# HYPERSPECTRAL IMAGE CLASSIFICATION BASED ON GENERATIVE ADVERSARIAL NETWORK WITH DROPBLOCK

*Jihao Yin*, Wenyue Li, Bingnan Han*

School of Astronautics, Beihang University, Beijing 100191, China

## ABSTRACT

Deep learning (DL) algorithms are widely applied in hyperspectral images (HSIs) classification. However, the insufficient utilization in spatial semantic information and inadequate number of HSIs samples both restrict the classification performance of DL-based HSIs algorithms. In this paper, we propose a novel method based on generative adversarial network (GAN) with DropBlock structure (DBGAN). Specifically, DropBlock enforces each unit in convolution neural network (CNN) to learn features by dropping contiguous regions of feature maps, therefore more spatial semantic information is capable to contribute in HSIs classification. Furthermore, GAN model can generate realistic samples by an adversarial game to mitigate HSIs data shortage. Extensive experimental comparisons demonstrate the effectiveness of the proposed method.

***Index Terms*—** Hyperspectral classification, generative adversarial networks, spatial semantic information

## 1. INTRODUCTION

Hyperspectral images (HSIs) contain abundant spectral information and spatial information, which is particularly beneficial to discriminate various materials in the observed area [1] [2]. As an important application of hyperspectral remote sensing, HSIs classification has been widely studied in recent years [3] [4] [5]. Among the proposed classification algorithms, deep-learning (DL) methods have received considerable attention [6] [7] [8]. Instead of using hand-craft features, DL algorithms adopt convolution neural networks (CNN) to automatically extract multifarious spectral and spatial features to assist classification. Although these DL algorithms have achieved acceptable classification performance, their effects are still restricted by two main shortcomings: the insufficient utilization of spatial semantic information and the limited quantity of HSIs samples.

Firstly, traditional DL methods fail to utilize spatial semantic information efficiently by only using fixed neighborhood region. For example, authors in [6] adopt a K*K neighborhood region for central pixel and then feed it to CNN. By this means, the spatial semantic information in a single neighborhood region, which is contributing in classification process, is limited. In other words, it is not guaranteed that the continuous spatial semantic information from each part of a neighborhood region can be learned and used sufficiently. Besides this, CNN receives massive overlapping areas from different K*K regions as input and extracts a great amount of identical spatial semantic features, those feature redundancies will seriously affect the effectiveness of the networks.

Secondly, the labeled HSIs data that can be used in training is inadequate. Generally, a great quantity of data is necessary for DL algorithms in training. However, collecting and labeling HSIs data is difficult, expensive and time-consuming. In this case, serious overfitting phenomenon will be occurred and thus leads to worse test performance [9].

In order to make full use of spatial semantic information in classification, we exploit DropBlock structures to assist CNN [10]. DropBlock is a novel DL method proposed by Ghiasi which can exploit the semantic information. DropBlock acts on the feature maps of the hidden layers, so that the units in a random contiguous region of feature maps are dropped together. The networks will fully use the remaining semantic information and force them to extract different features for classification by randomly dropping some units. Based on this, it is a considerable way to use DropBlock in HSIs classification problem.

To mitigate data shortage, Generative Adversarial Networks (GAN) [11] which is an emerging DL generative model can be used. The GAN model, first put forward by GoodFellow, contains a generator and a discriminator. These two parts can capture the data distribution by playing an adversarial game. Based on the classic GAN, a series of GAN models were proposed, such as CATGAN [12], INFOGAN [13]. Up to now, several GAN-based HSIs classification algorithms have been developed and obtain good performance [14] [15]. Therefore, it is reasonable to adopt GAN model to data augment in HSIs classification.

In this paper, we propose a novel GAN-based algorithm with DropBlock (DBGAN) for HSIs classification. DropBlock is applied for spatial semantic information utilization both in overlapping areas and different unblocked units. Meanwhile, the GAN model is provided to generate HSIs images which can distinctly extend sample set. We demonstrate that our model can offer significantly improved classification performance on several frequently used datasets compared to other related classifiers.

The rest of the paper is organized as follows. In section 2, we present our algorithm including DropBlock for HSIs and the structure of our model. Section 3 presents the details of experimental results. Conclusions and discussions are reported in section 4.

