This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Using Adversarial Network for Multiple Change Detection in Bitemporal Remote Sensing Imagery

Wenzhi Zhao, Xi Chen, Xiaoshan Ge, and Jiage Chen

***Abstract*— Change detection by comparing two bitemporal images is one of the most challenging tasks in remote sensing. At present, most related studies focus on change area detection while neglecting multiple change type identification. In this letter, an attention gates generative adversarial adaptation network (AG-GAAN) is proposed on multiple change detection. The AG-GAAN has the following contributions: 1) this method can automatically detect multiple changes; 2) it includes attention gates mechanism for spatial constraint and accelerates change area identification with finer contours; and 3) the domain similarity loss is introduced to improve the discriminability of the model so that the model can map out real changes more accurately. To demonstrate the robustness of this approach, we used the Google Earth data sets that include seasonal variations for change detection and understanding. The experimental results demonstrated that the proposed method can accurately detect the multiple change types from bitemporal imagery.**

***Index Terms*— Attention gates (AGs), bitemporal images, domain similarity loss, generative adversarial network (GAN), multiple-change detection.**

I. Introduction

CHANGE detection is one of the most important directions within the field of remote sensing. It aims to analyze and quantify changes in remote sensing images at the same geographical location over different periods [1]. Because of its unique characteristics, change detection has been widely used in vegetation monitoring, urban expansion, and disaster assessment [2]. With the development of remote sensing platforms and sensors, a huge amount of repeatable remote sensing imagery was acquired. Compared with dense time-series moderate spatial resolution imagery, the very-high-resolution (VHR) images have scarce revisit data but with much more detailed land cover information. With the increased spatial resolution, the intraclass variation also increases, which means that it is more difficult to obtain accurate detection results for conventional postclassification change detection methods [3], [4]. Besides, atmospheric interferences and illumination variation still pose challenges to extract useful information from the bitemporal image pair.

Over the past few decades, intensive studies have been devoted to identifying changes within image pairs. Among them, the most commonly used method is machine learning, such as support vector machine (SVM), change vector analysis (CVA) [5], and principal component analysis (PCA) [6]. However, the remote sensing binary images affected by solar illumination, seasonal differences, and intraclass changes will suppress the real change information, and it is difficult for machine learning methods to distinguish them. The rise of deep learning algorithm provides a new solution for remote sensing image change detection. Deep learning methods have natural advantages in applying change detection because it can identify semantically rich high-level features in images through hierarchical architecture [7], [8]. For instance, the fully convolutional network (FCN) uses multiple convolutional layers to formulate the transformation relationship between image pairs. However, FCN ignores the spatial relationship between pixels and is not sensitive to the targets' detailed information. Thus, the FCN-CRF model is proposed to solve the problem of spatial constraint. FCN-CRF uses a conditional random field (CRF) to extract context information to strengthen further the features extracted by FCN. Although FCN-CRF introduces extra spatial constraints to improve classification accuracy, it inevitably increases the computational complexity [9]. The U-Net model combines encoding–decoding structure with the jumping network; it increases the robustness in spatial feature extraction through multilayer upsampling and downsampling processes [10]. However, FCN, FCN-CRF, and U-Net are discriminative networks and need a large number of accurately labeled training samples. The change detection between VHR image pairs is often unable to meet this requirement.

Compared with the discriminative models, the generative adversarial network (GAN) is an unsupervised/semisupervised generative network that only requires a relatively small number of training samples [11], [12]. At present, GAN has been widely used in the field of remote sensing [13]. Ma *et al.* [14] designed a dense residual network in the GAN network to extract deep image features for improving the spatial resolution of remote sensing images. Lebedev *et al.* [15] conducted a detailed study on the change detection task by using pix2pix GAN, proving the usability of the pix2pix framework on the change detection task. Combined by the GAN and metric learning, Zhao *et al.* [16] proposed a MeGAN model, which

Manuscript received September 1, 2020; revised October 19, 2020 and October 28, 2020; accepted November 1, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFC1508903 and in part by the Fundamental Research Funds for the Central Universities under Grant 2018NTST01. *(Corresponding author: Xi Chen.)*

The authors are with the State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Institute of Remote Sensing Science and Engineering, Beijing Normal University, Beijing 100875, China, also with the Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Institute of Remote Sensing Science and Engineering, Beijing Normal University, Beijing 100875, China, also with the School of Surveying and Land Information Engineering, Henan Polytechnic University, Henan 454000, China, and also with the National Geomatics Center of China, Beijing 100830, China (e-mail: 211804020027@home.hpu.edu.cn).

