

Image Super-Resolution Algorithm Based on Deep Residual Network

Yang Liu

School of Computer Science, Guangdong University of Technology, China.

906328986@qq.com

Abstract. This paper puts forward to the image super-resolution algorithm based on the resent. The algorithm makes use of the depth residual network to reconstruct the low rate image, and the super-resolution images obtained by the cycle are, judged by the convolution neural network. The residual learning is introduced to reduce the complexity and degradation of deep network.

Key words: super-resolution; resent; convolution neural network.

INTRODUCTION

In paper [1], since the Alex Net convolutional neural network emerged in 2012, the Convolutional Neural Network (CNN) has an indispensable position in the current deep learning, and the network avoids the complex pre-processing of the image can directly input the original image. Only need to provide a large amount of data and set parameters, and then through training can achieve good results. With the rapid advancement of computer hardware and the promotion of deep learning, significant progress has been made in the field of imaging. In terms of image reconstruction and restoration, the literature [2] proposes the Super-Resolution Using Deep Convolutional Networks (SRCNN) based on convolutional neural network (CNN). The network consists of three convolutional layers, using 91-image dataset. Dataset to train the network, cut the image and then down sample 3 times into a low-resolution image as input, the original image as an output, learn the features to perform image super-resolution reconstruction according to the features

RELATED WORK

The possibility of introducing deep residual networks: The deep network validated in [3] can learn more features, and low-level features are extracted one-by-one to learn more advanced features like image classification [4] and target detection and semantic segmentation [5] benefit from deeper neural network models. Although deep-layer networks can converge, network degradation problems occur. That is, as the network deepens, the accuracy rate will drop. In 2015, Kaiming He et al. proposed residual network Residual Networks to effectively solve the problem of network degradation [6], and captured the first place in ImageNet's classification, detection, localization, and COCO's detection and segmentation.

In computer vision, the "level" of features increases as the depth of the network deepens. Studies have shown that network depth is an important factor in achieving good results. The more layers (mainly hidden layers) in the neural network, the higher the level of abstraction for the input features, and at least not lower than the shallow ones. This is because in the neural network, the input of the next layer of neurons is the weighted sum of the output of the previous layer, and the characteristics of the previous layer are abstracted in the latter layer. In fact, the learning process is to adjust and optimize the weight of each connection. And thresholding process. The level of abstraction that shallow neural networks can represent is not high, and the deeper the level, the higher the level of abstraction of features, that

is, the “better effect” on certain tasks, which is why deep neural networks can make many things that only humans can do need a high level of abstract understanding.

The facts also prove that the network depth increases and the network expresses more. Classic networks such as AlexNet, VGG-16, and VGG-19 increase the depth and have a significant effect on image classification and semantic segmentation. But when the network becomes deeper and deeper, there will be a phenomenon that the accuracy of the training set is reduced. This is not caused by overfitting, because the training set will perform well if there is an overfitting problem. Therefore, it can be determined that the deeper the network is, the more difficult it is to train, which will lead to a decrease in performance and increase the training time.

ALGORITHM SRWCGAN

This paper image super-resolution algorithm SRWCGAN network frame body using conditional generation against network. In order to improve the quality of image reconstruction and deepen the generator network layer, the network is too deep to be degraded. For this reason, we introduce the residual learning, namely, the deep residual network to solve the problem of network degradation.

Resnet

In order to solve the problem of network degradation, the algorithm uses a deep residual network structure, as shown in the figure. Residual learning considers that it is more difficult for the network to learn a mapping with $F(x)=x$ than learning $F(x)=0$, so we can design the network structure as $H(x) = F(x) + x$. This completes the learning of identity mapping and reduces the learning difficulty. Here x is the input of the residual structure, F is the mapping of the network learning in this layer, and H is the output of the entire residual structure. In contrast, learning residuals using the deep network is simpler, and for the same depth of the network, the residuals converge. Faster and easier to optimize. Therefore, the residual learning is introduced into the image super-resolution reconstruction problem to ensure that the deep network will not be degraded and learn more detailed features.

