Research on HRRP target recognition based on one-dimensional stack convolutional autoencoder

Guoling Zhang a, Xiaodan Wang*, Rui Li, Jie Lai, Qian Xiang
Cyberspace Security, Air Force Engineering University, Xian, Shanxi, 710051, China
ahejuxing@stu.xidian.edu.cn  *Corresponding author’s e-mail: afeu_wang@163.com

Abstract. Extracting key features from the radar high resolution range profile (HRRP) determines the accuracy and reliability of radar target recognition. Aiming at the problem of feature extraction and recognition in HRRP target recognition, a one-dimensional stack convolutional autoencoder (1D-sCAE) recognition method is proposed. Firstly, one-dimensional convolutional autoencoder is constructed to extract the deep features of one-dimensional HRRP signals by unsupervised learning. Then, multiple one-dimensional convolutional autoencoders are stacked to construct 1D-sCAE, and HRRP is classified by fine-tuning the network with label data. At the same time, for the overfitting problem of 1D-sCAE, dropout technology is added to optimize and improve the generalization performance. Through the simulated HRRP data of the target in the middle part of the trajectory, the experiment shows that the proposed algorithm has better feature extraction capabilities, higher recognition accuracy and stronger robustness than the typical deep neural networks.

1. Introduction
Since the high resolution range profile (HRRP) of radar contains important structural features such as target size and the distribution of scattering centers, HRRP target recognition has received extensive attention in the field of radar automatic target recognition [1]. In recent years, a variety of machine learning algorithms have been applied for the HRRP target recognition rate field, but these methods require manual feature extraction based on experience [2-4].

Deep learning algorithms can automatically extract deep features and have attracted extensive attention from researchers. As a typical supervised learning network, deep convolutional neural network (CNN) has a strong ability to extract local features, and is mostly used in image processing. Through the sharing of weights to extract common features, CNN can learn the deep structural features of the data [5]. And it’s pooling layer can extract features that have the characteristics of the translation, rotation and scaling invariance. The unsupervised learning network is also a branch of deep learning, such as stacked autoencoder (SAE) [6], generative adverse network (GAN) [7], Deep belief network [8] (DBN). These unsupervised learning algorithms can use unlabeled data to perform feature learning through data reconstruction. Due to the powerful feature extraction capabilities, deep learning algorithms are widely used in image classification [9], computer vision [10], face recognition [11] and other fields.

In recent years, some typical deep neural networks have been successfully applied for radar HRRP target recognition. Yan et al. [12] proposed an algorithm by combining sparse denoising autoencoder (SDAE) and multi-layer perception (MLP) to extract the robust features of HRRP. Zhao et al. [13] used SAE to learn the characteristics of HRRP through unsupervised training, and used ELM as a classifier to achieve better results when the number of samples was very small. Wan et al. [14]
constructed a 2D-DCNN to recognize the two-dimensional spectrum map converted from one-dimensional HRRP data, and achieved a better result. Lu et al. [15] used the bispectrum-spectral feature extraction of HRRP as input to the convolutional neural network to extract deep features. Yang et al. [16] proposed a HRRP target recognition method based on convolutional neural networks, and constructed two CNN networks. One CNN network directly extracts features from one-dimensional HRRP, and another CNN network extracts features from two-dimensional data after HRRP conversion. The results show that one-dimensional HRRP data is more conducive to extract features. Compared with fully connected networks, using convolutional neural networks for HRRP target recognition can better extract local structural features of targets.

From the recent research of HRRP target recognition, there are few studies on HRRP recognition using convolutional autoencoders. This study aims to construct a one-dimensional stack convolutional autoencoder network to extract effective deep features of HRRP and achieve high performance target recognition. In order to make full use of the internal structural features embedded in the one-dimensional HRRP data and further improve the HRRP target recognition accuracy, this paper proposes a HRRP target recognition method based on 1D-sCAE. The main contributions are as follows:

1. For the feature extraction problem, we proposed a one-dimensional convolution autoencoder to learn the structural features of HRRP. Compared with the fully connected network, the extracted features are more abstract and representative.

2. 1D-sCAE performs multi-level feature extraction by stacking one-dimensional convolutional autoencoders to enhance feature extraction capabilities and improve the recognition performance.

3. To solve the over-fitting problem of 1D-sCAE, the dropout [17] technology is used to optimize the fully connected layer and improve the generalization performance of 1D-sCAE.