Color versions of one or more of the figures in this letter are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2020.3035780

1545-598X © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Beijing Normal University. Downloaded on June 17,2021 at 05:58:27 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

successfully solved the problem of insufficient detection of regional edge details caused by a seasonal variation for image pairs. Experimental results have shown that MeGAN is more robust than conventional models. However, GAN networks still need multilayer convolution operation to extract the region of interest areas. This strategy often results in redundant calculations due to repeated feature extraction, and it is often difficult to locate change areas with accurate boundaries. Moreover, it is not enough to focus on change area detection, and the need to analyze the type of changes (multiple-change detection) is much more urgent. Liu *et al.* [5] added the interaction between pixels and their adjacent areas in the compressed change vector analysis (C^2VA) method and adopted the multiscale ensemble strategy to detect the various types. Saha *et al.* [17] construct a siamese framework and use the same pretrained CNN for semantic segmentation of bitemporal images and then conduct a per-pixel analysis to identify variation types. Saha *et al.* [17] used CycleGAN to conduct multiple-change detection, which proved the usability of GAN models in change detection tasks. However, GANs are easily to collapse due to the imbalance of samples.

In general, we find that the main difficulties for multiple-change detection are as follows: 1) how to accurately formulate the multitype feature transitions and capture the exact contours of change areas are quite challenging and 2) how to solve the problem of insufficient training or overfitting caused by unbalanced samples. To solve the above problems, we propose the attention gates' generative adversarial adaptation network (AG-GAAN) for multiple-change detection. Inspired by recent advances in attention mechanisms, we introduced the attention gates (AGs) [18] for better feature formulation and contour delineation of multiple-change detection. The AG we used is a modular mechanism that dynamically and implicitly generates the proposal regions of the convolution framework without any additional computation. Experiment results demonstrated that AGs could improve the ability to capture the target's robust transition features and make the detection results more accurate. Meanwhile, domain similarity loss is added to the proposed model to enhance the discriminator's discriminability, which helps GAN reach the Nash equilibrium [19].

The main contributions of this study are as follows: 1) an adversarial network-based multiple-change detection model is proposed, and it can automatically detect the changes according to feature transition patterns and 2) the AGs and domain similarity loss are introduced for change detection, which increases the ability to capture the targets' details and improve the detection accuracy. The remainder of this article is organized as follows. Section II introduces the algorithm of change type detection in detail. Experimental data sets, network settings, and experimental results are presented in Section III. Section IV provides conclusions.

II. Methodology

A. Generative Adversarial Networks

GAN is one of the most representative generative models and is widely used in unsupervised/semisupervised learning. The conventional GAN consists of a generator ($G(z)$) and a discriminator ($D(x)$). In the training process, they are in the process of mutual confrontation. The generator accepts random noise z and constantly upgrades itself to produce samples p_Z in order to deceive the discriminator, and the discriminator constantly updates itself to recognize the samples whether true or false. This process can be expressed as follows:

$$\min_G \max_D V(D,G) = E_{x\sim p_d}[\log D(x)] + E_{z\sim p_z}[\log(1-D(G(z)))]. \quad (1)$$

In this formula, p_d represents the real data distribution from real data x, and z represents the input item of the generator. The parameters θ_g and θ_d of generator and discriminator are represented by G and D. During the training process, the parameters of the discriminator are updated once, and then, the parameters of the generator are also updated once, until the convergence of the loss function. Compared with traditional GAN, the input of CGAN is no longer a random noise z but a picture and a control condition y. In addition, CGAN adds L1 loss to the loss function to make the images of the source domain and target domain as close as possible. Pix2pix GAN is further improved by adding a skipping network structure into the generator, making the GAN model more suitable for the Image2Image translation task. The skipping network structure shares the high- and low-level semantic information so that the transformed image has better detail performance.

