A Review of Image Compressed Sensing in Deep Learning

Kaiguo Xia*, Lei Hu and Pengqiang Mao
Army Engineering University of PLA, Nanjing, China

*Corresponding author. Email: 857257221@qq.com

Abstract. In recent years, deep learning has developed rapidly in the field of image recognition, which provides a new idea for the reconstruction of compressed sensing. The new method based on deep learning can measures the correlation between the measurement signal and the original signal through network, which not only has high reconstruction accuracy, but also significantly reduces the time consuming, showing the great potential of deep learning in the field of compressed sensing reconstruction. This paper sorts out the current image compressed sensing reconstruction methods based on deep learning, analyzes the characteristics and key steps of the algorithm according to three different deep network structures, and finally looks forward to the development direction of compressed sensing reconstruction based on deep learning.

Keywords: compressed sensing, deep learning, network structure.

1. Introduction

In 2006, David L. Donoho proposed the compressed sensing theory [1], it used sparsity of signal to break through the Shannon-Nyquist specific restrictions on sampling rate, the basic conclusion that [2] : when measuring satisfy certain conditions, can use a sparse recovery algorithm to reconstruct compressed signals accurately from its high undersampling (that is far lower than the Nyquist sampling rate) in linear measurement sample. As for the image signal, compared with the traditional compression technology, which follows the process of sampling, compression and decompression, the compressed sensing is far lower than the sampling frequency twice the maximum frequency of the signal. This greatly reduces the resource consumption of the sampling and compression process. However, there are two main problems in traditional compressed sensing reconstruction methods :(1) real signals such as natural images do not exactly satisfy sparsity in the transformation domain, so it may lead to the reconstruction accuracy decreases when using reconstruction algorithm in real signals .(2) due to the reconstruction algorithm using multiple iterations to solve the original signal, the computational complexity is high and it is difficult to achieve real-time performance, which limits the application breadth and depth of compressed sensing technology.

In recent years, deep learning has been widely studied and applied. Many scholars have also tried to solve the problem of compressed sensing with deep learning. Deep learning adopts a data-driven approach to learn the feature of the signal and the specific structure of the actual signal by adaptively adjusting the network weight. With the support of parallel GPU computing hardware, the compressed sensing reconstruction can realize real-time processing, reducing process of the multiple iterations of traditional algorithms.
2. Image reconstruction of deep learning method

2.1. Compressed sensing reconstruction based on multi-layer fully connected neural network

The great success of deep learning in image recognition and speech recognition makes researchers focus on how to use deep learning model to realize compressed sensing reconstruction algorithm. In 2015, Ali Mousavi et al. [3] realized the reconstruction of compressed sensing for the first time by Stacked Denoising Autoencoder (SDA), which can learn the mapping relationship between the measurement and the original signal. Compared with the traditional iterative reconstruction method D-AMP [4], this method not only basically as good as the reconstruction accuracy of the traditional method, but also performed better in higher compression ratios. Because of fully connected neural network, the speed of reconstruction is more than 10,000 times faster than the traditional method.

This method adopts two different signal measurement methods, the traditional linear measurement and the network-based nonlinear measurement. In figure 1(a), \( \Phi \) is the measurement matrix, \( x \) is the input signal, \( y \) is the measurement. The coding process can be described as \( y = \Phi x \), Figure 1.linear and non-linear measurement.

![Figure 1. Linear and non-linear measurement.](image)

The signal measurement is linear. Then, the three-layer SDA network is used to output the reconstructed signal. In figure 1(b), the measurement matrix is replaced by this nonlinear process, which can be a sigmoid function or other nonlinear functions. The last three layers of the network are the same as figure 1(a).

Using SDA for reconstruction shows obvious advantages in computing time, but there are still some problems. Since it is a fully connected network, when the input picture dimension grows, the network parameters will also increase rapidly and the network complexity will increase synchronously. At the same time, too many redundant connections will easily lead to overfitting. In consideration of this problem, they also adopted the image block, but the reconstructed picture produced block effect, thus reducing the reconstruction accuracy. Moreover, the network is relatively simple, and the learned signal distribution characteristics are relatively simple, which also affects the reconstruction effect.

