# Survey on Remote Sensing Data Augmentation: Advances, Challenges, and Future Perspectives

Amel Oubara, Falin Wu, Abdenour Amamra and Gongliu Yang



April 2, 2022

# Survey on Remote Sensing Data Augmentation: Advances, Challenges, and Future Perspectives

Amel Oubara[1], Falin Wu[1], Abdenour Amamra[2], and Gongliu Yang[1]

[1] SNARS Laboratory, School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing, China
[2] Ecole Militaire Polytechnique, Bordj El-Bahri BP 17, Algiers, Algeria

**Abstract.** Deep learning-based methods have shown great progress in remote sensing applications. The performance of such methods can significantly outperform traditional remote sensing methods under the condition of the availability of large datasets for training. Unfortunately, some remote sensing tasks, such as the change detection task, lack large established datasets. This issue is due to the limited access to some remote sensing data and the absence of a sufficient labeled dataset. Data augmentation techniques are generally used to tackle this issue by increasing the number of samples and enhancing the quality of the training data. These techniques have shown performance improvement for general data and have recently been applied to remote sensing data. The present review synthesizes the recent data augmentation works contributed to the remote sensing field. It briefly describes data-level issues, existing data augmentation techniques used to address these issues, and challenges facing these techniques. This review provides the reader with an idea about the influence of data augmentation techniques on the performances of deep learning models, especially while using a small amount of data.

**Keywords:** Remote Sensing · Change Detection· Deep Learning · Data Augmentation · GANs.

## 1 Introduction

Remote Sensing (RS) science has known a great growth in recent years due to the development of different types of optical and radar sensors. Placed on satellites or airplanes, these sensors allow the collection and acquisition of a huge amount of remote sensing data. The efficient exploitation of this data has pushed researchers to apply Deep Learning (DL) methods to solve several remote sensing tasks. With the growth of remote sensing data and the need to make the use of this important new data stream effective, DL has recently been widly used to solve complex remote sensing tasks, such as classification, segmentation, target detection, and change detection. Therefore, several new approaches integrating DL techniques were developed to improve the accuracy and automation of these tasks. Since it is very huge to mention all the tasks that have been treated in remote sensing, we take the change detection task as an example and discuss the application of DL methods on it.

Considering the importance of change detection task in earth observation, several traditional and DL-based methods have been proposed to perform this task accurately. Compared to traditional hand-crafted feature-based algorithms, DL-based methods have shown superior efficiency in terms of feature extraction used to detect changes. In addition, due to the meaningful data representation and learning capabilities, DL techniques can model the relationship between the image object and its real-world geographical feature as closely as possible, which enables the production of a real-world change map accurately [19]. This change map can be generated to detect segmented changed or unchanged pixels (binary change detection) or to detect multiclass change detection (multiclass change map). Depending on whether the training data is labeled or not, DL techniques are divided into supervised and unsupervised learning algorithms. In each algorithm, several architectures have been developed based on different deep network models such as Deep Belief Networks (DBNs), Autoencoders (AEs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Generative Adversarial Networks (GANs) [11].

Despite the fact that DL-based methods have shown their effectiveness in different applications, the mentioned networks are deeply dependent on big data to avoid overfitting. Overfitting appears when the model is over-optimized for training data but performs poorly with new data. Unfortunately, many application domains do not have access to big and diverse data, such as remote sensing change detection task. According to the literature, there are two main issue-levels to be addressed in implementing remote sensing change detection task using deep learning methods. The first one concerns data-level issues (small labeled data and unbalanced change classes). The second level is related to feature extraction and feature discrimination issues. Since the emergence of DL-based methods in the change detection area, most research has focused on the second-level issues by trying to improve the feature extraction and discrimination power of the neural networks. However, less attention has been paid to tackling data-level issues.