2. Related works

2.1. One-dimensional convolutional neural network

Convolutional neural network is a deep neural network that imitates the design of biological visual cortex. It has the characteristics of obvious hierarchical structure, local connection and parameter sharing [18]. The one-dimensional convolutional neural network consists of a one-dimensional convolutional layer, a one-dimensional pooling layer, a fully connected layer, and a softmax classifier, the network structure is shown in Figure 1. 1D-CNN uses a one-dimensional convolution kernel as the basic calculation unit, and has strong ability to extract local features.

2.2. Autoencoder

Autoencoder (AE) is a fully connected unsupervised neural network [13]. The training process includes encoding and decoding. The network structure of AE is shown in figure 2.

The encoder map input to the latent representation space for representation learning, the decoder attempts to map the latent representation to output for reconstruction the input. Auto-encoder attempts to reconstruct the input with output as much as possible to learn underlying explanatory representation of data. Once the auto-encoder is trained, the encoder module is stacked to construct the stacked auto-encoder, the final representations provide useful information which can be used as features for building classifiers. Stacking up auto-encoder can learn more abstract and complicated representations.
3. HRRP target recognition method based on 1D-sCAE

3.1. One-dimensional stack convolutional autoencoder

According to the idea of auto-encoder unsupervised learning, we propose a one-dimensional convolutional autoencoder (1D-CAE). It extracts the deep essential features of the data through one-dimensional convolution and pooling, and reconstructs the input through one-dimensional deconvolution and upsampling operations. Its network structure is shown in Figure 3.

1D-CAE unsupervised learning includes encoding and decoding stage:

(1) encoding stage: It consists of one-dimensional convolution layer and one-dimensional pooling layer. The output of the i-th one-dimensional convolution is:

$$C_i = ReLU\left(\sum X \omega_i + b_i\right)$$

In the formula, $ReLU(x) = \max(0,x)$ activation function is a nonlinear activation function.

The one-dimensional pooling layer has the function of data dimensionality reduction. If the length is 1, the characteristic output of the i-th channel after pooling is

$$P_i(n) = \max_{0 \leq t \leq \frac{N}{S}} \{T(nW_i(n+1)W)\}$$

(2) Decoding stage: It consists of a one-dimensional deconvolution layer and an upsampling layer, which is equivalent to the inverse calculation process of encoding. The output of the channel feature after deconvolution is:

$$D_i = ReLU\left(\sum X \otimes \sigma_i + b_i\right)$$

In the formula, $\sigma_i$ is the deconvolution kernel, and $\otimes$ means deconvolution operation.

The output after upsampling is:

$$U_i = \begin{cases} 0 & k \neq j_k \quad k \in [t, 2t], t=1,2,\ldots,l \\ X_i & k = j_k \end{cases}$$

In the formula, $j_k$ represents the position of the maximum value in the one-dimensional pooling process. The feature channel is $l$.

Assuming the training set $D = \{X(i)\}_{i=1}^N$, the loss function of 1D-CAE is:

$$L(X, \hat{X}) = \frac{1}{N} \sum_{x \in X} \|X - \hat{X}\|^2$$

1D-sCAE is obtained by stacking 1D-CAE, namely, the one-dimensional pooling layer of the previous layer is used as the input of the next layer of 1D-CAE. The one-dimensional stacked convolution auto-encoder model in this article is constructed by stacking 4 1D-CAE is shown in Figure 4. Then we use Adam method optimization of 1D-sCAE network.

![Figure 3. The network structure of 1D-CAE](image)

![Figure 4. The network structure of 1D-sCAE](image)

3.2. HRRP target recognition method based on 1D-sCAE

HRRP target recognition based on 1D-sCAE mainly includes pre-processing stage, training stage and testing stage. In the pre-processing stage, the HRRP data is denoised and normalized. In the training
phase, several one-dimensional convolution auto-encoders are trained and stacked up, and then the whole network is fine-tuned using the label data. In the testing phase, the trained feature extractor and classifier are used to classify the test data to obtain the target label.

4. Analysis of experimental results

4.1. Experimental data
This paper uses FEKO and MATLAB joint simulation to get five types of target simulation data [19]. The simulation target is shown in Figure 5. After simulation, we get 18005 HRRP, each target has 3601 HRRP. Then we construct two data sets A and B by random extraction. Data set A contains 12004 training data and 6001 test sample data, and data set B contains 6001 training data and 12004 test sample data.