B. AGs Generative Adversarial Adaptation Network

The general workflow of AG-GAAN is shown in Fig. 1. We randomly crop the original image pairs into 256×256 patches and then input them to the generator to generate predictive change maps that can deceive the discriminator. The discriminator distinguishes the change maps coming from the generator or real labels. In the AG-GAAN model, AGs are added to enhance the robustness of the GAN model in determining sensitive features. AGs utilize a modular mechanism, which can adjust the number of layers according to requirements. It is worth noting that the AGs are a spatial constraint mechanism that gradually locates change areas during the training process and enhances the discrimination ability on change types. At the same time, domain similarity loss is added to improve the stability of GAN model performances and ensure the generated maps and reference ones to be similar. For the generator, we replaced the original skip network structure with AGs, as shown in Fig. 1(a). AGs take a shallow network as a gating vector k_i, determine the change area of each layer i from the deep network, and remove the corresponding low-level feature responses in the deep network q_i. To formulate this process

$$\alpha_i = \sigma\left(W_v^T\left(W_q^T q_i + W_k^T k_i + b_{q,k}\right) + b_v\right) \quad (2)$$

where $\sigma(x) = (1/(1+\exp(x)))$ represents the sigmoid activation function. $W^T \in R^{\hat{C}\times C}$ stands for channelwise $1\times 1\times 1$ convolutions for the linear transformations of the input tensor, and b is the bias term. Attention coefficients $\alpha_i \in [0,1]$ is used to focus on the region of change areas. For each AG, the activation can be represented as

$$\text{attention}_i = \alpha_i \times k_i. \quad (3)$$

Authorized licensed use limited to: Beijing Normal University. Downloaded on June 17,2021 at 05:58:27 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

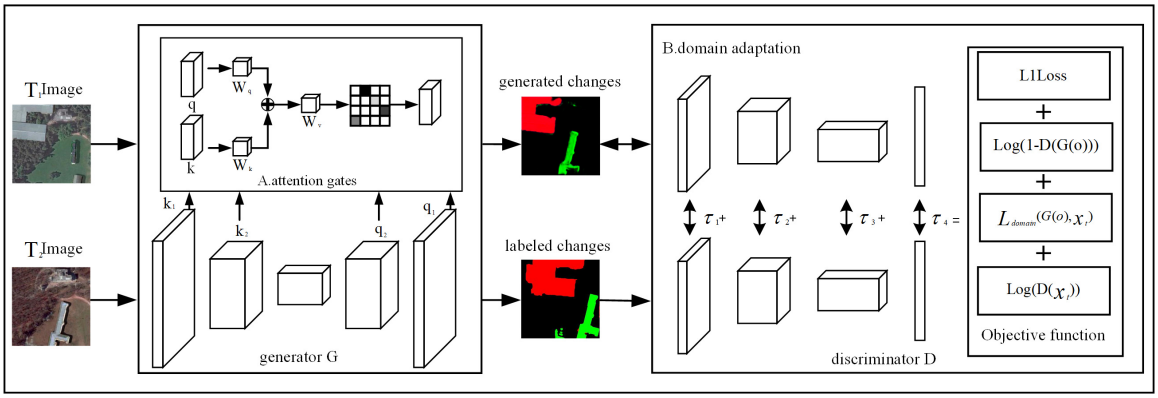

Fig. 1. General workflow of AG-GAAN.