Network Structure

The network structure of the algorithm in this paper is shown in the figure.1. It uses conditional generation against the neural network structure, including generators and discriminators.

The generator network is a deep residual network. The figure shows a residual operation for every three convolutional layers. The generator uses an end-to-end mapping model and does not require data preprocessing. Low-resolution images are mapped to high-resolution images through a deep residual network, normalized Batch Norm [20] is added after each convolution operation to solve the deep network gradient disappearance/explosion problem, and the PReLU[21] activation function is used to pass the output. $\times 1$ The convolutional layer adjusts the output dimension to make it the same as the residual input dimension. The size of the image output from each convolutional layer is the same (3x3 size convolution kernel, "padding" is set to 1) to learn more features and maintain the integrity of the feature information. The dropout [22] is added to the final conv of the generator network. Dropout refers to that the weights of some hidden layer nodes in the network are not allowed to work randomly during the training of the model. Those nodes that are not working can temporarily assume that they are not network structures. Part of it, but its weight has been retained (just temporarily not updated). Due to the complexity of the network and the limited amount of training data, adding dropout can effectively prevent over-fitting. The traditional cGAN generator input is noise and constraint y . The generator input in this paper replaces the constraint y and replaces it with a low-resolution map. The low-resolution map is generated by the bicubic interpolation of the original image. In order to compare with the existing image super-resolution algorithm, the Y-channel (YCrCb color space, Y represents the luminance channel) of the image is trained and tested.

The discriminator network is a five-layer convolutional layer, and the use of fewer convolutional layers is to increase efficiency and reduce training time. As shown in the figure, the input is a "real" high-resolution image and the resulting super-resolution image is output as a probability.

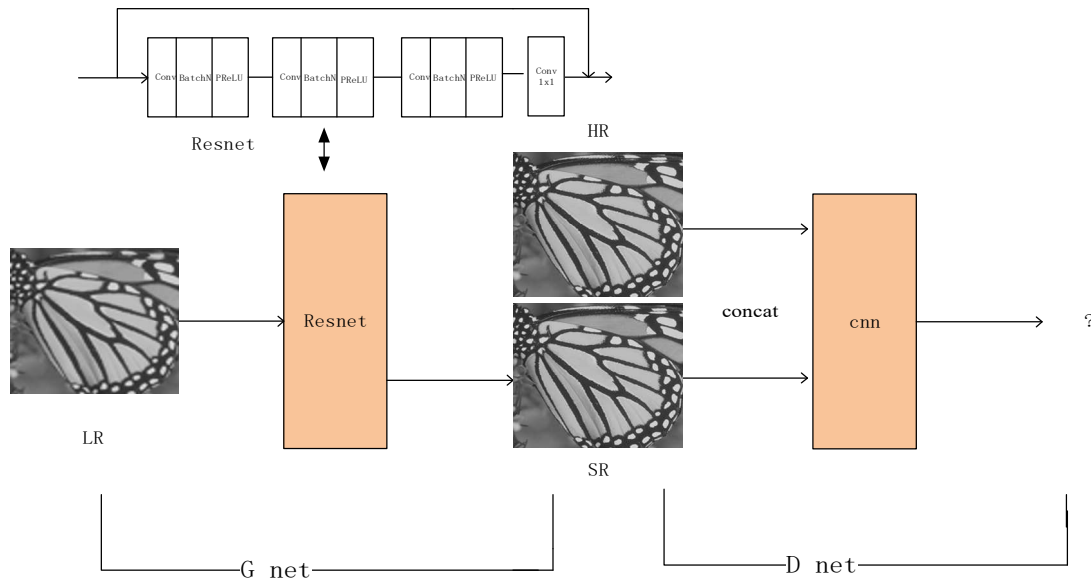


FIGURE.1. SRWCGAN Network structure

CONCLUSION

In order to verify whether the method proposed in this chapter is effective in reconstructing images, this section provides experimental tests. The test images for the aforementioned test data set are Set5 (5) + Set14 (14 sheets) + BSD200 (100 sheets), A total of 119 images. Among them, the objective evaluation index PSNR and SSIM test are single-channel grayscale images. At the same time, the three-channel color map is tested and the reconstruction effect of the algorithm is observed by human eyes. The experimental platform is TensorFlow 1.4, the experimental environment is Ubuntu 16.04 operating system, and the GPU is configured as NVIDIA GeForce GTX1080. The experiment in this chapter includes testing the effects of three different depth generators and observing the generator loss changes during training.