![Fig. 5 Five types of ballistic simulation targets](image)

4.2. Model parameter setting
The 1D-sCAE is stacked up by 4 1D-CAE, then we add a fully connected layer and a softmax classifier. The convolution size is 3, the stride is 2. The pooling size of each 1D-CAE is 2, and the stride is 2. The channel of four 1D-CAE convolution layer is 32, 48, 64, 96, The fully connected layer contains 400 hidden nodes and uses dropout technology to improve overfitting. The dropout rate is 0.2.

4.3. Model training
In this experiment, the number of epoch is set as 100. The change of HRRP recognition accuracy on dataset A with the number of epochs is shown in Figure 6. It can be found that the 1D-sCAE recognition accuracy converge quickly to more than 90% after 10 epochs, then rises steadily and slowly, and finally reaches 95.66%. Experiments prove that the 1D-sCAE algorithm has a faster convergence speed and a higher recognition rate.

The confusion matrix of HRRP target recognition is shown in Figure 7. It is not difficult to find that 1D-sCAE achieves better recognition results for five targets, and can better identify warheads and other targets.

![Figure 6. Change Curve of recognition rate](image)

![Figure 7. Recognition confusion matrix of 1D-sCAE](image)

4.4. Visualize analysis
In order to verify the ability of 1D-sCAE to extract features layer by layer, this experiment used t-SNE [20] on dataset A to visualize the original input data, the output of 3th 1D-CAE, and the output of fully connected layer. The results are shown in Figure 8.
Through comparison, it is found that the feature data of the same category after convolutional coding is similar, and the similarity becomes more obvious as the number of stacked layers increases. After the feature transformation of the fully connected layer, the separability between categories is stronger, and the similarity within the category is increased. It shows a good classification effect. Experiment shows that 1D-sCAE has a strong feature extraction capability layer by layer.

4.5. Dropout effect on recognition

In order to analyse the impact of dropout on HRRP recognition rate, the recognition accuracy of 1D-sCAE under different dropout rates was compared. When the dropout rate is set as different values, the change curve of the recognition rate on the A test set is shown in Figure 9.

It can be seen from Figure 9 that the recognition accuracy of 1D-sCAE increases firstly and then decreases with the increase of the dropout rate. When p=0.2, the recognition accuracy reaches the optimal value. It shows that proper simplification of network can improve network performance and alleviate the problem of overfitting to a certain extent. A too large dropout rate will reduce the performance of the algorithm.

4.6. Classification accuracy performance comparison

In order to verify the comprehensive performance of the 1D-sCAE algorithm for HRRP target recognition, in this section, we compare the performance of 1D-sCAE with SAE [21], SDAE [22], SpAE [23], and 1D-DCNN algorithms. The experiment was conducted on two data sets A and B with different degrees of noise. The SAE, SDAE, and SpAE is set as 4 hidden layers, and the number of nodes in each hidden layer is 400,300,200,100. 1D-CNN consists of 4 convolutional layers, 4 pooling layers and 1 fully connected layer. The experimental results are obtained by the method of five-fold cross-validation, which are shown in Figure 10.
Observation found that the recognition accuracy of HRRP recognition based on 1D-sCAE is generally higher than that of SAE, SDAE and SpAE. It shows that the feature extraction ability of convolutional neural network is stronger than that of fully connected network. Introducing the one-dimensional convolution kernel into the autoencoder can fully learn the local structural features of the HRRP, effectively enhance the network's ability to extract deep essential features and improve the target recognition rate. Comparing the recognition rate of 1D-sCAE on the noisy data set and the non-noise data set, it is found that the recognition rate of 1D-sCAE is higher obviously on the noisy data set. It shows that the 1D-sCAE algorithm has certain anti-noise ability and good robustness when processing one-dimensional HRRP data.

5. Conclusion
In order to effectively extract features and improve HRRP recognition accuracy, this paper proposes a HRRP target recognition method based on 1D-sCAE. 1D-CAE can learn the deep structure information of HRRP, and perform multi-level feature extraction by stacking 1D-CAE. At the same time, dropout technology is used to improve the network structure of the one-dimensional stack convolutional autoencoder, effectively alleviate the problem of overfitting and improve its generalization performance. According to the results, the features extracted by 1D-sCAE are more abstract and representative than those extracted by fully connected networks, and the recognition accuracy is better than that based on 1D-DCNN and SDAE.