Once AGs are activated, gradients originating in the background area are determined during the backward process. In the remote sensing change detection task, the changed objects often have different sizes and colors. It is not enough to rely solely on a discriminator to determine whether a region has changed. Therefore, domain similarity loss is added to enhance the discriminator's ability, as shown in Fig. 1(b). Especially, the maximum mean discrepancy (MMD) is added to the discriminator to measure the distance between the predicted result and the reference map

$$L_{\text{domain}}\left(G(o), x_t\right) = \sum_{n=1}^{N} \| D_n(x_t) - D_n(G(o)) \|^2. \tag{4}$$

x_t and $G(o)$ are change labels and prediction results by unlabeled images respectively, and n represents the nth layer of the discriminator. We use τ_n to represent the nth layer computations $\| D_n(x_t) - D_n(G(o)) \|^2$ in the formula. The domain similarity can also be written as $L_{\text{domain}}\left(G(o), x_t\right) = \tau_1 + \tau_2 + \cdots + \tau_n$. The objective function of the discriminator can be expressed as

$$\begin{aligned} &\max_D L_{\text{dom}} \\ &\quad = E_{x,x'}\left[k\left(x, x'\right)\right] + 2E_{x,y}[k(x, y)] + E_{y,y'}\left[k\left(y, y'\right)\right] \end{aligned} \tag{5}$$

where x and x' represent the change map predicted by two random image pairs, and y and y' are their corresponding labels. $k(*)$ is the kernel function that measures the similarity between x and y. In order to enhance the discriminant ability, the discriminator maximizes $E_{x,y}$ to make the change map generated by the generator to be more realistic. Meanwhile, the discriminator minimizes intraclass variance through $E_{xx'}$ and $E_{yy'}$. Similarly, the objective function of the generator can be expressed as

$$\begin{aligned} &\min_G L_{\text{dom}} \\ &\quad = E_{x,x'}\left[k\left(x, x'\right)\right] - 2E_{x,y}[k(x, y)] + E_{y,y'}\left[k\left(y, y'\right)\right]. \end{aligned} \tag{6}$$

Therefore, the objective function of the AG-GAAN for multiple change detection has the following formulation:

$$\begin{aligned} &\min_G \max_D V(D, G) \\ &\quad = E_{x\sim p_d}\left[\log D(x_t)\right] + E_{z\sim p_z}[\log(1 - D(G(o)))] \\ &\qquad + L_{\text{domain}}(G(o), x_t) + L1. \end{aligned} \tag{7}$$

Fig. 2. Google Earth bitemporal data set with seasonal changes.

TABLE I

DETAILED CONFIGURATION ABOUT THE AG-GAAN G AND D, RESPECTIVELY, REPRESENT THE GENERATORS AND DISCRIMINATORS; q, k, AND v ARE THE CONVOLUTION LAYER OF AGS, WHICH ARE W_q, W_k, AND W_v

Name	Layer	Kernel	Stride	Features	Activation
G	1,16	4×4	2×2	20	ReLu
	2,15	4×4	2×2	40	ReLu
	3,14	4×4	2×2	80	ReLu
	4-13	4×4	2×2	160	ReLu
	17	4×4	2×2	3	tanh
	q	1×1	1×1	512	ReLu
	k	1×1	1×1	512	ReLu
	v	1×1	1×1	1	Sigmoid
D	1	4×4	2×2	20	ReLu
	2	4×4	2×2	40	ReLu
	3	4×4	2×2	80	ReLu
	4	4×4	2×2	160	ReLu
	5	4×4	1×1	1	Sigmoid

The first two losses are the optimization terms of GANs, and the second is the additional optimization of domain similarity loss. The last term is L1 loss $= \sum |x_t - G(o)|$. Finally, loss terms are calculated and minimized by Adam optimizer during the training process.

