

# Polarimetric SAR image classification using 3D generative adversarial network

Lu Liu<sup>1,2,\*</sup> and Guobao Feng<sup>3</sup>

<sup>1</sup>School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China

<sup>2</sup>Shaanxi Key Laboratory for Network Computing and Security Technology, Xi'an University of Technology, Xi'an 710048, China

<sup>3</sup>National Key Lab of Science and Technology on Space Microwave, China Academy of Space Technology, Xi'an 710100, China

**Keywords:** Polarimetric SAR image classification, Generative adversarial network, Three-dimensional convolutional.

**Abstract.** In this paper, a new architecture of three-dimensional deep convolutional generative adversarial network(3D-DCGAN) is specially defined to solve the unstable training problem of GAN and make full use of the information involved in polarimetric data. Firstly, a data cube with nine components of polarimetric coherency matrix are directly used as the input features of DCGAN. After that, a 3D convolutional model is designed as the components of generator and discriminator to construct the 3D-DCGAN, which considers the effective feature extraction capability of 3D convolutional neural network(CNN). Finally parameters of the network are fine-tuned to realize the polarimetric SAR image classification. The experiments results show the feasibility and efficiency of the proposed method.

## 1 Introduction

Polarimetric SAR(PolSAR) data has more characteristics of polarimetric, phase and space, providing reliable data resources for precise land cover classification. However it is difficult to mark samples artificially, so there is a lack of labeled samples in PolSAR images.

At present, many kinds of classification methods based deep neural network have been used in the field of PolSAR image classification, such as Deep Belief Network (DBN) [3], Convolutional Neural Network (CNN)[4] and Generative Adversarial Network (GAN) [5], etc. As an unsupervised learning method, GAN are widely used because they can expand polarimetric data, learn the potential distribution of real data, and generate fake data with the same distribution as real data. However the training process of GAN[6] is unstable and easily falling to the local optimum, various various variants of GAN have been proposed to solve the problem, such as Deep Convolutional GAN(DCGAN)[7], Conditional GAN(CGAN)[8], Wasserstein GAN[9], etc.

---

\* Corresponding author: [liulu0613@163.com](mailto:liulu0613@163.com)

Combining the advantages of CNN and GAN, DCGAN has a high feature representation ability. The hidden layer of the discriminator model is all realized by the convolutional layer, and the hidden layer of the generator model is realized by the deconvolution layer. However the PolSAR data has three channels which contains amplitude, phase and spatial information, these operations in 2D space tend to lose the spatial and phase information

In this paper, a new type of three-dimensional(3D) DCGAN is specially defined to solve the unstable training problem of GAN and make full use of the information involved in polarimetric data. First of all, a data cube with nine components of polarimetric coherency matrix are directly used as the input features of DCGAN. And then 3D convolutional model is designed as the compoents of generator and discriminator to construct the 3D-DCGAN, which combines the 3D convolutional, pooling and nonlinear operators, as well as backpropagation. Finally parameters of the network are fine-tuned to realize the classification of PolSAR images. The experiments results show the feasibility and efficiency of the proposed method.

## 2 Related Background

### 2.1 Deep Convolutional GAN

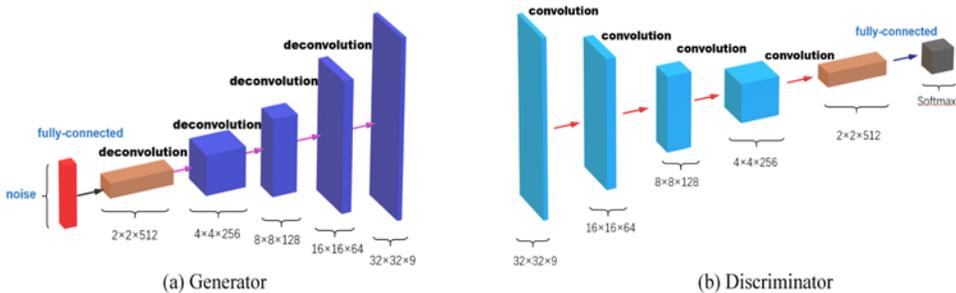
DCGAN combines CNN and GAN. On the basis of the original GAN, the hidden layer of the discriminative model is all realized by the convolutional layer, and the hidden layer of the generated model is realized by the deconvolution layer[10].

GAN is composed of generator  $G$  and discriminator  $D$ . Its objective function is [5]:

$$\min_G \max_D L_{GAN}(D, G) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

where,  $x$  is the real polarimetric data,  $z$  is the random noise, and  $G(z)$  is the sample generated for  $G$  that conforms to the distribution of real data as far as possible.

