HYPERSPECTRAL TARGET DETECTION BASED ON ONE-DIMENSIONAL GENERATIVE ADVERSARIAL NETWORK

Yulei Wang, Hongzhou Wang, Yuchao Yang, Enyu Zhao, Meiping Song, and Chunyan Yu

Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Information Science and Technology College, Dalian Maritime University, Dalian 116026, China

ABSTRACT

Hyperspectral images provide spectral curves that reflect the "fingerprint" properties of substances, making them suitable for many applications. Thanks to the rapid development of computing resources, deep learning algorithms can significantly improve the cognitive ability of the network by extracting hidden features, and have been successfully applied to hyperspectral image processing, such as classification and detection. In this paper, a new hyperspectral target detection model based on one-dimensional generative adversarial networks (1D-GAN) is proposed. The proposed 1D-GAN network is designed to extract HSI features, and the probability is calculated accordingly whether the pixel to be detected is a target or background. In order to capture the spatial features, the guided filter is then used to obtain the final detection map. Performance comparison with several state-of-the-art methods demonstrate the effectiveness and efficiency of the proposed 1D-GAN algorithm.

Index Terms— Hyperspectral image (HSI), Generative adversarial network (GAN), Semi-supervised learning, target detection

1. INTRODUCTION

Benefiting from the combination of rich spatial and spectral information, hyperspectral images have a wide range of applications in many fields, where target detection has shown great advantages in military investigation, agriculture, forestry, ocean exploration, et al . With a view to distinguish interested targets from a large number of complex backgrounds, hyperspectral target detection (HTD) is can find artificial targets different from natural backgrounds from the perspective of spectral characteristics, which is a common practical need.

Deep neural network (DNN), such as convolutional neural network (CNN) and generative adversarial network (GAN), has been proved with outstanding effect in mining and identifying hidden features from high-dimensional complex hyperspectral images. DNN has been successfully applied to HTD [1],[2],[3],[4], because of its excellent nonlinear recognition and generalization ability, where the DNN with cascaded structure and nonlinear functions has a strong representation ability and can be used for hyperspectral target detection with complex background. With a combination of generative and adversarial, a new deep learning network has emerged recently, called generative adversarial network (GAN), where the algorithms based on GAN framework have become the most advanced methods for two-dimensional natural image processing. However, the traditional GAN framework cannot be directly applied to HSI due to the structure of one-dimensional spectrum of natural images.

This paper proposes a hyperspectral target detection model with one-dimensional structure for generating adversarial networks (ID-GAN). Firstly, a one-dimensional GAN framework is designed for hyperspectral image processing, and is trained using unlabeled samples to obtain the characteristics of samples. Then, the trained GAN is transformed into a binary classification network to distinguish targets of interest and background pixels in a semi-supervised manner by only fine-tuning the last layer of the network with relatively few labeled samples while keeping the parameters of other layers unchanged. Experimental results show that the proposed semi-supervised 1D-GAN model can achieve satisfactory results compared with several state-of-the-art methods.

2. GENERATIVE ADVERSARIAL NETWORK FOR HYPERSPECTRAL TARGET DETECTION

The GAN framework is proposed by Good fellow et al. [8], with a combination of generative and adversarial viewpoints. It consists of two adversarial models: one is the generative model G, whose distribution (p_g) attempts to approximate the distribution of training data (p_r). The other one is the discriminant model D, whose goal is to determine whether the sample is from the training data or from the generative model. The models G and D can be nonlinear mapping functions such as CNN. In order to learn the distribution of real data (p_r), the model G constructs a mapping function from the random noise distribution p_z to the data space G(z; θ_g), where G is a differentiable function represented by the CNN of the parameter θ_g. The second model D can be

IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-2010-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/IGARSS52108.2023.10281894

expressed as D(x; θ_d), where the parameter θ_d outputs a scalar with the indication whether the input is "*real*" or "*fake*" . D(x) is the probability that x came from real data. In the training phase, both G and D are trained simultaneously.

$$\min_G \max_D V(D,G) = E_{x\sim p_r}\left[\log D(x)\right] + E_{x\sim p_g}\left[\log\left(1 - D(G(x))\right)\right] \quad (1)$$

where E denotes the expectation, and D and G are like the two-player min–max game with value function V (D, G). Through several epochs of training, the random noise distribution p_g is approximate to the training data distribution p_r, and the discrimination model D cannot distinguish the difference between the two distributions well.

