Image Recognition Technology of Synthetic Aperture Radar

Rongxia Duan¹, Jiaru Liu¹, Baocai Xu²,∗ and Tianchen Huang¹

¹Army engineering University of PLA Shijiazhuang campus, Shijiazhuang, China
²Hebei College of Industry and Technology, Shijiazhuang, China

*Corresponding author e-mail: Zhaoxiaoyan@ustb.edu.cn

Abstract. The rapid development of artificial intelligence in recent years, especially the rise of deep learning, provides a new path for radar target recognition. This paper first introduces the development of synthetic aperture radar and deep learning, and then analyzes the challenges and difficulties of deep neural networks. Then the self-encoder is used for target recognition. The model uses self-encoder to realize feature selection and data Manacor to classify the final object. In this paper, the advantages and disadvantages of the model are analyzed, and the improvement ideas are put forward. Finally, the original synthetic aperture radar image is processed by Median filter, mean filter and Gaussian filter. After comparison and analysis, it is concluded that the filtering effect is best when the Median filter size is 22.

Keywords: Deep Learning, Convolutional Neural Network, Radar, Image Recognition

1. Introduction

Since the 1980s and 1990s, when the U.S. and the Soviet Union launched a series of spaceborne radars, the synthetic aperture radar has become very rich in image data, and the recognition of synthetic aperture radar images has become particularly important. In recent years, researchers at home and abroad have done a lot of research on synthetic aperture radar image recognition. In 2016, a scholar named s Wang proposed a complementary new coding method for synthetic aperture radar recognition, and his pyramid coding approach yielded features that are more discriminating and robust than ever before, the method is tested on the MSTAR data set, and the feasibility of the method is verified [1]. In the same year, another researcher s Wang proposed a method of radar target recognition based on sparse feature representation, which can extract structural information from radar images, to capture the subtle differences between images [2]. In 2017, b Ding proposed a feature extraction technique for images of synthetic aperture radar objects called 2 D compression sensing [3]. The idea is to represent the original image as a sparse image, the sparse image can be reconstructed with a low-dimensional Matrix. Since the 1970s, when China developed its first synthetic aperture radar system, it has been catching up with foreign countries in synthetic aperture radar. After decades of development, the SAR image recognition technology has also made great progress. Since the 2008 Sichuan earthquake, SAR imaging systems have played a huge role in disaster relief. The satellite with high-resolution SAR system launched in China has made great contributions to the detection of natural resources and disaster early warning.

Deep learning was proposed long ago, and has made breakthroughs in recent years. In 2016, Google's AI team's Alpha Go beat out Go masters, including world champions, using techniques
including deep learning. At home, the research institutions of the state and enterprises also attach great importance to artificial intelligence. In 2013, for example, Baidu set up the first deep learning institute, becoming the First Institute in China in this field. Immediately afterwards, Tencent also established artificial intelligence research institute, begin to explore artificial intelligence foundation and application direction. In 2017, Alibaba also began research into artificial intelligence and established the ALIBABA Bardamo Institute. The State Council, China's cabinet, has also outlined a strategy to catch up with the West in artificial intelligence. At present, the biggest disadvantage of deep learning is that the training within the network is a black box for us, so its interpretability is very poor, people do not trust. Deep learning model does not have a complete theoretical foundation, and it is often used to construct a network in the way of past experience and continuous experiment, sometimes it is time-consuming and effort-consuming and can not get the desired results. And convolution networks today are a lot different in their overall structure than they were a few decades ago, so why are they so good and new. There is only one reason, and that is that the advent of big data and the development of computer technology now provide the best conditions. At the moment, the convolutional neural network still needs to adjust its parameters empirically, so the theory of deep learning has yet to be broken.

2. Deep Neural Networks
The deep neural network, which literally means 'deep network', can be understood as a perceptron with many hidden layers. Therefore, deep neural network is characterized by a large number of hidden layers. Deep learning is valued because traditional machine learning can not solve the core problems of artificial intelligence such as speech recognition, and other drawbacks that traditional machine learning can't solve, but there are many challenges and difficulties with deep learning [4-10].The first problem is the dimensionality disaster, which makes many ML algorithms very difficult to solve because of the high dimensionality of the data they are now dealing with. The dimensionality disaster can occur in statistics because the number of variables (dimensions) in the original input data is greater than the number of samples we have collected. The second problem is the modeling problem. For a task, we should establish what kind of structure model, is a simple or deep network model, many times look at the accumulation of experience. The third problem is the acquisition of training samples for a very deep neural network, we know that it takes a lot of data to train it well. The easier data are fine, but the more difficult the training samples are, the more expensive they can be. For supervised learning, sample labels have to be manually labeled. The third problem is the opacity of the deep neural network. It's like a black box, we don't know how it works, so we can't explain how it works, and it's this opacity that creates a certain distrust of neural networks. The fourth problem is the high demand for computer hardware. With the increase of the number of neural network layers and the geometric increase of parameters, the requirement of computing resources is very high, and ordinary computers can not meet the requirement. The fifth problem is poor portability. Models that perform well on one task become less effective when ported to another.