III. EXPERIMENTS AND RESULTS

A. Data Description

We use a bitemporal data set obtained from Google Earth to demonstrate the detection capability of our proposed model, which is provided by Lebedev *et al.* [15]. The data set has three bands that describe two different seasonal variations

Authorized licensed use limited to: Beijing Normal University. Downloaded on June 17,2021 at 05:58:27 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE II
MULTIPLE-CHANGE DETECTION ACCURACY NODSLOSS-GAN:GAN ONLY ADDS AGS; NO AG-GAN:GAN ONLY ADDS DOMAIN SIMILARITY LOSS

Methods	SVM		CNN		Pix2pix		noDSloss-GAN		noAG-GAN		AG-GAAN		Count
–	precision	recall	precision	recall	precision	recall	precision	recall	precision	recall	precision	recall	–
C1	0.93	0.77	**0.95**	0.79	0.91	0.91	0.93	0.90	0.92	**0.93**	0.93	0.92	836869
C2	0.30	0.67	0.25	0.64	0.44	0.61	0.43	0.62	0.40	**0.70**	**0.47**	0.65	34547
C3	0.16	0.37	0.28	**0.57**	0.20	0.17	0.26	0.30	0.33	0.21	**0.36**	0.32	78584
OA	0.73		0.77		0.84		0.84		**0.86**		**0.86**		–
Kappa	0.21		0.30		0.26		0.32		0.33		**0.38**		–

with a spatial resolution of 0.3 m, and the sizes of the image are 2700 × 4275 pixels. As shown in Fig. 2, there are a large number of targets that have been "increased" and "decreased" within the image pairs that allow us to test the ability of the model for multiple-change detection. Although high spatial resolution images allow us to detect small changes, still images over different seasons produce a large number of pseudochange that interferes with detection results. We define the emerged targets in the second image relative to the first one as "increased," and conversely, it is defined as "decreased." This letter tests the multiple changes of buildings, including unchanged buildings, increased buildings, and reduced buildings, which are expressed as C1–C3, respectively.

In order to verify the performance of the proposed method, we use SVM, CNN, and pix2pix GAN as the comparison algorithms. It is worth noting that pix2pix GAN has a similar structure as the proposed method, and the specific configuration is shown in Table I. At the same time, in order to prove the effectiveness of AGs or domain similarity loss, we tested the model that only uses AGs and domain similarity loss for comparison. To quantitatively illustrate the accuracy of change detection, overall accuracy (OA), recall rate, and kappa coefficient were added for accuracy evaluation.

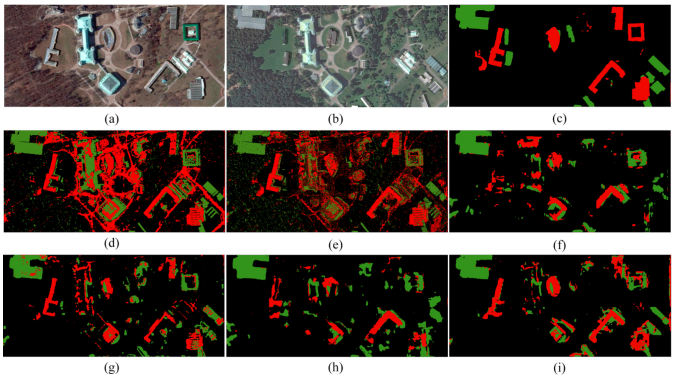

Fig. 3. Comparison on multiple-change detection results. (a) Image 1. (b) Image 2. (c) Reference map. (d) SVM. (e) CNN. (f) Pix2pix. (g) GAN only adds AGs. (h) GAN only adds domain similarity loss. (i) AG-GAAN. Red refers to decreased objects, and green refers to increased objects.

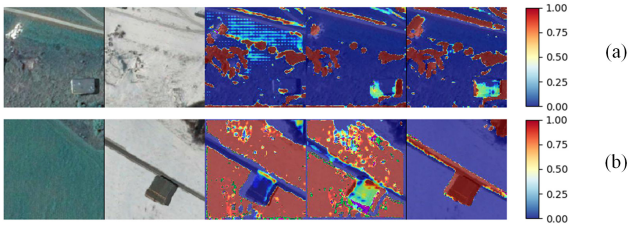

Fig. 4. Spatial awareness maps of the proposed AGs. (a) C3: 50, 2000, and 2500 and (b) C2: 50, 500, and 1000 of the coefficient of attention. During the training process, the model gradually locates the increase and decrease areas.