Inspired by [5], a one-dimensional generative adversarial networks (1D-GAN) is designed in this paper, where the framework G contains four 1-D convolution with size of 1 × 5 and step size of 1, and D consists of four 1-D convolution with size of 1 × 3 and step size of 1, as shown in Figure 1. In order to avoid the loss of detailed semantic features caused by maximum pooling, the 1-D convolution layer is used instead of pooling layer, whose convolution kernel is 1 × 3 with step size of 2, which can better extract higher-level semantic features and distinguish the similarity and difference between spectra. FC represents the fully connected layers, which take random noise distribution p_z as a one-dimensional input. Right next to the FC layer is the batch normalization layer, which stabilizes the learning process of deep GAN by normalizing the input to each cell into a distribution of 0 means and unit variance. This helps to deal with training problems arising from poor initialization of G and helps to implement gradient flows in deeper models. The final layer of G outputs a 1 ×n_{bs} vector, where n_{bs} is the number of spectral bands. It is then offered to D as a “fake” input. When the model training is complete, the distribution of the model over “fake” data p_g will approximate to the distribution of real data p_r. The last layer of the discrimination model D is very important, where the nonlinear mapping layer is composed of ReLU activation functions.

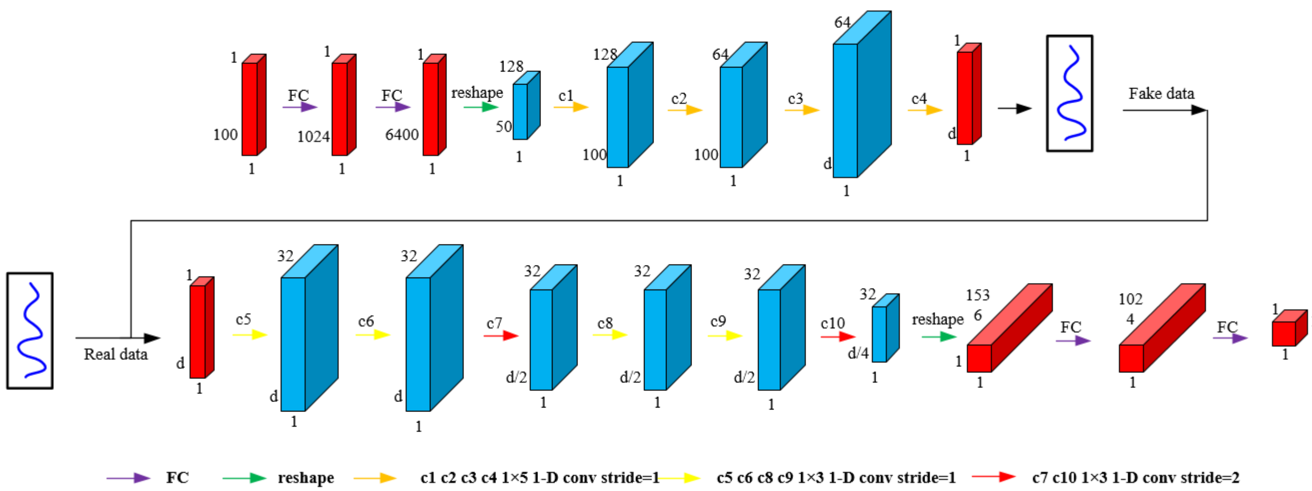


Figure.1 One-dimensional depth generation adversarial network (1D-GAN) structure

In the training stage, a large number of unlabeled samples are used to train the 1-D GAN in order to enable D to distinguish “*real*” and “*fake*” data using the Sigmoid activation function, with a scale output indicating whether the input is "*true*" or "*false*".