It is known that the quantity and quality of training data have a great influence on the training phase of any deep learning network [4]. Thus, the large amount of high-quality training samples is indispensable to guarantee the accuracy and robustness of change detection task based on a deep learning approach. Seeing the importance and the influence of the dataset on the accuracy of DL-based change detection methods, some researchers focus on improving the reliability of the original data by solving some data-level issues to increase the reliability of DL-based change detection. This review resumes data augmentation techniques, the challenges faced while applying these techniques on remote sensing data, and the work done to address these challenges.

## 2 Data-level Issues

Whatever their application, the robustness of deep learning methods is often limited by the amount of data available for the training step. For remote sensing applications, this issue is more perceived especially for change detection task due to the rarity and sparsity of changes in the real world which makes it difficult to collect a great number of effective bitemporal images. In addition, limited labeled change map samples and open benchmarks are provided due to labor cost and time-consuming while labeling high-quality samples for training models [3].

To tackle these issues, transfer learning using pre-trained models was proposed but still faces the problem of model adaptation [11]. Semi-supervised and unsupervised learning with AEs and GANs were also used to reduce the dependence on ground truth [11]. The results of the proposed methods still need to be improved. These methods are used to tackle the small labeled data issue however they try to overpass this problem and not eliminate it.

Another data-level issue, which appears in the learning phase of a deep neural network, is the class imbalance phenomenon. This issue is more observed in CD task due to the low frequency of changes in RS images which creates the imbalance between changed and unchanged pixels. Usually, the targets of interest (in our case, targets belong to changes) only cover a much smaller number of pixels than the background [21]. Naive machine learning algorithms applied on such imbalanced data have a bias toward the no-change class and tend to ignore the change class. Several studies have been proposed to solve this issue such as adjusting the optimization objectives based on the weighted loss [1].

## 3 Data Augmentation Techniques

In order to solve the overfitting problem that appears in the learning phase due to insufficient data samples, functional solutions such as reducing the capacity of the network, adding weight regularization and including dropout and batch normalization, using transfer learning and pretrained models have been developed to improve the application of Deep Learning methods on smaller datasets. However, getting more training data is considered a powerful technique to mitigate the overfitting problem. Therefore, data augmentation has been proposed to artificially enhance and diversify the existing training dataset.

Data augmentation can be performed based on basic methods, imaging simulation system-based methods, or DL-based methods. Basic techniques rely mainly on image processing procedures such as geometric and color transformations, mixing images [12]. Imaging simulation system-based methods are used to generate simulated samples and combine them with real ones [16]. DL-based methods generate synthetic samples based on several techniques such as Variational Auto Encoding (VAE), adversarial learning, and generative adversarial networks (GANs) [12].

### 3.1 Basic Techniques

Basic data augmentation techniques are based on image processing tools that apply some transformations to an image to get a newly transformed image [12]. These transformations can be geometric such as rotation, scaling, translating, and flipping. It can be intensity transformation like grayscale and color transformation. Noise injection and filtering are popular data augmentation to get noisy and filtered new images, respectively. Mixing images is another data augmentation technique that combines two input images to get a new single image. Although these basic techniques have shown good performance for computer vision tasks, they are either geometric transformations or simply changing the pixel values of RGB channels which cannot improve the semantic information quality of remote sensing images in interpretation applications such as change detection task [9].

### 3.2 Imaging Simulation System Based Methods

With the evolution of imaging simulation systems, using simulated or virtual data has been heavily investigated in computer vision tasks. For instance, Richter *et al.* [10] produce simulated semantic segmentation datasets using a photorealistic open-world computer game. Li *et al.* [8] use CityEngine and Unity3-D to build a virtual image dataset used for traffic vision research in the large-scale urban street scene. These works use the simulated data directly to train the model. However, the difference between real and simulated images can affect the network model especially when the number of simulated data is greater than the real ones. The best solution to this issue is to try to produce as much as possible realistic simulated data. Aligning the style-level features (backgrounds, colors, and textures) between the simulated and real images is performed using Neural Style Transfer to reduce the gap domain between real and simulated world [16].