B. Configuration and Analysis

In the experiment, we randomly cropped the original bitemporal data set into small patches with sizes of 256 × 256. We took 50% of the sample data set, which is 1:500 rows and 1:1900 columns of the original image, as the training data. Especially, the numbers of "increased," " decreased," and background samples in the training set were 19013, 14468, and 916519 pixels. The remaining samples were used for testing. In order to make full use of the spectral information of the image, the six bands of the original bitemporal data were stacked as input data for change detection.

C. Results and Comparison

Fig. 3(a)–(c) shows the prechange image, postchange image, and manual annotation samples of the Google Earth data set, respectively. Fig. 3(d)–(f) demonstrates the detection results of different comparison methods, and Fig. 3(g) and (h) shows the effects of the two modules that we proposed. The evaluation values are listed in Table II. Although high-resolution images provide more detailed information, they are followed by more pseudochanges and random noise. With this complex background, traditional detection methods cannot identify real changes, such as SVM in Fig. 3(d). In Fig. 3(e), the CNN method can capture the multiple change features, but the single use of the convolutional network is not enough to accurately identify the "increase" and "decrease" regions, so the precision value of "increase" and "decrease" types reached only 0.25 and 0.28. In addition, salt-and-pepper noise is also a problem. In Fig. 3(f), pix2pix GAN is a typical generative model, which does not need a large number of training samples. It can automatically learn high-level robust features through the deep convolution network during the confrontational process between the discriminator and the generator. Compared with CNN, the detection effect of pix2pix is smoother, with the overall detection accuracy value improved by nearly 0.1, but it is still difficult to find multiple changes. Fig. 3(g) shows the AG-GAAN detection results only with AGs, and it can be seen that the changed building has a more complete contour. We also present the AGs' spatial awareness map in Fig. 4. During the training process, AGs gradually update and locate the boundary of the change targets. In addition, we added domain similarity loss to our model. Fig. 3(h) shows the AG-GAAN detection results with only domain similarity loss; the model can effectively reduce the seasonal noise through the measurement between different domains. Therefore, it has

Authorized licensed use limited to: Beijing Normal University. Downloaded on June 17,2021 at 05:58:27 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

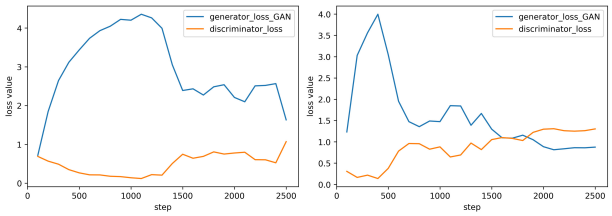

Fig. 5. Curves of loss functions with 2500 iterations. (a) AG-GAAN. (b) Pix2pix.

a relatively high recall value of 0.93, 0.70, and 0.21. The final detection results of the proposed method are shown in Fig. 3(i). The proposed method combines the advantages of AGs and domain similarity loss and has a better detection result for each type of change. Therefore, the recall rate of each type has a good performance, and the average precision and kappa are the highest among all methods. The loss curves of pix2pix GAN and AG-GAAN are shown in Fig. 5, which can demonstrate the efficiency and effectiveness of the proposed method.

IV. Conclusion

We propose an AG-GAAN model for multiple-change detection with high-resolution bitemporal images. Compared with conventional multiple change detection, our model focuses on the transformation of the same objects in different directions, such as increase and decrease. Understanding the increase and decrease of the same objects makes their changes easier to explain, especially in disaster assessment. Especially, we add the domain similarity loss to the pix2pix GAN model to improve the discriminator's ability and help the model to achieve the Nash equilibrium. At the same time, we added AGs to gradually locate the change areas during the training and suppress background interference. Experiments demonstrated that our model can significantly improve the detection rate of multiple change detection and understand the change for ground targets. However, the model is greatly challenged by the hazardous environments, such as snow cover, which needs to be improved. In the future, we still need to pay attention to improve the robustness recognition of target changes between bitemporal images.