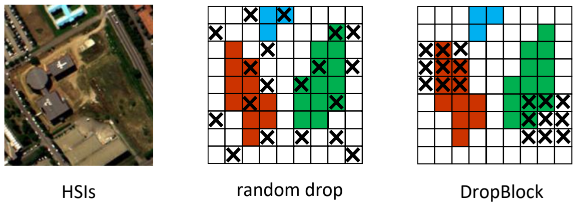


**Fig. 1** The comparison between random drop and DropBlock on a feature map of HSIs.

## 2. METHOD

This section introduces the proposed DBGAN. We will present the DropBlock for HSIs and the structure of DBGAN respectively.

### 2.1. DropBlock for HSIs

Spatial semantic information plays an important role in HSIs classification and has been utilized in DL-based HSIs algorithms. The $K \times K$ neighborhood region for central pixel is a common and effective method to employ spatial semantic information. However, based on the previous analysis, the spatial semantic information in neighborhood region is not fully used. Therefore, we adopt DropBlock structure to guarantee that spatial semantic information can take effect in HSIs classification.

DropBlock works by dropping continuous regions of units in the feature maps from hidden layers, therefore, some certain semantic information can be removed and remaining units are enforced to be effective in classification. Fig. 1 demonstrates the comparison between random drop, which drops discrete units randomly on feature maps, and DropBlock. As shown in this figure, compared to dropping continuous regions in DropBlock structure, dropping discrete units randomly will destroy the relation between nearby units which contain closely related information. In the process of DropBlock, a mask $M$ is firstly sampled of a feature map and each element $M_{i,j}$ in $M$ subjects to Bernoulli distribution, then the zeros in $M$ is expended to a zero block. There are two important parameters of DropBlock, $Block\_size$ is the size of the block to be dropped, and $\gamma$ controls how many units to drop. We can obtain $M$ by $\gamma$:

$$M: M_{i,j} \sim Bernoulli(\gamma) \quad (1)$$

and $\gamma$ can be calculated by the proportion of units to retain $keep\_prob$ and the size of feature map $feat\_size$:

$$\gamma = \frac{1 - keep\_prob}{block\_size^2} \frac{feat\_size^2}{(feat\_size - block\_size + 1)^2} \quad (2)$$

The advantages of DropBlock in HSIs classification can be summarized as follows. Firstly, with the application of DropBlock, most of spatial semantic information in neighborhood region can be focused and make decision in test procedure, rather than a limited certain part that decides the final category of the central pixel principally. Fig. 2 shows the principle of DropBlock structure. When we employ DropBlock structure to remove some spatial semantic information in every epoch of training process, the remaining units will be enforced to impact training. Meanwhile, it is also notable in Fig. 2 that the dropped information is not lost, due to the reason that the mask M for DropBlock is constantly changing, instead of fixed, in training process. Furthermore, in different epochs of training, the M subjects to a new Bernoulli distribution, so the remaining functional units are entirely different. Following this approach, the spatial semantic information is not only retained but also adequately learned and used.

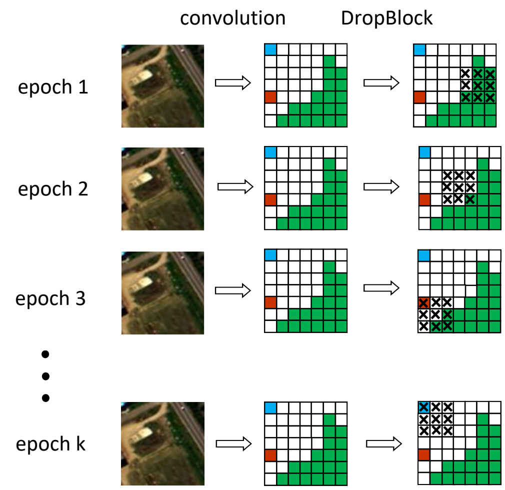


**Fig. 2** The principle of DropBlock in training process.

Secondly, DropBlock can solve the identical feature redundancy and increase diversity of features in HSIs classification. Because of the low image spatial size that most available HSIs data has and more importantly the spatial

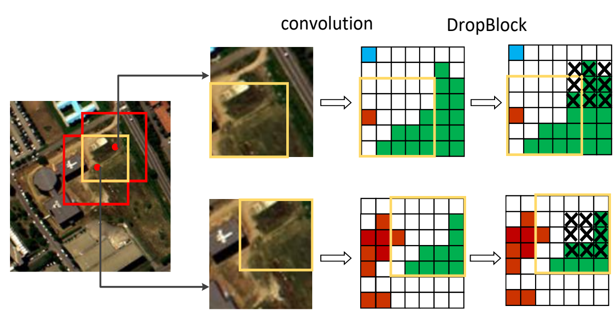


**Fig. 3** The illustration of DropBlock in processing overlapping area of HSIs. Red windows represent neighborhood regions for central pixels. Yellow square represents the overlapping area of two neighborhood regions.

