CBS-GAN: A Band Selection Based Generative Adversarial Net for Hyperspectral Sample Generation

Yulin Qiao1,2,*, Mostofa Zaman Mohammad1,2, Yi Li1,2, Xiaobo Liu1,2,4, and Zhihua Cai3,4

1School of Automation, China University of Geosciences (Wuhan), 430074 Wuhan, China
2Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, 430074 Wuhan, China
3School of Computer Science, China University of Geosciences (Wuhan), 430078, Wuhan, China
4Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430078, China

* Corresponding author: ylqiao@cug.edu.cn

Abstract. Sample generation is an effective method to improve the performance of hyperspectral image classification by generating virtual samples for training sample expansion in the training process of classification. However, there are some defects existing in the previous sample generation methods including the lack of spatial information, the redundant generation and the damage of the original spectral components. In this paper, we propose conditional band selection generative adversarial net, named CBS-GAN, to handle this problem. Firstly, the band selection net of CBS-GAN is utilized to avoid redundant bands and keep original spectral information, then the generation net of CBS-GAN generates the spatial-spectral data blocks by selected bands for sample generation. The experiments of classification are also used to demonstrate the availability of virtual samples generated by our method.

1. Introduction

The labeled data of hyperspectral images is often scarce, causing the phenomenon of inadequate training in hyperspectral deep learning methods. It is because the labeling process of hyperspectral image whose categories are represented by the pixels is more exhausting than that of other images[1]. With the deep learning methods widely used in hyperspectral image filed, sample generation which generates virtual samples for training process of deep learning model has been playing a more decisive role.

The generative adversarial network (GAN)[2] is an effective method in the field of sample generation. The core idea of GAN is the adversarial learning which can generate better virtual samples than other methods. Adversarial learning means the generator G of GAN and the discriminator D of GAN confront each other during the training process, improving both the models of G and D continually. At last, the model can generate virtual samples with extremely high similarity to real samples. In addition, conditional generative adversarial nets (CGAN)[3] improves the structure of GAN so that it can generate under the condition of label information which is more suitable for hyperspectral sample generation.
In order to expand the training samples for hyperspectral classification, some GAN-based models were proposed. Zhu et al. [4] proposed 3D-GAN for hyperspectral image classification and used the generated samples whose band dimensions were reduced by principal component analysis (PCA) for data augment, in this case, the original spectral information was broken by PCA. Xu et al. [5] proposed SpecGAN which was based on CGAN to generate virtual spectrum of hyperspectral images for the first time, and the comparison of classification results with generated samples was used to demonstrate the ability of their virtual spectrum. However, it cannot generate the spatial information of hyperspectral images. Wang et al. [6] proposed TripleGAN for training data generation to improve the classification model CapsNet. However, the spatial information was incomplete which was firstly reduced by PCA then turned into a 1-D spatial vector. Liu et al. [7] proposed C2GAN to generate complete spatial-spectral samples for the first time. The spatial information and the spectral information were generated by the stage one and stage two of C2GAN respectively. Although the complete spatial-spectral data blocks were generated, the high dimension of spectrum led to redundant generation which increased time complexity of training process and extracted redundant features for classification. Thus it can be seen that the problems of the lack of spatial information, the damage of the original spectral components, and the redundant generation in hyperspectral sample generation field need to be improved effectively.

In this paper, a novel conditional band selection generative adversarial net, named CBS-GAN is proposed for hyperspectral sample generation. In CBS-GAN, the band selection net, generation net and discrimination net are designed to select efficient bands, generate virtual samples and improve the ability of generation, respectively. The main contributions of this paper are summarized as follows.

1. To solve the problems of the redundant generation and the damage of the original spectral components, the model of band selection is used to reduce the redundant bands and keep the original spectral information for hyperspectral sample generation.
2. To keep the original spatial information, the generation module of CBS-GAN is adjusted with suitable size of convolutional layers to generate hyperspectral virtual data blocks with spectral-spatial information.

2. Proposed method
The main framework of the proposed CBS-GAN is shown in Figure 1, which consists of the band selection net and the generation net (Generator $G$ and Discriminator $D$). The band selection net is used to select informative band subset with less redundancy for sample generation. The generation net is used to generate virtual data blocks with spatial information and spectral information after band selection, in addition, the generation net can generate corresponding samples according to the inputs of labels.

2.1. The band selection net
The band selection net aims to reduce redundant bands without breaking up the spectral components. In order to gain better informative band subset, a unified band selection framework, Band Selection...
Network (BS-Net) [8] is applied to the band selection net of CBS-GAN. The framework consists of a band attention module (BAM), which aims to explicitly model the nonlinear interdependence between spectral bands, and a reconstruction network (RecNet), which is used to restore the original hyperspectral image from the learned informative bands, resulting in a flexible architecture. The resulting framework is end-to-end trainable, making it easier to train from scratch and to combine with many existing networks. In this way, k bands are selected as informative band subset.