References

[1] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2401, 2004.

[2] D. M. Browning and C. M. Steele, "Vegetation index differencing for broad-scale assessment of productivity under prolonged drought and sequential high rainfall conditions," *Remote Sens.*, vol. 5, no. 1, pp. 327–341, Jan. 2013.

[3] F. Luo, L. Zhang, B. Du, and L. Zhang, "Dimensionality reduction with enhanced hybrid-graph discriminant learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5336–5353, Aug. 2020.

[4] F. Luo, L. Zhang, X. Zhou, T. Guo, Y. Cheng, and T. Yin, "Sparse-adaptive hypergraph discriminant analysis for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1082–1086, Jun. 2020.

[5] S. Liu, Q. Du, X. Tong, A. Samat, L. Bruzzone, and F. Bovolo, "Multiscale morphological compressed change vector analysis for unsupervised multiple change detection," *IEEE J. Sel. Topics Appl. Earth Observ., Remote Sens.*, vol. 10, no. 9, pp. 4124–4137, Sep. 2017.

[6] N. Neeti and J. R. Eastman, "Novel approaches in extended principal component analysis to compare spatio-temporal patterns among multiple image time series," *Remote Sens. Environ.*, vol. 148, pp. 84–96, May 2014.

[7] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[8] M. Lan, Y. Zhang, L. Zhang, and B. Du, "Global context based automatic road segmentation via dilated convolutional neural network," *Inf. Sci.*, vol. 535, pp. 156–171, Oct. 2020.

[9] H. Zhou, J. Zhang, J. Lei, S. Li, and D. Tu, "Image semantic segmentation based on FCN-CRF model," in *Proc. Int. Conf. Image, Vis. Comput. (ICIVC)*, Aug. 2016, pp. 9–14.

[10] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018.

[11] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[12] R. Hang, F. Zhou, Q. Liu, and P. Ghamisi, "Classification of hyperspectral images via multitask generative adversarial networks," *IEEE Trans. Geosci. Remote Sens.*, early access, Jun. 25, 2020, doi: 10.1109/TGRS.2020.3003341.

[13] R. Hang, Z. Li, P. Ghamisi, D. Hong, G. Xia, and Q. Liu, "Classification of hyperspectral and LiDAR data using coupled CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4939–4950, Jul. 2020.

[14] W. Ma, Z. Pan, F. Yuan, and B. Lei, "Super-resolution of remote sensing images via a dense residual generative adversarial network," *Remote Sens.*, vol. 11, no. 21, p. 2578, Nov. 2019.

[15] M. A. Lebedev, Y. V. Vizilter, O. V. Vygolov, V. A. Knyaz, and A. Y. Rubis, "Change detection in remote sensing images using conditional adversarial networks." *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 42, no. 2, pp. 565–571, 2018.

[16] W. Zhao, L. Mou, J. Chen, Y. Bo, and W. J. Emery, "Incorporating metric learning and adversarial network for seasonal invariant change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2720–2731, Apr. 2020.

[17] S. Saha, F. Bovolo, and L. Bruzzone, "Unsupervised deep change vector analysis for multiple-change detection in VHR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3677–3693, Jun. 2019.

[18] R. Hang, Z. Li, Q. Liu, P. Ghamisi, and S. S. Bhattacharyya, "Hyperspectral image classification with attention-aided CNNs," *IEEE Trans. Geosci. Remote Sens.*, early access, Jul. 16, 2020, doi: 10.1109/TGRS.2020.3007921.

[19] W. Zhao, X. Chen, Y. Bo, and J. Chen, "Semisupervised hyperspectral image classification with cluster-based conditional generative adversarial net," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 539–543, Mar. 2020.

Authorized licensed use limited to: Beijing Normal University. Downloaded on June 17,2021 at 05:58:27 UTC from IEEE Xplore. Restrictions apply.