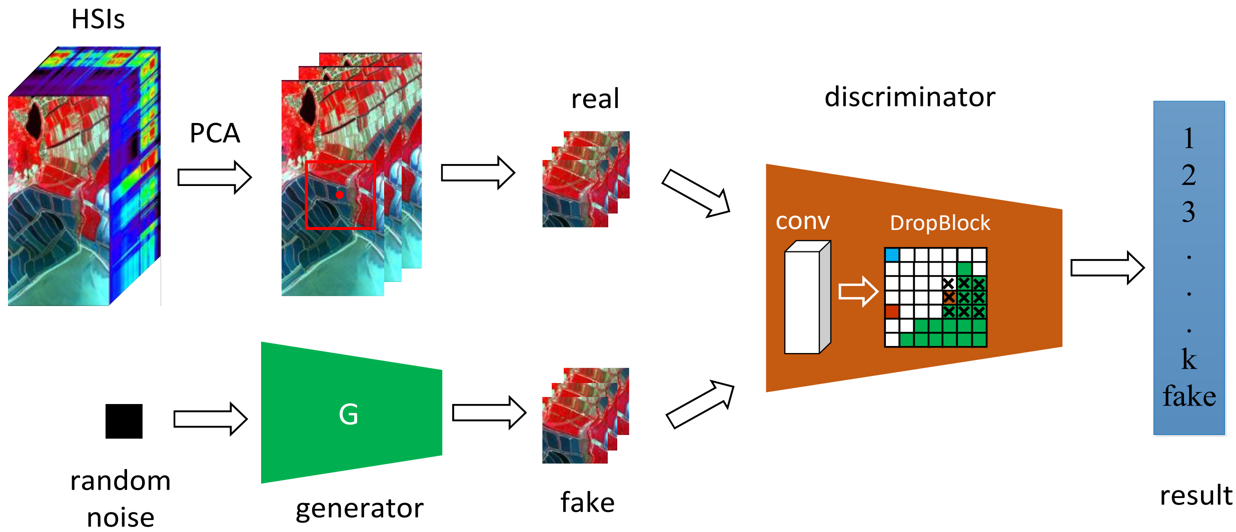


**Fig. 4** The framework of DBGAN.

correlation utilization method using $K \times K$ neighborhood region, the serious overlapping phenomenon is occurred between neighborhood regions of different central pixels. As shown in Fig. 3, in convolutional layer, the spatial semantic features which extracted from overlapping area are identical. As a consequence, a great number of redundant features would significantly influence the network efficiency. However, by randomly dropping unit blocks in different feature maps, DropBlock can extract entirely different spatial semantic features by analyzing unique unblocked remaining units in overlapping area. In this way, the feature redundancy is reduced meanwhile the diversity of spatial feature can be increased.

### 2.2. DropBlock generative adversarial networks

GAN is a novel deep learning framework which trains deep generative models using a minimax game. To learn a generative distribution that matches the real data distribution, two sub-models of GAN, a generator and a discriminator, are trained adversarially. The generator produces samples from the generative distribution by transforming a noise variable into a sample, while the discriminator aims to distinguish between samples from the true data distribution and the generative distribution. As the training process continues, generative ability of the generator and the discriminative ability of the discriminator both continuously enhance. Finally, the generator will output realistic samples and the discriminator will be unable to differentiate the generative data and real data.

Inspired by GAN, we design a DropBlock GAN (DBGAN) which sets DropBlock structures in an end-to-end GAN-based HSIs classification model. A large number of generated HSIs images are provided by the generator of DBGAN to data augment and the DropBlock structure can assist the discriminator in adequate utilization of spatial semantic information. The data augment method of DBGAN mitigates overfitting phenomenon in training and it is a completely different strategy compared with traditional data augmentation techniques such as rotate, translate, scale. The generator and the discriminator are both modified from CNN. As shown in Table.1, the generator contains fully-connected layers, convolution layers and up-sampling layers, while the discriminator contains only fully-connected layers and convolution layers. It is worth noting that dropping units of deep layers which have large receptive fields will loss too much information in classification, so DropBlock (DB) structures are set behind the first three shallow convolution