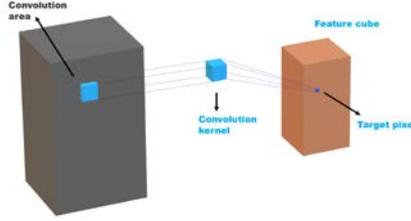
The structure of the generator and discriminator is set as shown in Fig 1(a) and (b).



**Fig. 1.** The framework of generator and discriminator.

### 2.2 3 D Convolutional GAN

3D convolutional model is employed instead of 2D convolutional model in DCGAN. The input data used in the proposed method are the elements of  $3 \times 3$  complex coherency matrix  $T$  forms a feature matrix with the size  $m \times n \times 9$  ( $m$  and  $n$  are the width and heigh of the image respectively). The process of 3D convolution is shown in Fig. 2.



**Fig. 2.** 3D convolution process.

The output value of the neuron at the position of  $(x, y, z)$  in the hidden layer  $i$  is as follows[11]:

$$v_{ij}^{xyz} = f \left( b_{ij} + \sum_m \sum_{h=0}^{H_i-1} \sum_{n=0}^{N_i-1} \sum_{d=0}^{D_i-1} w_{ijm}^{hnd} v_{(i-1)m}^{(x+h)(y+n)(z+d)} \right) \quad (7)$$

where,  $v$  is the output of position  $(x,y,z)$  of layer  $i$ , the convolution kernel size of the  $i$ th layer is  $H_i \times N_i \times D_i$ ,  $f(\cdot)$  is the activation function,  $b_{ij}$  is the bias of the  $j$ th data map in the  $i$ th layer,  $m$  represents a data map in the  $(i-1)$ th layer that connects the  $j$ th feature map,  $w_{ijm}^{hnd}$  is the weight between the neuron on position of  $(h,n,d)$  of the  $j$ th feature map on the  $i$ th layer and the  $m$ th feature map on the  $(i-1)$ th layer.

The formula of 3D maximum pooling is as follows:

$$v_{x,y,z} = \max_{0 \leq i \leq M, 0 \leq j \leq N, 0 \leq k \leq K} \left( u_{x \times m+i, y \times n+j, z \times p+k} \right) \quad (8)$$

where,  $u$  represents 3D input vector of pooling layer,  $v$  is output of pooling,  $m, n, p$  represents sampling steps and  $M \times N \times K$  represents the 3D area.

### 3 Proposed Method

The details of the proposed algorithm for PolSAR data classification are described as follows:

(1) The input data used in the proposed method are the elements of  $3 \times 3$  complex coherency matrix  $T$  including  $T_{11}, T_{22}, T_{33}$  and the real part and imaginary part of  $T_{12}, T_{13}$  and  $T_{23}$ .

(2) The sliding window method with step size 1 is adopted to take  $32 \times 32$  image blocks around each pixel to represent the original pixel value, thus forming  $N$  image blocks with size  $32 \times 32 \times 9$ .

(3) The unlabeled training data set, the labeled training data set and the test data set were constructed.

(4) The 3D-DCGAN model was trained with no label training data set. The feature map of each layer is given, and the dimensions of each deconvolution layer are determined, and weights and biases are randomly initialized.

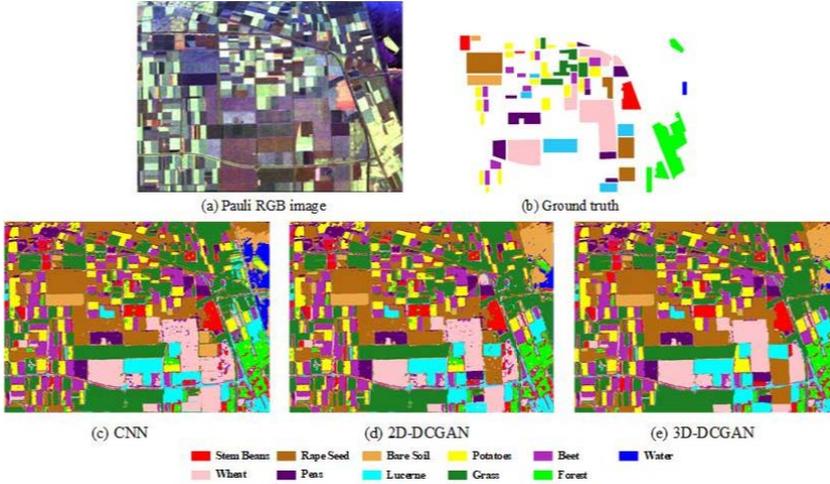
(5) After data is input, each step of training updates the parameters of the discriminant model.