3. Convolution Network and Self-Encoder
Convolution networks, also known as convolutional neural network, are neural networks that have one or more convolution layers (using convolution operations). The network input data can be one-dimensional (one-dimensional audio signal sampled in time sequence) or two-dimensional (two-dimensional Matrix represented by a gray-scale image) [11-12]. Today, convolutional neural network excel in a variety of areas, particularly image recognition (object recognition, target location, face recognition, etc.). In short, convolutional neural network is a generic term for a series of networks in which the ordinary matrix multiplication of at least one network is replaced by convolution on the basis of the original conventional neural network. Those with more convolution layers are called deep convolution networks. As a special member of the neural network family, the job of the self-encoder is to convert the input data through the hidden layer of the network and output the same result as the input in the output layer. We call the process from the input layer to the hidden layer the encoding
process, and the process from the hidden layer to the output layer the decoding process. We can see that the network coding process is to convert the input data into another form (in general, the data dimension decreases), can be written as \( h = f(x) \); The decoding process of the network is to encode the data through the function to restore the original data, this process can be written as \( y = g(H) \). Their complete network process can be expressed as \( g(F(X)) = Y \), as shown in Figure 1.

![Figure 1. Self-coding network](image)

If we simply want the input \( x \) to be equal to the output \( Y \), then it's still like we're doing nothing, because the input is its output or its output, and it doesn't change. But we left out the hidden layer \( h \) for the intermediate transformation. If we set the dimensions of the Middle Layer \( H \) to be less than the input layer, and we know that the middle layer to the output layer is a process of restoring as much of the input data as possible, is the middle layer the most important part of the input data? So when the hidden layer unit is less than the input layer unit, the self-encoder is not equivalent to a dimensionality reduction or feature selector. For the large loss function in the process of self-encoder learning, we can give \( l(X, Y) = l(X, G(F(X))) \), \( l \) can be mean square error and so on. When the hidden layer unit is less than the input layer unit, the self-encoder is called under-complete self-encoder. However, the self-encoder has its disadvantages. When the capacity of the self-encoder is too large, the self-encoder will not learn any useful information. This limits the number of hidden layer units in an ordinary self-encoder to be less than the number of input layer units. In order to solve this problem, we design a new self-encoder, namely, regular self-encoder, including sparse self-encoder, de-noising self-encoder and so on. The key to sparse self encoders is the word sparse, and we think of the weight thinning effect of L1 regularization in regularization, which is the same here. Sparse self-encoder is a training loss function plus a regularization term.

\[
L(x, g(f(x))) + \Omega(h)
\]

Where \( h \) is the encoder output.

The purpose of sparse self-encoder is to learn the hidden layer data, that is, a series of features to achieve the final classification and other tasks. For example, we now have a 64x64 grayscale image. First we expand it to a vector of 4096 elements in one dimension and use it as input vector \( x \), then we pass through a hidden layer, which has far fewer neurons than the input layer Then, on the output layer, 4096 neurons matched the number of input layers. Next, we introduce the de-noising self-encoder. The biggest difference between the de-noising self-encoder and the previous one is that the data input by the de-noising self-encoder is not the original pollution-free data, but the noise data is added artificially, but the predicted output is uncorrupted raw data. The original training data is \( x \), then we add noise to destroy the original data as \( x' \); \( (X, x') \) as the training sample input, get the loss function

\[
L(x, x', y) = L(x, g(f(x')))
\]
Due to the addition of noise, the anti-jamming ability of the de-noising self-encoder is stronger than that of the general self-encoder. We generally use the self-encoder is only a coding layer and a decoding layer, that is, only a hidden layer, an input layer and an output layer. But there are also depth encoders and depth decoders. Depth self-encoders tend to produce data with better compression ratios. However, when the self-encoder is cascaded with other deep neural networks (such as deep neural networks), the single-layer self-encoder is sufficient, because the deep convolutional network that the deep self-encoder can do can also be completed. Self-encoder is an unsupervised deep learning method. We can use the self encoder to reduce the dimension and extract the useful features in the recognition of the synthetic aperture radar image, which has a good effect when the sample data is not abundant.