In the detection stage, a small number of labeled samples are used firstly to fine-tune the well-trained model D with features of unlabeled samples. In the process of fine-tuning, only the weight of the last layer is updated with weights of all other layers unchanged, so as to obtain a well-trained and fine-tuned classifier C. Finally, the classifier C is used to classify the data to be detected, and the classification results are input into the guided filter to obtain the final detection result.

3. JOINT SPATIAL INFORMATION FOR TARGET DETECTION

The HSI is a 3-D cube with spatial and spectral information. However, the previous 1-D GAN only uses the spectral characteristic to detect the target, while ignoring the spatial information of the HSI. Using spatial information to modify the detection map obtained by 1-D GAN framework could further improve the detection accuracy.

In this paper, the guided filter is used to make use of the spatial information. The principal component analysis (PCA) is firstly performed on the HSI to obtain the guide image, and the first principal component is selected as the guide image. Define a filter that includes a guide image I, an input image K, and an output image O of the joint space-spectral detection result. The filtering output of pixel i is represented as a weighted average:

$$\mathrm{O}_i = \sum_j W_{ij}(I)\mathrm{K}_j \quad (2)$$

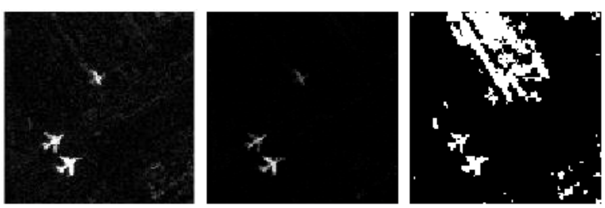
and the filter kernel weight $W_{ij}(I)$ can be expressed as:

$$W_{ij}(I)=\frac{1}{|e|^2}\sum_{k:(i,j)\in e_k}\left(1+\frac{(I_i-\mu_k)(I_j-\mu_k)}{\sigma_k^2+\varepsilon}\right) \quad (3)$$

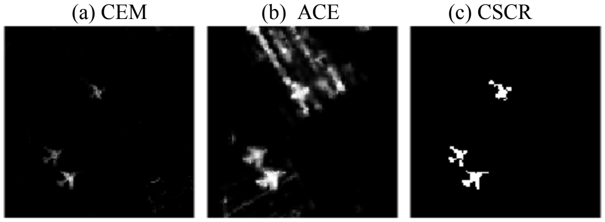
Where e_k is the window centered on the kth pixel with the window size of $(2r+1)\times(2r+1)$ (r is the radius of the window), μ_k and σ_k^2 are the mean and variance of the guiding image respectively, $|e|$ is the number of pixels in e_k, and ε is a penalty value. I_i and I_j refer to the values of two adjacent pixels in the guiding image. After the guided filter, the detection map K obtained by the previous section can be smoother and with the boundary of the target maintained.

4. EXPERIMENTAL RESULTS

In this section, four real hyperspectral images are used to verify the performance of the proposed algorithm 1D-GAN, as shown in Figure 2-5 (AVIRIS1, AVIRIS2, El Segundo, Beach), where the sizes of the images are100 × 100 × 189, 120 × 120 × 189, 100 × 100 × 224, and 90 × 90 × 188, respectively. Four state-of-the-art HTD algorithms are used for comparison, which are three traditional algorithms CEM [6], ACE [7],[8], and CSCR [9], and one deep learning based algorithm BLTSC [10]. The detection results are shown in Figure 2-5 (b)-(f). As can be seen from the detection maps, the result obtained by the proposed 1D-GAN has the most significant difference between the target and the background pixels, with the edge shape of the target maintained well.