(6) Softmax classifier was trained by using the labeled training data set as the input of the discriminative classification model and the category of each sample in the labeled training data set as the output of the discriminative classification model.

(7) The parameters of the whole discriminator are fine-tuned with the training data set to form the trained discriminator.

## 4 Experiments

The polarimetric data used in this experiment was acquired by AIRSAR over Flevoland, with a size of  $750 \times 1024$ . Pauli RGB image and the ground truth map are shown in Fig.4(a) and (b). The area contains eleven types of crops: stem beans, rapeseed, bare soil, potatoes, beet, wheat, peas, lucerne, grass, forest and water. 2D-DCGAN and CNN method is used for comparison. Classification results are evaluated by overall accuracy (OA) and Kappa coefficient.



**Fig. 4.** Classification results of Flevoland image: (a)Pauli RGB image, (b)Ground truth, (c) CNN, (d)2D-DCGAN, (e) 3D-DCGAN.

**Table 1.** Classification Accuracy Comparison.

Methods	CNN	2D-DCGAN	3D-DCGAN
Stem Beans	0.8032	0.8740	0.8665
Rape Seed	0.9308	0.9235	0.9486
Bare Soil	0.7562	0.8680	0.9883
Potatoes	0.8693	0.9092	0.9440
Beet	0.8843	0.9257	0.9734
Wheat	0.8026	0.8831	0.8868
Peas	0.8500	0.8993	0.9496
Lucerne	0.8776	0.8884	0.9870
Grass	0.7629	0.8026	0.8960
Forest	0.9307	0.9308	0.9574
Water	0.8993	0.9003	0.8895
OA	0.9090	0.9480	0.9832
Kappa Coff.	0.9180	0.9370	0.9490

Fig.4 (c)-(e) illustrates that all the three methods can achieve a satisfactory partition of rape seed and forest. However CNN cannot reliably classify grass, wheat, stem beans and bare soil, which directly leading to the lower OA and Kappa coefficient. 2D-DCGAN and the proposed 3D-DCGAN have improved the classification of wheat and beet. The proposed method which adopts the DCGAN model, is superior to CNN and achieves notable improvement in region connectivity, uniformity and robustness. Compared with DCGAN, our method get a higher OA for adopting 3D convolution model instead of 2D convolution model, which can make full use of the spatial information in PolSAR data. At the same time, table1 represents that the OA of the proposed method achieves 98.32%, which improves 3.52% and 7.42% respectively, while Kappa coefficient reaches 94.90%.

## 5 Conclusions

In this paper, a new architecture of 3D-DCGAN is proposed. We apply a 3D convolution model attribute to the feature extraction capability of 3D convolutional network is stronger than that of 2D convolutional network. Meanwhile, the introduction of 3D convolution model follows the physical scattering mechanism and holds the polarimetric and spatial information involved in polarimetric data, which can also alleviate the unstable training problem of GAN. The comparative experiments results show the feasibility and efficiency of the proposed method.

## Acknowledgements

This work was supported in part by the Natural Science Foundation of Shaanxi Province under Grant 2020JQ-644, in part by the Scientific Research Plan Projects of Shaanxi Education Department [20JK0808], and in part by the China Postdoctoral Science Foundation Funded Project [2017M623212].

## References

1. K. Ersahin, I. G. Cumming, R. K. Ward, IEEE Trans. Geosci. Remote Sens., **48**,1 (2010)
2. B. Hou, C. Chen, X. Liu, L.C. Jiao, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., **8**, 11 (2015)
3. F. Liu, L.C. Jiao, B. Hou, S. Yang, IEEE Trans. Geosci. Remote Sens., **54**, 6 (2016)
4. Y. Zhou, H. Wang, F. Xu, Y. Jin, IEEE Geosci. Remote Sens. Lett., 13, 12 (2016)
5. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, NIPS(2014)
6. F. Liu, L.C. Jiao, X. Tang, IEEE Trans. Neural Netw. Learn. Syst., **30**, 9 (2019)
7. W. Fang, Y. Ding, F. Zhang, J. Sheng, IEEE Access, **7** (2019)
8. L. Li, C. Wang, H. Zhang, B. Zhang, IEEE Geosci. Remote Sens. Lett., (to be published)
9. M. Zhang, M. Gong, Y. Mao, J. Li, Y. Wu, IEEE Trans. Geosci. Remote Sens., **57**, 5 (2019)
10. Q. Sun, X. Li, L. Li, X. Liu, F. Liu, L. Jiao, IGARSS (2019)
11. M. Liu, Y. Hu, S. Wang, Y. Guo, B. Hou, L.C. Jiao, X. Hou, IGARSS (2018)
12. L. Zhang, Z. Chen, B. Zou, Y. Gao, IGARSS (2018)