### 3.3 Deep Learning-based Methods

The DL-based augmentation approaches can automatically learn the representations of images and generate realistic images to increase the model's generalizability and reduce overfitting during training. This process of generating synthetic images is often quoted as image synthesis. The most commonly stated DL networks for data augmentation are the generative adversarial network (GAN) and its variants. Recently, GANs based methods are widely used in remote sensing change detection data generation and have shown accuracy improvement as presented in Section 4.

**Generative Adversarial Network (GAN)** was introduced, for the first time, in 2014 by Goodfellow *et al.*, which consists of a generator and a discriminator model. The generator ($G$) is a deep network that learns to produce novel data as negative training examples for the discriminator. The discriminator ($D$) is a network that learns to distinguish these fake generated data from original data. The generator is trained to fool the discriminator (to incorrectly classify fake generated data as real) and the discriminator is trained not to be fooled by the generator (correctly classify data as real or fake) [6]. The loss function which manages this model is formulated as below:

$$V(D,G) = \min_{G}\max_{D} \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_{Z}(z)}[1 - \log D(G(z))] \quad (1)$$

where $x$ is sampled from real data distribution $p_{data}(x)$, $z$ is sampled from the prior distribution $p_Z(z)$ such as uniform or Gaussian distribution, and $\mathbb{E}$ represents the expectation [6].

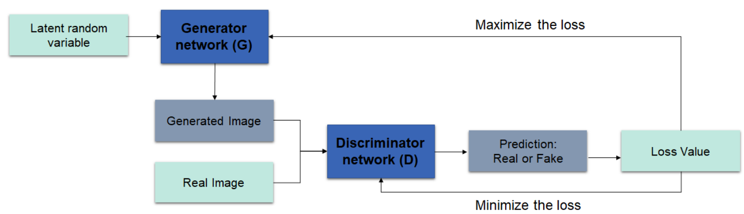


**Fig. 1.** Basic architecture of GAN.

The success of the generator to overcome the discriminator makes it very powerful for generative modeling. The impressive performance of GAN makes it the most promising generative modeling technique used in Data Augmentation. GAN was proposed as an unsupervised method to generate images and it has been improved to be applied for solving insufficient training datasets. The main challenges for GAN-based image generation are mainly related to conserving training stability and generating high-quality images. Several research works have been done to overcome these challenges by modifying the GAN architecture. These works have significantly improved the quality of samples created by GANs. Among these new architectures, we can list the Deep Convolutional GAN, Progressively Growing GANs, Cycle GANs, and conditional GANs.

The GAN architecture was applied to generate low-resolution images and revealed high performance in generating new images. However, it failed to show the same performance for high-resolution images. Deep Convolutional GAN [18] was created to enlarge the complexity of the generator and discriminator networks using convolutional layers in order to produce higher resolution images. Progressively Growing GAN [14] was then proposed to train a series of networks with progressive resolution complexity. Improving the quality of the generated images is as important as their resolution. Therefore, Cycle GAN [20] was proposed to enhance the quality of outputs by introducing an additional Cycle-Consistency loss function for image-to-image translation. Another new promising architecture was proposed to have more controllable generated images regarding the output images number, shape, and type. This architecture called conditional GAN [5] adds conditional vector to both the generator and discriminator to get more sophisticated produced data.

Compared to basic data augmentation techniques which try to raise the number of images by generating correlated image data, DL-based methods can produce new synthetic data and augment image data with different imaging modalities automatically such as changing the size, shape, color, and location of targets.

The gap domain between real and generated images that heavily appears while using imaging simulation systems is systematically reduced using DL-based methods. These methods are mainly based on GAN where the discriminator measures automatically the gap between fake and real images so that making the generator produces as real images as possible. However, the doubt of whether the generated synthetic images can represent realistic remote sensing features is still the main concern for DL-based augmentation since they are not from a real-world environment.