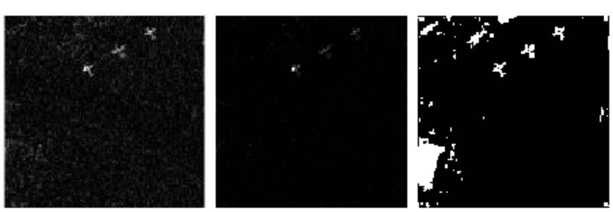


(a) CEM (b) ACE (c) CSCR



(d) BLTSC (e) 1D-GAN (f) Ground-truth

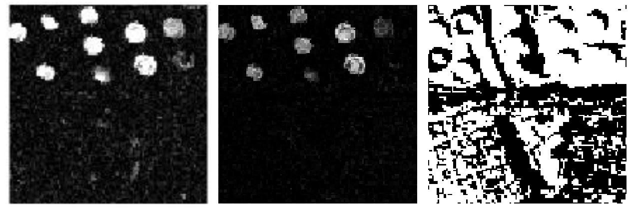
Figure.2 AVIRIS1 data set detection results by different algorithms.



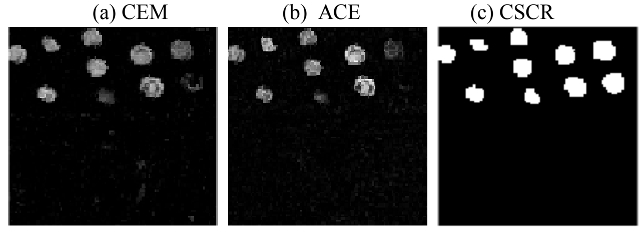
(a) CEM (b) ACE (c) CSCR

(d) BLTSC (e) 1D-GAN (f) Ground-truth

Figure.3 AVIRIS2 data set detection results by different algorithms.

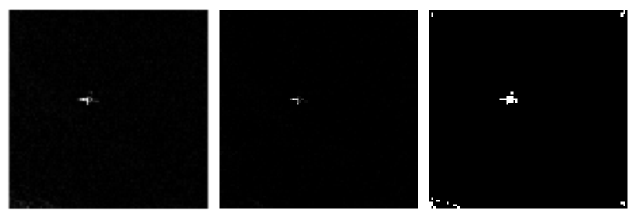


(a) CEM (b) ACE (c) CSCR

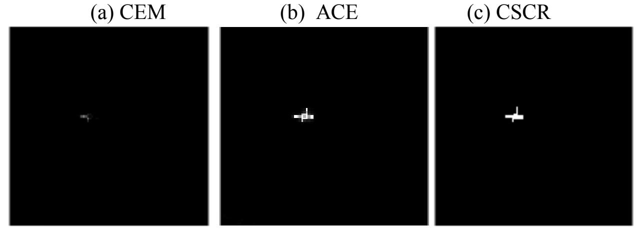


(d) BLTSC (e) 1D-GAN (f) Ground-truth

Figure.4 EI Segundo data set detection results by different algorithms.



(a) CEM (b) ACE (c) CSCR



(d) BLTSC (e) 1D-GAN (f) Ground-truth

Figure.5 Beach data set detection results by different algorithms.

In order for a quantitative comparison, the receiver operating characteristic (ROC) curve and the area under curve (AUC) are used in this section, which are widely used in HTD. Figure 6 shows the ROC curves of five detection algorithms, where each subgraph includes all algorithms with the same HSI data. It could be seen that, the ROC curve of the proposed 1D-GAN detector is always higher than those of the other four algorithms.

Table I gives the AUC values of five algorithms using four different HSI data sets, which further prove the superior comprehensive performance of the proposed 1D-GAN detector with best AUC for all four data sets, highlighted in coarsened form.