2.2 Compressed sensing reconstruction based on Convolutional Neural Network

In the past few years, convolutional neural network has shown remarkable performance over traditional methods in many fields of computer vision and image processing, such as image segmentation [5], object detection [6], object classification and recognition [7], etc. However, the problems solved in these fields are different from those of compressed sensing, the former is the inference problem, the latter is reconstruction problem.

The general process of convolutional neural network (CNN) is to input high-dimensional images, go through several convolution layers and pooling layers, and then output low-dimensional vectors.
The network removes unnecessary details from the original signal, and refines the key features of the target, so as to facilitate machine recognition and classification. The purpose of the reconstruction algorithm is to restore the original signal as much as possible and obtain good signal reconstruction accuracy. It should not only extract the key features of the target signal, but also retain as much signal details as possible in the process. Therefore, it is necessary to improve the classic CNN network to meet the requirements of compressed sensing reconstruction.

K.ulkarni et al. [8] proposed Recon-Net network to solve the reconstruction of compressed sensing by CNN network for the first time. The network structure is shown in figure 3. In order to reduce the complexity of the network, the image is still divided into small blocks, and the random Gaussian measurement matrix is used for the measurement matrix. The network removes the pooling layer of traditional CNN and only retains the Convolution and ReLU layer, which reduces the information loss caused by pooling layer. Meanwhile, it uses multi-layer convolutional layer to change the dimension of the feature to realize the reconstruction process of measurement signal. Compared with the SDA method, this network reduces the complexity of the network, retains a lot of details, and achieves better reconstruction accuracy at a lower compression ratio.

Yao et al.[9] improved on recon-net, proposed that Dr2-Net first used the linear mapping of the full connection layer to generate the reconstruction image. They absorbed the advantages of Resnet [10], and introduce multi-layer fusion of features by using the residual layer to further improve the reconstruction accuracy. Due to the addition of residual layers in Dr2-Net, the computational complexity was greatly increased. On this basis, densenet block was added in [11], and a DC-Net with double channels was proposed, which not only retained the high precision reconstruction brought by resnet block, but also accelerated the network computing speed by densenet block, and further improved the reconstruction accuracy, visual quality and robustness.

2.2. Compressed sensing reconstruction based on Generative Adversarial Network

Generative Adversarial Networks[12] was firstly proposed by Ian Goodfellow in 2014, and now it is the best generative model in the deep learning area, with striking performance in image generation and migration image style, image super resolution. Currently, researchers are beginning to utilize GAN network to solve the problem of compressed sensing reconstruction.
Article [13] was the first to apply GAN network to the problem of compressed sensing reconstruction, and proposed DAGAN network structure, which aims to solve the problem of restoring MRI images in the medical field, as shown in figure 4. Compared with the traditional multi-coil rapid MRI technology, the compressed sensing MRI technology can break through the Nyquist sampling law and use less raw data to reconstruct the MRI image, thus speeding up the whole imaging process. GAN network consists of a generator G and a discriminator D, which can be used to learn the potential distribution of actual data X, generate pseudo-samples, and ultimately make the discriminator unable to distinguish the pseudo-samples from the real data. In DAGAN, generator G consists of a U-net which contain 8-layer convolution and an 8-layer deconvolution, and the discriminator is a standard CNN network with an 11-layer convolution, which directly outputs true or false. In the training, the input $x_u$, the under-sampled measurement with zero padding, feed in generator G and output the pseudo-sample $\tilde{x}_u$, then the discriminator will compare $\tilde{x}_u$ with the real value $x$ until it fails to distinguish the pseudo-sample from the real value. This method utilized GAN network to try to find the mapping relationship between under-sampled signal and fully sampled image, so as to realize the reconstruction of under-sampled signal, which is not only better than the traditional method in image details, but also has a significant reduction in processing time.

3. Conclusion

In this paper, we sort out the image compressed sensing reconstruction method in three different network structures. Compared with the traditional method, which adopts the iterative method to solve the measurement signal, the deep learning method USES a large amount of data to learn the correlation between the measurement signal and the original signal, which improves the reconstruction accuracy, greatly shortens the calculation time, and has certain real-time performance. At present, researches on deep learn-based compressed sensing reconstruction mainly focus on image reconstruction, but some researchers have extended the dimension to time domain, and compressed sensing reconstruction for video is also an important research field.

References