The band subset selected by the band selection net is used to create the training dataset of the generation net. The other bands are eliminated from hyperspectral data except the bands in the subset, in this way, the redundant bands are tremendously reduced to decrease time cost. The remaining bands equip higher characterization ability of hyperspectral data for classification, so they are more valuable to be generated. After band selection, each pixel circles a window of (32, 32) as the spatial information. In this way, each training sample \( x \) of spatial-spectral data block in the shape of (32, 32, k) is created, labeled as the corresponding label \( c \). The training dataset is composed of the all training samples, which is the combination of data blocks \( x \) and labels \( c \).

### 2.2. The generation net

The generation net which consists of generator \( G \) and discriminator \( D \) is used to generate the virtual spatial-spectral data blocks \( G(z, c) \) with the training dataset mentioned above. The generator \( G \) is composed of the conditional deconvolutional structure which deconvolves the combination of Gaussian noise \( z \) and input label \( c \) into deep features and 3 deconvolutional layers which deconvolve the features into \( G(z, c) \) of required size. The discriminator \( D \) consists of the conditional convolutional structure which convolves the combination of data blocks \( x \) or \( G(z, c) \) and input label \( c \) into deep features and 3 convolutional layers which convolve the features into discrimination for model training.

The process of adversarial learning to generate virtual sample is described as follows. Firstly, the generator \( G \) deconvolves the combination of Gaussian noise \( z \) and input label \( c \) into a virtual sample \( G(z, c) \), but the virtual sample is useless without training in the beginning. Secondly, the discriminator \( D \) discriminates whether the training sample \( x \) is the real object, and the gradient descent method is used to optimize the parameters of discriminator \( D \) to improve the discrimination ability of discriminator \( D \). At last, the discriminator \( D \) discriminates whether the virtual sample \( G(z, c) \) is the real object or fake object, as we all know, the generator \( G \) wants the virtual sample \( G(z, c) \) to be the real object and the discriminator \( D \) wants the virtual sample \( G(z, c) \) to be the fake object, so the parameters of generator \( G \) and discriminator \( D \) can be optimized by the loss function which is written as

\[
\min_G \max_D V(D, G) = \mathbb{E}_{x \sim Pr}[\log(D(x, c))] \\
+ \mathbb{E}_{z \sim Pg}[\log(1 - D(G(z, c)))]
\]

(1)

where \( Pr \) is the real sample distribution, and \( Pg \) is the sample distribution produced by the generator \( G \).
Figure 2 Visualization of generated samples from 9 labels in total on the Pavia University dataset, band 90, 42 and 16.

In this way, the generation ability of generator $G$ and the discrimination ability of discriminator $D$ can be improved by the process of adversarial learning. To relieve the trouble of model collapse during the training process, the spectral normalization[9] is utilized to optimize the model parameters during the convolutional process.

3. Experimental result

3.1. Dataset and experimental setting

In our study, the standard hyperspectral dataset, i.e., the Pavia University is used for the hyperspectral image sample generation experiment. The Pavia University dataset was gathered by the Reflective Optics System Imaging Spectrometer (ROSIS-3) over the city of Pavia, northern Italy. The original dataset has 115 bands in the range of 0.43-0.86 μm with the size of 610×340 pixels, which has high spatial resolution of 1.3 m/pixel. In the experiment, 12 noisy bands have been removed and the remaining 103 channels were used for generation. There are a total of 9 labels of ground truth.

The experimental settings of the proposed CBS-GAN are described as follows. In the band selection net of CBS-GAN, the hyperparameters of $\lambda$, $\eta$ and $\text{maxiter}$ which control the ability of band selection are set as 0.01, 0.002 and 100 respectively. In the initialization of the generation net, the size of Gaussian noise is set as 100, and input labels need to be scattered into the shape of (1, 9) as there are 9 classes in the Pavia University dataset. To facilitate network training, the training data is scaled to the range of (0, 1). In the training process of the generation net, the number of epochs is set as 300, and the optimizers of $G$ and $D$ are both RMSprop with the learning rate of 0.0002 which decays per 100 epochs with the gamma of 0.1. After the spatial-spectral information generation, all virtual samples are scaled to the range of (0, Max) as real hyperspectral image samples.

3.2. Visualization of generated sample

In order to better display the significant samples generated by the proposed CBS-GAN model, we provide the visualization experiment to show the virtual samples of spatial-spectral information.

The visualization of the fake image generated by CBS-GAN is presented in Figure 2. Firstly, the band selection net selects the informative band subset [90, 42, 16, 48, 71, 3, 78, 38, 80, 53] which are ranked according to their expression ability. Secondly, create the training dataset with the window of
(32,32) for spatial information and selected band subset for spectral information to train the generation net of CBS-GAN. Then, after 300 epochs of adversarial training of $G$ and $D$, the generator $G$ which has the ability to independently generate realistic samples, is taken out separately to be fed into required label to generate the spatial-spectral information sample fixed with the label style. Figure 2 shows the spatial-spectral information samples after band selection from label 1 to label 9 that equip the ability to show the spatial-spectral information of each label on the Pavia University dataset. It needs to be supplemented that to create RGB image for visualization, the RGB channels are corresponding to generated bands [90, 42, 16].