The test results after magnified the image by three times are shown in Table 1. In Table 1, the depth residual network of the algorithm generator in this chapter uses 50 layers, 101 layers and 152 layers respectively, and existing ones include Bicubic, A+[27], and SRCNN[8], VDSR [28] four super-resolution algorithm for comparison, the test set for Set5, Set14, BSD200. The PSNR and SSIM in this chapter's algorithm are higher than those in the other four algorithms, and although the training speed of this chapter decreases with the increase of the generator network layer, both PSNR and SSIM increase with the network depth.

As shown in Figure 2, the loss values of three types of network training changes, the experimental results of this paper increase with the increase of the number of network layers, and the generation of confrontation networks will cause the loss value to be unstable.

TABLE 1. Mean PSNR (dB), a verage SSIM, and average runtime (S) test results

Datasets	Scale	SRWCGAN50			SRWCGAN101			SRWCGAN152		
		PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME
Set5	×3	33.89	0.921	0.31	34.16	0.930	0.37	34.22	0.935	0.53
Set14	×3	29.91	0.815	0.33	30.29	0.829	0.39	30.46	0.853	0.52
BSD200	×3	29.58	0.798	0.39	29.87	0.810	0.50	30.01	0.823	0.65

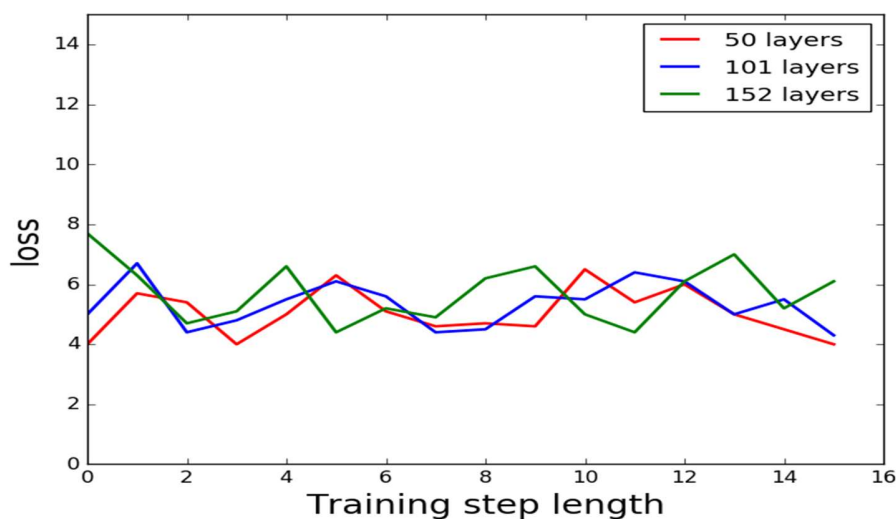


FIGURE.2. Loss value

ACKNOWLEDGMENTS

This work is supported by funds from Department of Science and Technology at Guangdong Province and GuangzhouCity(No.201604010051,2015B090901060,2016B090903001,2016B090904001,2016B090918126,2016KZ010101).

REFERENCES

1. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
2. Dong C, Loy C C, He K, et al. Image Super-Resolution Using Deep Convolutional Networks. [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 38(2):295-307.
3. Du B, Wang Z, Zhang L, et al. Robust and Discriminative Labeling for Multi-Label Active Learning Based on Maximum Correntropy Criterion[J]. IEEE Transactions on Image Processing, 2017, 26(4): 1694-1707.
4. Li J, Mei X, Prokhorov D, et al. Deep neural network for structural prediction and lane detection in traffic scene[J]. IEEE transactions on neural networks and learning systems, 2017, 28(3): 690-703.
5. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3431-3440.
6. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.