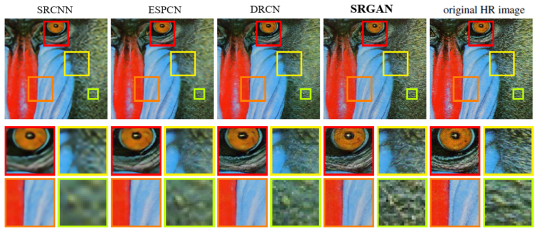

Fig.6. Output picture detail comparison diagram

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 SR networks. It can be seen from the experimental results that although the PSNR and SSIM values of this network are not the highest, the reconstructed image is closer to the real image. It's a more subjective visual network.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700), and the Industry-University Cooperation and Education Project (201802286009).

References

1. Harris J L, Diffraction and resolving power, *JOSA*, 1964, 54(7): pp. 931-933.
2. Goodman J W, Cox M E, Introduction to Fourier Optics, *Physics Today*, 1969, 22: pp 97.
3. Dong C, Loy C C, He K, et al, Learning a deep convolutional network for image super-resolution, *Proc of European Conference on Computer Vision*, 2014: pp 184-199.
4. Dong C, Loy C C, He K, et al, Image super-resolution using deep convolutional networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): pp 295-307.
5. Wenzhe Shi, Jose Caballero, et al, Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, arXiv preprint arXiv:1609.05158, 2016.
6. Kim J, Lee J K, Lee K M, Deeply-Recursive Convolutional Network for Image Super-Resolution, arXiv preprint arXiv:1511.04491, 2016.
7. Christian Ledig, Lucas Theis, et al, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, arXiv preprint arXiv:1609.04802, 2016.