## 4 Remote Sensing Data Augmentation Challenges

The submentioned techniques and their development are mostly applied to general and natural images, few studies have adopted these techniques for remote sensing imagery. compared to general images, remote sensing images have more specific characteristics which need to be taken into consideration while constructing a generation model. Therefore, the existing findings on data generation cannot be applied directly to remote sensing data and need to satisfy some requirements that are resumed as follow.

- **Data labeling:** most remote sensing tasks such as interpretation, classification, segmentation, object detection, and change detection belong to supervised learning that needs a huge amount of labeled data. However, manual data labeling is time-consuming and labor cost due to the large size of remote sensing data. Therefore, it would be better to have a data generation network that can automatically generate data with their labels.
- **Data generation quality:** computer vision tasks such as segmentation, classification, and detection are more complex while performing remote sensing images than natural images. Therefore, high-quality image generation is required to maintain high accuracy while using remote sensing data. Moreover, some remote sensing image characteristics should be taken into consideration while building the generating model, such as spatial and spectral resolution, geometric and radiometric characteristics of the generated images. All these characteristics considerations can help to generate more realistic samples.
- **Time-consuming:** remote sensing images are generally big size data compared to usual images, so processing such data is time-consuming. This point should be involved while developing a data generation model which needs to be less computation complexity in order to reduce the generation time. The generation speed should be very high, especially in the case of models that require real-time prediction.

## 5 Remote Sensing Data Augmentation Works

The aforementioned data level issues are mainly caused by the low quantity and poor quality of remote sensing images. Researchers agreed that data augmentation is an effective technique to solve deep learning data-level problems. It can increase the number of data samples and satisfy the requirement of deep learning algorithms. Therefore, this technique has been applied to remote sensing images and has shown satisfactory results. In [17], the author proposed a new data augmentation technique for aircraft detection by creating simulated targets using aircraft three-dimensional models. This method has shown improvement compared to traditional data augmentation methods such as flipping and rotating. This technique brings new training samples to the deep neural network which helps to improve its detection performances.

The same idea of generating simulated targets has been exploited in [16] for the remote sensing ship classification task. In order to diversify the dataset, ship simulation samples were first generated through three-dimensional models of real images using a visible light imaging simulation system. Then, Neural Style Transfer (NST) network called Sim2RealNet was used to eliminate the gap domain between real and simulated images. The proposed method tackles the issues of poor diversity, blurring, and distortion. However, the quality of the generated images is poor and the amount of these generated data still leads to overfitting. Moreover, the time consuming while performing Sim2RealNet is too long.

Generating data samples has also been performed using GANs that have shown promising performance on the quality and diversity of generated images. Therefore, it has been applied to remote sensing CD datasets to address the issue of an absence of training samples [2, 3]. It has been noticed that the use of GAN is more effective than Variational Auto Encoder (VAE) which requires prior knowledge to map low-dimensional inputs to high-dimensional data [2]. Therefore, GAN was proposed as an unsupervised method that does not need prior information on complicated reasoning processes. It has been gradually improved by researchers to generate high-quality samples [9]. However, the images generated by GAN models are generally not controllable and random. Conditional GAN has been proposed to add some constraints to the training process in order to produce controllable generated images. Therefore, modified versions of GANs have been applied to remote sensing images to address the challenges listed before. Few works that perform remote sensing data augmentation have been conducted recently based on adversarial training such as [9, 2, 3]. These new studies are based on modified GANs to solve image generation problems like poor generation quality, lack of generation label, and slow generation speed of remote sensing image generation task.

In [9], the authors propose an improved data samples generation using a Generative Adversarial Network (GAN) called deeply supervised GAN (D-sGAN) to produce training samples for remote sensing images that cover Anhui Province, China. The application of this method on detecting soil-moving has shown an improvement of 5% compared to the results with no data augmentation technique.