Acknowledgments
This work was financially supported by the National Natural Science Foundation of China (61876189, 61273275,)

References
[1] LIAO, K, SI, J.X., ZHU, F.Q., et al. (2018) Radar HRRP Target Recognition Based on Concatenated Deep Neural Networks. IEEE Access, 6: 29211-29218.
[2] LEI, L, WANG, X.D., XING, Y.Q., BI, K. (2013) Multi-polarized HRRP classification by SVM and DS evidence theory. Control and Decision, 28: 861-866.
[3] XU, D.L. (2015) Study of radar automatic target recognition based on Sparse Bayesian Learning. XIDIAN UNIVERSITY.
[4] LI, R., WANG, X.D., LEI, L., et al. (2017) Ballistic Target HRRP Fusion Recognition Combining Multi-class Relevance Vector Machine and DS. Information and Control, 46: 0065-0071.
[5] KRIZHEVSKY, A., SUTSKEVER, I., HINTON, G.E. (2012) ImageNet classification with deep convolutional neural networks. In: Neural Information Processing Systems. Lake Tahoe. pp. 1106-1114.
[6] RANZATO, M., BOUREAU, Y.L., LECUN, Y. (2008) Sparse feature learning for deep belief networks. In: Neural Information Processing Systems. Vancouver. pp.1185-1192.

[7] GOODFELLOW, I.I., POUGET-ABADIE, J., MIRZA, M., et al. (2014) Generative adversarial nets. In: Neural Information Processing Systems. Palais des Congrès de Montréal. pp: 2672-2680.

[8] HINTON, G.E., SALAKHUTDINOV, R.R. (2006) Reducing the dimensionality of data with neural networks. Science, 313: 504-507.

[9] HUANG, G., LIU, Z., MAATEN, L.V.D., et al. (2017) Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition. Hawaii. pp: 4700-4708.

[10] SZEGEDY, C., VANHOUCKE, V., IOFFE, S., et al. (2016) Rethinking the inception architecture for computer vision. In: IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas. pp: 2818-2826.

[11] Daniel, O.M., Luige, V. (2020) Facial Expressions Recognition for Human–Robot Interaction Using Deep Convolutional Neural Networks with Rectified Adam Optimizer. Sensors, 20. DOI:10.3390/s20082393.

[12] Yan, H., Zhang, Z., Xiong, G., et al. (2016) Radar HRRP recognition based on sparse denoising autoencoder and multi-layer perceptron deep model. In: IEEE Conference on Ubiquitous Positioning Indoor Navigation and Location Based Services. Shanghai. pp: 283-288.

[13] ZHAO, F.X., LIU, Y.X., HUO, K., et al. (2018) Radar HRRP Target Recognition Based on Stacked Autoencoder and Extreme Learning Machine. Sensors, 18: 173-187.

[14] WAN, J., CHEN, B., XU, B., et al. (2019) Convolutional neural networks for radar HRRP target recognition and rejection. EURASIP Journal on Advances in Signal Processing, 2019:1-9.

[15] LU, W., ZHANG, Y.S., XU, C. (2020) HRRP target recognition method based on bispectrum-spectrogram feature and deep convolutional neural network. Systems Engineering and Electronics, http://kns.cnki.net/kcms/detail/11.2422.TN.20200220.0926.002.html.

[16] YANG, Y.H., SUN, J.M., YU, S.K., et al. (2017) High Resolution Range Profile Target Recognition Based on Convolutional Neural Network. Modern Radar, 3(12): 24-28.

[17] Wahlbeck, K., Tuunainen, A., Ahokas, A., et al. (2001) Dropout rates in randomised antipsychotic drug trials. Psychopharmacology, 155: 230-233.

[18] HONG, S., YOU, T., Kwak, S., et al. (2015) Online Tracking by Learning Discriminative Saliency Map with Convolutional Neural Network. Computer Science, 13: 597-606.

[19] SHUAI, W.Y., WANG, X.D., XUE, A.J. (2014) Simulation of Complete Polarized Range Resolution Profile Based on FEKO. Journal of Projectiles, Rockets, Missiles and Guidance, 34: 173-175.

[20] PAN, M., JIANG, J., KONG, Q., et al. (2017) Radar HRRP target recognition based on t-SNE segmentation and discriminant deep belief network. IEEE Geoscience and Remote Sensing Letters, 14: 1609-1613.

[21] Bengio, Y., Lamblin, P., Dan, P., Hugo, L. (2006) Greedy layer-wise training of deep networks. Proceedings of the Advances in Neural Information Processing Systems, 19: 153-160.

[22] Vincent, P., Larochelle, H., Lajoie, I., ey al. (2010) Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. Journal of Machine Learning Research, 11: 3371-3408.

[23] Le, Q.V., Ngiam, I., Coates, A., et al. (2011) On optimization methods for deep learning. In: The 28th International Conference on Machine Learning. Bellevue. pp: 265-272.