In general, the visualization experiments demonstrate that the fake samples we generated equip obvious spatial-spectral features of each label. the CBS-GAN that we proposed is demonstrated to be competent.

### 3.3. Verification of generated sample

To verify the similarity of the generated sample, we use the classification network for comparative experiments. We use CNN and ResNet for classification and use CBS-GAN for data augmentation to improve the training process of CNN and ResNet.

| Method          | Trainset-1%     | Trainset-0.5%  |
|-----------------|-----------------|----------------|
|                 | Overall Accuracy(%) | Kappa×100 | Overall Accuracy(%) | Kappa×100 |
| CNN             | 94.27±0.62       | 92.04±0.84    | 86.00±0.91       | 81.41±1.31 |
| CBS-GAN + CNN   | 95.16±0.25       | 93.46±0.20    | 88.57±1.25       | 84.81±1.67 |
| ResNet          | 92.78±1.36       | 90.46±1.79    | 79.33±1.38       | 72.66±2.14 |
| CBS-GAN + ResNet| 93.25±0.96       | 91.10±1.24    | 83.64±2.58       | 78.44±3.26 |

The CNN model consists of four convolutional layers with batch normalization and ReLU, and the ResNet model is ResNet18[10] adapted to hyperspectral images. The learning rate and epochs of both models are set as 0.0002 and 500, respectively. The optimizer of both models is Adam optimizer. The training data Trainset-1% and Trainset-0.5% are created by randomly extracting 1% or 0.5% samples from each class in the original samples, the rest of the samples are all integrated as the test datasets where the classification results are gained. In addition, CBS-GAN generates 5 and 3 samples per class for data augmentation on Trainset-1% and Trainset-0.5% respectively. At last, the verification results on the Pavia University dataset are shown in Table 1.

From Table 1, we can see that the united models with data augmentation gain higher overall accuracy and kappa than the original CNN and ResNet methods on the Pavia University dataset, which demonstrates the realistic generation ability of the proposed CBS-GAN method. Also we can see from Table 1 that the complex model like ResNet performs worse in the situation of the lack of training dataset, and the virtual samples generated by CBS-GAN for data augmentation are effective for improving the training process of the classification model.

### 4. Conclusion

In this paper, a significant method CBS-GAN is proposed for the generation of hyperspectral samples with spatial-spectral information. In CBS-GAN, the model of band selection and the model of generative adversarial net are combined for the first time to generate effective virtual samples. In this way, the virtual samples generated by CBS-GAN are equipped with both the undamaged spatial and spectral information, and can be generated as the designated label by inputting the label information. In addition, the redundant bands are eliminated by the band selection net so that the virtual samples can be more suitable for sample expansion in hyperspectral classification. In addition, the convolutional layers are set with the spectral normalization to optimize model parameters, reducing the probability of model
collapse. Experimental results including visualization and verification on the Pavia University dataset demonstrate the realistic generation ability of the proposed CBS-GAN method. Furthermore, there is an existing problem for our generated samples: the existing noise in the original samples leads to the visible noise existing in the sample image generated by the model of CBS-GAN. In the future, an innovative denoising algorithm will be reasonably considered in CBS-GAN to generate more precise and clear samples with less noise.

Acknowledgments
This work was supported by the National Nature Science Foundation of China (grant numbers 61973285, 62076226, 61873249, and 61773355) and the Open Research Project of the Hubei Key Laboratory of Intelligent Geo-Information Processing (grant number KLIGIP-2019A04).

References
[1] Li S.T., Song W.W., Fang L.Y. (2019) Deep learning for hyperspectral image classification: an overview. IEEE Transactions On Geoscience and Remote Sensing., 57:6690-6709.
[2] Goodfellow I., Pouget-Abadie J., Mirza M., et al. (2014) Generative adversarial nets. In: Advances in neural information processing systems. Montreal. 2672-2680.
[3] Mirza M., Osindero S. (2014) Conditional generative adversarial nets. https://arxiv.org/abs/1411.1784v1.
[4] Zhu L., Chen Y.S., Ghamisi P., et al. (2018) Generative adversarial networks for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing., 56: 5046-5063.
[5] Xu Y.H., Du B., Zhang L. (2018) Can we generate good samples for hyperspectral classification? a generative adversarial network based method. In: 2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia. 5752-5755.
[6] Wang X., Tan K., Du Q., et al. (2019) Caps-Triple GAN: GAN-Assisted CapsNet for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing., 57: 7232-7245.
[7] Liu X.B., Qiao Y.L., Xiong Y.H., et al. (2020) Cascade conditional generative adversarial nets for spatial-spectral hyperspectral sample generation. Science China Information Science., 63(4): 140306.
[8] Cai Y.M., Liu X.B., Cai Z.H. (2019) BS-Nets: An End-to-End framework for band selection of hyperspectral image. IEEE Transactions on Geoscience and Remote Sensing., 58(3): 1969-1984.
[9] Miyato T., Kataoka T., Koyama M., et al. (2018) Spectral normalization for generative adversarial networks. https://arxiv.org/abs/1802.05957v1.
[10] He K.M., Zhang X.Y., Ren S.Q., et al. (2016) Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas. 770-778.