The D-sGAN architecture was developed based on the requirements mentioned before. It adds conditional constraint information which consists of random noise and a rough image segmentation map. This constraint solves the problem of label generation. The random noise based on Gaussian distribution helps to diversify the generated background and the rough image segmentation map can be used as the label for the generated image. In order to generate high-quality remote sensing imagery, the designed model needs to make full use of the original features of the image and reduce the semantic loss. Therefore, the D-sGAN uses the Unet++ network with a new downsampling module that uses the image segmentation map to reduce the semantic loss in the downsampling operation. Inference time is also reduced using deep supervision with the use of the idea of "pruning" to simplify the network by splitting the generator and discriminator into subnetworks with different depths.

In [3], the author addresses the issue of unbalanced classes by efficiently synthesizing effective building CD samples by exploiting generative adversarial training and image blending. The proposed Instance-level change Augmentation (IAug) algorithm creates plenty and diverse buildings changes in the synthesis images. According to the authors, adding more changes to the image can increase the number of positive classes and reduce the risk of class imbalance. This method shows higher performance even with using only 20% of the training data set. It has been noticed that solving class unbalance issue helps also to reduce the training time.

All the aforementioned methods, developed for generating change detection samples, used only RGB channels of remote sensing images due to the limitation of existing change detection datasets. The training phase of these models has not been applied to multispectral (MSI) or hyperspectral (HSI) imagery (more than three channels). Moreover, the generated images still need quality improvement to be more realistic. Therefore, applying DL-based data augmentation methods to generate artificial remote sensing imagery is viewed as a fresh research area and faces some challenges which need further studies to be faced. Some works proposed alternative solutions to perform data augmentation techniques with multichannel remote sensing imagery. For instance, Haut et al. adopted the random occlusion method for generating HSI [7]. In addition, for the same type of imagery, Wang et al. used a data mixture model to extend the labeled training HSI set quadratically [15]. For MSI, a new data augmentation model called spectral index generative adversarial network (SIGAN) has recently been proposed [13]. This model generates synthetic MS images by exploiting class-specific properties through the normalized spectral index. The SIGAN outperforms other GANs in terms of image generation quality.

## 6 Conclusion and Future Works

Data augmentation is one of the best solutions to overpass overfitting problems while training deep learning models on a small number of samples. It becomes the most effective technique, especially for the fields that have limited datasets such as the remote sensing field and its applications. Currently, the research on DL-based remote sensing data augmentation is still at an early stage and faces some serious challenges. However, it is rapidly progressing and has shown promising results in many studies in comparison to traditional augmentation methods as presented in this review. Therefore, many future works can be conducted to improve data generation techniques such as:

– **The first future direction is to improve a DL-learning model that takes into consideration the specific characteristics of RS imagery:** almost all the DL-based data augmentation methods are applied to RGB imagery and the need to be applied to a specific RS imagery type (HSI, MSI, or Radar) is crucial.
– **The quality improvement of generated images by building a DL model** having high performance to generate more realistic RS samples. The impact study of the implementation of the generated images in different remote sensing applications is another very important research area for future work.
– **The development of low computation cost and fast data generation algorithms** is also a fresh and involved research area, especially for real-time data generation and prediction.

## References


1. Cao, Z., Wu, M., Yan, R., Zhang, F., Wan, X.: Detection of small changed regions in remote sensing imagery using convolutional neural network. In: IOP Conference Series: Earth and Environmental Science. vol. 502, p. 012017. IOP Publishing, Beijing, China (2020). https://doi.org/10.1088/1755-1315/502/1/012017
2. Chen, C., Ma, H., Yao, G., Lv, N., Yang, H., Li, C., Wan, S.: Remote sensing image augmentation based on text description for waterside change detection. Remote Sensing **13**(10), 1894 (2021). https://doi.org/10.3390/rs13101894
3. Chen, H., Li, W., Shi, Z.: Adversarial instance augmentation for building change detection in remote sensing images. IEEE Transactions on Geoscience and Remote Sensing **60**, 5603216 (2021). https://doi.org/10.1109/TGRS.2021.3066802
4. Dietterich, T.G.: Steps toward robust artificial intelligence. AI Magazine **38**(3), 3–24 (2017). https://doi.org/10.1609/aimag.v38i3.2756
5. Gauthier, J.: Conditional generative adversarial nets for convolutional face generation. Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, Winter semester **2014**(5), 2 (2014)
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Communications of the ACM **63**(11), 139–144 (2020). https://doi.org/10.1145/3422622
7. Haut, J.M., Paoletti, M.E., Plaza, J., Plaza, A., Li, J.: Hyperspectral image classification using random occlusion data augmentation. IEEE Geoscience and Remote Sensing Letters **16**(11), 1751–1755 (2019)
8. Li, X., Wang, Y., Wang, K., Yan, L., Wang, F.Y.: The paralleleye-cs dataset: Constructing artificial scenes for evaluating the visual intelligence of intelligent vehicles. In: 2018 IEEE Intelligent Vehicles Symposium (IV). pp. 37–42. IEEE (2018). https://doi.org/10.1109/IVS.2018.8500459