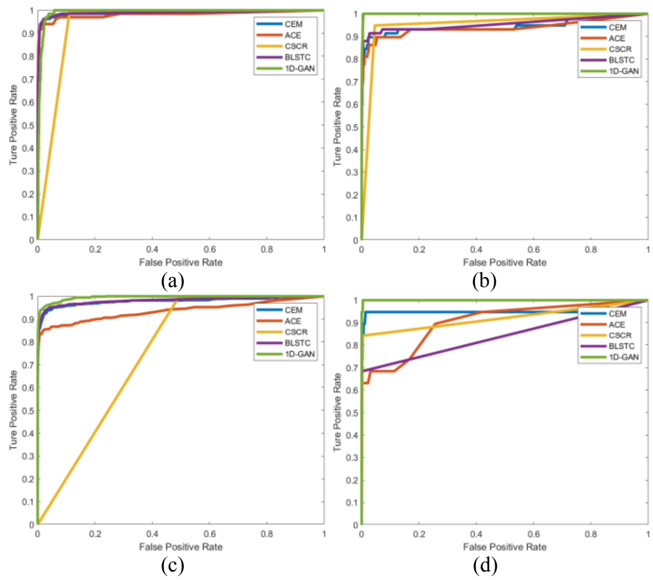


Figure.6 ROC curve of different HSI datasets. (a) AVIRIS1, (b) AVIRIS2, (c) El Segundo, (d) Beach.

TABLE I
AUC VALUES FOR DIFFERENT METHODS ON FOUR DATASETS

Method	AVIRIS1	AVIRIS1	El Segundo	Beach
CEM	0.9909	0.9457	0.9785	0.9534
ACE	0.9818	0.9376	0.9359	0.9026
CSCR	0.9408	0.9512	0.7516	0.9196
BLTSC	0.9891	0.9551	0.9813	0.9512
ODGAN	**0.9920**	**0.9999**	**0.9951**	**0.9998**

5. CONCLUSIONS

This paper proposes a new deep learning based hyperspectral target detection method using 1-D generative adversarial network (1D-GAN). A large number of unlabeled samples are used firstly to train 1-D GAN framework to extract the image features, and calculate the probability whether the tested pixel is target or background. The well-trained network is then fine-tuned through a small number of labeled samples to get the target-background classifier, and finally a guided image filter is used to get the final detection results with spatial-spectral information. Compared with other GAN models, the proposed 1D-GAN uses Batch Normalization layer in its network design, which deals with the training problems caused by poor initialization of the generator and helps to implement gradient flow in the deeper model and stabilize the learning process of the deep GAN. At the same time, compared with other deep learning model, the 1D-GAN uses fewer super-parameters and only needs to manually set the parameters of the guided filter in the detection stage. Experimental results show that the proposed 1D-GAN is superior to several traditional and deep learning-based detection methods.

6. ACKNOWLEDGEMENTS

This work is supported by the National Nature Science Foundation of China (61801075, 42271355), the Natural Science Foundation of Liaoning Province (2022-MS-160), the China Postdoctoral Science Foundation (No. 2020M670723), and the Fundamental Research Funds for the Central Universities (3132023238).

REFERENCES

[1] J.-F. Hu and T.-Z. Huang, "Hyperspectral image super-resolution via deep spatiospectral attention convolutional neural networks," *IEEE Trans.Neural Netw. Learn. Syst.*, Jun. 9, 2021, doi: 10.1109/TNNLS.-2021.3084682.

[2] W. Li, G. Wu, and Q. Du, "Transferred deep learning for hyperspectral target detection," *IEEE Int. Geosci. Remote Sens. Symp.(IGARSS)*, Jul. 2017, pp. 5177–5180.

[3] G. Zhang and S. Zhao, "HTD-net: A deep convolutional neural network for target detection in hyperspectral imagery," *Remote Sens.*, vol. 12, no. 9, p. 1489, 2020.

[4] Y. Wang and X. Chen, "Meta-Learning Based Hyperspectral Target Detection Using Siamese Network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, doi: 10.1109/TGRS.2022.3169970

[5] C.-I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Trans. Geographical science. Remote Sens.*, vol. 44, no. 6, pp. Jun. 2006.

[6] C.-I. Chang and D. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1114-1159, May 2000.

[7] D. Manolakis and D. Marden, "Hyperspectral image processing for automatic target detection applications," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 79–116, 2003.

[8] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. Sep. 1999.

[9] W. Li and Q. Du, "Combined sparse and collaborative representation for hyperspectral target detection," *Pattern Recognit.*, vol. 48, no. 12, pp. 3904-3916, Dec. 2015

[10] W. Xie and X. Zhang et al., "Background learning based on target suppression constraint for hyperspectral target detection," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5887–5897, 2020.