**Table 1** The structure of DBGAN.

| Net | No. | Conv | Pooling | DB |
|---|---|---|---|---|
| G | 1 | FC:1280 | No | No |
| | 2 | Up: $2\times 2$ | No | No |
| | 3 | $3\times 3\times 64$ | No | No |
| | 4 | Up: $2\times 2$ | No | No |
| | 5 | $3\times 3\times 3$ | No | No |
| D | 1 | $3\times 3\times 32$ | No | Yes |
| | 2 | $3\times 3\times 32$ | $2\times 2$ | Yes |
| | 3 | $3\times 3\times 32$ | No | Yes |
| | 4 | $3\times 3\times 32$ | $2\times 2$ | No |
| | 5 | $3\times 3\times 32$ | No | No |
| | 6 | FC:k+1 | No | No |

layers in the discriminator, and we set $Block\_size$ and $keep\_prob$ as 3 and 0.85, respectively. Based on the end-to-end GAN model and DropBlock structures, the classification performance is capable to be promoted drastically.

Fig. 4 shows the framework of DBGAN and in order to speed up the training and stabilize DBGAN model, we condense the whole HSIs to a suitable dimension using PCA. The generator receives random noise $\rho_z(z)$ as input and generates fake neighborhood regions, and the discriminator obtains real neighborhood regions combining with generated fake neighborhood regions as input. Convolutional layers extract features with the assistance of DropBlock structures and the features are finally fed to the last layer for classification. The last layer is no longer a binary classifier which just distinguish whether the sample is real or fake, a K+1 classifier is adopted in this paper which K is the real classes and another one category is the class for fake. The loss function includes two parts, in which the $L_{real}$ is obtained by the cross-entropy between the in neighborhood regions $\rho_{\text{data}}(x,y)$ training set and their predicted class $\rho_{model}(x,y)$, and the $L_{fake}$ is designed according to both real neighborhood regions $\rho_{data}(x)$ and fake neighborhood regions $G(z)$ into K real classes and class of fake, respectively. Based on these, the loss function becomes:

$$L = L_{real} + L_{fake} \tag{3}$$

$$L_{real} = -E_{x,y\sim\rho_{data}(x,y)} \log_{\rho_{model}}(y \mid x, y < k+1) \tag{4}$$

$$\begin{aligned} L_{fake} = -\{&E_{x\sim\rho_{data}(x,y)} \log[1-\rho_{model}(y=k+1 \mid x)] \\ &+ E_{x\sim G(z)} \log[\rho_{model}(y=k+1 \mid x)]\} \end{aligned} \tag{5}$$

## 3. EXPERIMENT

In this section, two widely used HSIs datasets with different environment settings are adopted to validate the proposed DBGAN. The first dataset was captured by Reflective Optics System Imaging Spectrometer sensor over Pavia, Italy. It is called here Pavia University. There are 9 classes over the land in Pavia University with a size of $610\times340\times103$. Another

**Table 2** The classification results on Pavia University.

| Method | CNN | GAN | CNN(DB) | DBGAN(ours) |
|---|---|---|---|---|
| OA(%) | 85.49±0.51 | 89.69±0.33 | 90.74±0.33 | **92.75±0.13** |
| AA(%) | 78.54±1.03 | 83.47±0.32 | 84.74±1.29 | **87.77±1.19** |
| k*100 | 80.77±0.68 | 86.35±0.44 | 87.73±0.43 | **90.41±0.19** |

**Table 3** The classification results on Indiana Pines.

| Method | CNN | GAN | CNN(DB) | DBGAN(ours) |
|---|---|---|---|---|
| OA(%) | 84.59±0.58 | 88.15±0.44 | 88.02±0.34 | **90.60±0.44** |
| AA(%) | 78.01±0.75 | 83.33±0.39 | 81.51±0.86 | **85.21±0.90** |
| k*100 | 82.46±0.58 | 86.49±0.37 | 86.36±0.66 | **89.29±0.50** |

dataset, Indiana Pines, was collected by AVIRIS sensors in June 1992 over the Indian Pines region in Northwestern Indiana. Size of the data is $145\times145\times220$. There are 16 classes of land covers in this data.