9. Lv, N., Ma, H., Chen, C., Pei, Q., Zhou, Y., Xiao, F., Li, J.: Remote sensing data augmentation through adversarial training. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **14**, 9318–9333 (2021). https://doi.org/10.1109/JSTARS.2021.3110842
10. Richter, S.R., Vineet, V., Roth, S., Koltun, V.: Playing for data: Ground truth from computer games. In: European conference on computer vision. pp. 102–118. Springer, Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_7
11. Shi, W., Zhang, M., Zhang, R., Chen, S., Zhan, Z.: Change detection based on artificial intelligence: State-of-the-art and challenges. Remote Sensing **12**(10), 1688 (2020). https://doi.org/10.3390/rs12101688
12. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. Journal of Big Data **6**(1), 60 (2019). https://doi.org/10.1186/s40537-019-0197-0
13. Singh, A., Bruzzone, L.: Sigan: Spectral index generative adversarial network for data augmentation in multispectral remote sensing images. IEEE Geoscience and Remote Sensing Letters **19**, 1–5 (2021)
14. Tero, K., Timo, A., Samuli, L., Jaakko, L.: Progressive growing of gans for improved quality, stability, and variation. In: Sixth International Conference on Learning Representations (ICLR 2018). Vancouver, Canada (2018)
15. Wang, C., Zhang, L., Wei, W., Zhang, Y.: Hyperspectral image classification with data augmentation and classifier fusion. IEEE Geoscience and Remote Sensing Letters **17**(8), 1420–1424 (2019)
16. Xiao, Q., Liu, B., Li, Z., Ni, W., Yang, Z., Li, L.: Progressive data augmentation method for remote sensing ship image classification based on imaging simulation system and neural style transfer. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **14**, 9176–9186 (2021). https://doi.org/10.1109/JSTARS.2021.3109600
17. Yan, Y., Zhang, Y., Su, N.: A novel data augmentation method for detection of specific aircraft in remote sensing RGB images. IEEE Access **7**, 56051–56061 (2019). https://doi.org/10.1109/ACCESS.2019.2913191
18. Yu, Y., Gong, Z., Zhong, P., Shan, J.: Unsupervised representation learning with deep convolutional neural network for remote sensing images. In: Zhao, Y., Kong, X., D., T. (eds.) International Conference on Image and Graphics. vol. 10667, pp. 97–108. Springer, Springer, Cham (2017). https://doi.org/10.1007/978-3-319-71589-6_9
19. Zhang, W., Lu, X.: The spectral-spatial joint learning for change detection in multispectral imagery. Remote Sensing **11**(3), 240 (2019). https://doi.org/10.3390/rs11030240
20. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232. Venice, Italy (2017). https://doi.org/10.1109/ICCV.2017.244
21. Zou, Z., Shi, Z.: Random access memories: A new paradigm for target detection in high resolution aerial remote sensing images. IEEE Transactions on Image Processing **27**(3), 1100–1111 (2017). https://doi.org/10.1109/TIP.2017.2773199