4. Image Noise Processing of Synthetic Aperture Radar
Since we are going to classify and study the synthetic aperture radar image, which is a microwave image, the final image result is equivalent to the vector superposition of the echoes reflected by the scattering points in the target's large number of scattering units. It is because of the small difference of the amplitude and phase of these scattering points that the interference spots of the final image will result in the degradation of the imaging quality. So we should reduce the speckle and noise as much as possible before recognizing the image. To reduce the effect of noise, we use Median filter, mean filter and Gaussian filter to process the original image. Median filtering is a kind of nonlinear filtering technique, which can eliminate the random noise of image effectively. The idea of median filtering is to replace the gray value of every pixel in the whole image with the median value of the gray value of the adjacent pixel. Filter size selection is important, if it is too small will make the effect of noise removal is not good, and too large will be the image. The idea of the mean filter is to replace the gray value of each pixel in the whole image with the mean value of the neighboring pixels. The size of the mean filter also has a considerable impact on the final results. Gaussian filter is a linear smoothing filter that is widely used to remove noise from images. The idea of gauss filtering is to replace the gray value of each pixel in the entire image with a weighted average of the gray value of the adjacent pixel. As shown in Figure 2 is each filter on the same image filtering results.
Figure 2. Filter image

The top left image is the original image, the top right image is the median filtered image, the bottom left image is the mean filtered image, and the bottom right image is the gauss filtered image. It can be seen that the worst effect of the mean filter, gauss filter makes the image become blurred a bit, and the median filter achieved a better effect. Next, we filter the original image using Median filter of size 2×2, 3×3, and 4×4, respectively, and analyze the results, as shown in Figure 3.

Figure 3. Median filter result
The top left image in the image above is the original image, the top right image is the result of a Median filter of size $2\times2$, and the bottom left image is the result of a Median filter of size $3\times3$, the final image is filtered through a Median filter of size $4\times4$. We can see that the effect of the filter is best when the filter size is $2\times2$, when the filter size is $3\times3$ the filter result begins to be lost, and when the filter size is $4\times4$ the filtered image is more lost than the original image. So the Median filter size we chose is a $2\times2$ matrix.

5. Conclusion
In this paper, the principle of synthetic aperture radar and the principle of deep learning are studied in detail, and the advantages and disadvantages of deep convolution network in image recognition are analyzed, the function of self-encoder in dimension reduction and feature extraction is verified. In this paper, the feasibility of the model is verified by experiments, but some improvements are also found. First is the self-encoder, the self-encoder in this paper does not use the picture structure information, so the final recognition result is not as good as the deep convolutional network model. Therefore, we can improve the self-encoder so that it can handle multi-dimensional spatial information.

References
[1] S. Wang, L. Jiao, S. Yang, et al. SAR image target recognition via complementary spatial pyramid coding [J]. Neurocomputing, 2016, 196(C):125-132
S. Song, B. Xu, J. Yang. SAR target recognition via supervised discriminative dictionary learning and sparse Representation of the SAR-HOG feature [J]. Remote Sensing, 2016, 8(8)
[2] B. Ding, G. Wen, F. Ye, et al. Feature extraction based on 2D compressive sensing for SAR automatic target recognition[C]. IEEE European Conference on Antennas and Propagation, Paris, 2017:1219-1223
[3] Z. Liao, G. Carneiro. On the importance of normalisation layers in deep learning with piecewise linear activation units [J]. 2015, 3(3):1-8.
[4] L. Wan, M. Zeiler, S. Zhang, et al. Regularization of neural networks using dropconnect [C]. International Conference on Machine Learning. 2013:1058-1066
[5] K. He, X. Zhang, S. Ren, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification [C] Proceedings of the IEEE international conference on computer vision, 2015: 1026-1034.
[6] Y. LeCun, Y. Bengio, G. Hinton. Deep learning [J]. nature, 2015, 521(7553): 436
[7] G. Hinton, L. Deng, D. Yu, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups [J]. IEEE Signal Processing Magazine, 2012, 29(6): 82-97
[8] P. Vincent, H. Larochelle, I. Lajoie, et al. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion [J]. Journal of Machine Learning Research, 2010, 11(12):3371-3408.
[9] N. Srivastava, R. Salakhutdinov. Learning representations for multimodal data with deep belief nets[C]. Proc 29th ICML, 2012, 594-605.
[10] S. Chen, H. Wang, F. Xu, et al. Target classification using the deep convolutional networks for SAR images [J]. IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(8):4806-4817.
[11] S. A. Wagner. SAR ATR by a combination of convolutional neural network and support vector machines [J]. IEEE Transactions on Aerospace & Electronic Systems. 2017, 52(6): 2861-2872.