Table.2 and Table.3 display the performance of DBGAN compared with other HSIs classification algorithms. To be specific, CNN is traditional DL-based classification proposed in [6], GAN is a GAN-based classification method proposed in [14], the method in [15] which only uses spectral information performs worse compared with algorithms which combine spatial information. And CNN-DropBlock is that we combine the CNN model and DropBlock structures to compare with DBGAN. We randomly choose 5% of data in Pavia University and 1% of data in Indiana Pines from each class for training respectively. From the two table, we can notice that DBGAN can obtain better overall accuracy (OA), average accuracy (AA) and Kappa (k) than other algorithms. Besides, the classification accuracies for Pavia University are evaluated from a visual perspective. Fig .5 shows the classification maps for different methods on Pavia University.

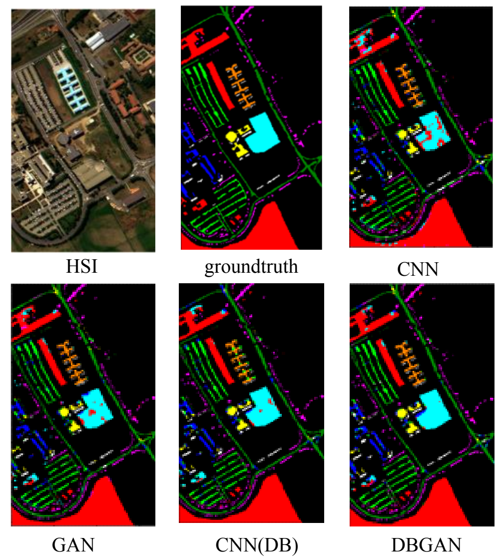


**Fig. 5** The classification maps for different methods on Pavia University.

## 4. CONCLUSION

In this paper, we propose a novel GAN-based HSIs classification algorithm DBGAN. In this method, DropBlock structure is employed to make the best use of spatial semantic information and the GAN model can generate samples to assist classification. In the future, we will exploit the DropBlock in the spectral domain of HSIs and adopt suitable GAN model for spectrum to further improve our classification performance.

## 5. REFERENCE


[1] R. M. Cavalli, S. Pascucci and S. Pignatti, "Hyperspectral remote sensing data to map hazardous materials in a rural and industrial district: The Podgorica dwellings case studies," *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Grenoble, 2009, pp. 1-4.

[2] K. D. Singh, D. Ramakrishnan and L. Mansinha, "A novel method for automatic minerals identification and their abundance estimation for material type discrimination using hyperspectral data," *2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Shanghai, 2012, pp. 1-4.

[3] Bingnan Han, Jihao Yin, Aiqun Xiao, Xiaoyan Luo, "Hyperspectral hierarchical representation based on density peak ranking," *J. Appl. Remote Sens*. 13(1), 016508 (2019).

[4] H. Qv, J. Yin, X. Luo and X. Jia, "Band Dual Density Discrimination Analysis for Hyperspectral Image Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 7257-7271, Dec. 2018.

[5] J. Yin, H. Qv, X. Luo and X. Jia, "Segment-Oriented Depiction and Analysis for Hyperspectral Image Data," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3982-3996, July 2017.

[6] Y. Chen, H. Jiang, C. Li, X. Jia and P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232-6251, Oct. 2016.

[7] M. He, B. Li and H. Chen, "Multi-scale 3D deep convolutional neural network for hyperspectral image classification," *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, 2017, pp. 3904-3908.

[8] P. Jia, M. Zhang, W. Yu, F. Shen and Y. Shen, "Convolutional neural network based classification for hyperspectral data," *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Beijing, 2016, pp. 5075-5078.

[9] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.

[10] Ghiasi, Golnaz, Tsung-Yi Lin, and Quoc V. Le. "DropBlock: A regularization method for convolutional networks." Advances in Neural Information Processing Systems. 2018.

[11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David WardeFarley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *In Advances in neural information processing systems*, pages 2672–2680, 2014.

[12] Jost Tobias Springenberg. "Unsupervised and semi-supervised learning with categorical generative adversarial networks," *arXiv preprint arXiv: 1511.06390*, 2015.

[13] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.

[14] L. Zhu, Y. Chen, P. Ghamisi and J. A. Benediktsson, "Generative Adversarial Networks for Hyperspectral Image Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5046-5063, Sept. 2018.

[15] Y. Zhan, D. Hu, Y. Wang and X. Yu, "Semisupervised Hyperspectral Image Classification Based on Generative Adversarial Networks," in *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 212-216, Feb. 2018.