A New Remote Sensing Change Detection Data Augmentation Method Based on Mosaic Simulation and Haze Image Simulation

Zhipan Wang, Di Liu, Zhongwu Wang, Xiang Liao, and Qingling Zhang

***Abstract*— The quality of optical remote sensing images is largely affected by clouds and haze. In addition, the mosaicking image of multiple remote sensing images, due to objective factors such as acquiring time or climate conditions, will lead to large spectral differences in the area around the seamline. The aforementioned scenarios will seriously affect the accuracy of change detection models based on deep learning. However, there is still a lack of methods to address such issues. To solve these problems, from the perspective of training samples, this article proposed a simple but effective data augmentation method to improve the generalization ability of the deep change detection model in the region of haze cover and the seamline. First, from the characteristics of the optical remote sensing image itself, two image simulation methods are designed to conduct data augmentation, named mosaic simulation and haze image simulation. Then, the newly augmented training samples are mixed with the original training samples and then input into a deep learning model for model training. Finally, the change detection results indicate that the proposed data augmentation method can effectively improve the generalization ability of the change detection model in the region of haze cover and seamline, which has high practical value for improving the deep learning model's performance in real-world scenarios and also provides a simple but effective algorithm reference for other intelligent interpretation tasks from the perspective of training data.**

***Index Terms*—Change detection, data augmentation, high-resolution remote sensing image.**

I. Introduction

REMOTE sensing images have the advantages of large width, continuous and frequent observation, and low cost, making them become the main data source for monitoring large-area change detection. Deep learning-based change detection algorithms have become themainstream algorithms for remote sensing change detection [1] due to their excellent performance and generalization ability, especially compared with traditional machine learning methods.

At present, most of the existing deep learning-based change detection methods are applied to high-resolution remote sensing imagery without cloud contamination. Researchers have paid a lot of attention to model architecture design [2] but less considering that optical remote sensing images are usually polluted by haze or thin clouds in real-world scenarios, which greatly reduces the clarity of objects and then affects the model detection accuracy. In addition, several high-resolution remote sensing images are often required to mosaic into a large image for the following large-area applications tasks, and the spectral difference along the seamlines regions in mosaicked images is often very obvious. As a result, the performance of change detection models trained on ideal quality samples usually decreases significantly when applied in complex real-world scenarios. Therefore, how to improve the generalization ability of deep learning-based change detection models in complex scenarios is a very valuable but challenging scientific and engineering problem.

Deep learning methods are essentially data-driven, and their performance largely depends on the labeling accuracy and the quantity of training samples. However, in practical application scenarios, it is very challenging to obtain enough training samples in most cases, and high-quality training samples require a lot of manpower and material resources. To address the problem of model overfitting, caused by insufficient training samples, data augmentation has been proven to be a very effective means [3]. Existing data augmentation methods can be divided into two categories: basic image manipulation and deep learning methods [3]. Basic image manipulation mainly includes image processing methods such as rotation, flipping, Gaussian noise, random erasing, and mixing images [4], [5], [6], [7], [8], [9], [10]. Deep learning methods improve model generalization ability by learning, such as the generative adversarial network (GAN), which learns the latent probability distribution of existing samples [11], [12], [13], [14], [15] to generate new samples by oversampling. However, data augmentation based on deep learning methods has two limitations. On the one hand, the quality of new samples is difficult to guarantee, and the generated new training samples may deviate from expectations. On the other hand, the newly generated samples are poorly interpretable, which may lead to even worse detection accuracy.

Manuscript received 10 November 2022; revised 14 March 2023; accepted 17 April 2023. Date of publication 25 April 2023; date of current version 22 May 2023. This work was supported by the Shenzhen Science and Technology Innovation Project under Grant ZDSYS20210623091808026, in part by the National Natural Science Foundation of China (General Program), under Grant 42071351, in part by the National Key Research and Development Program of China under Grant 2020YFA0608501, and in part by the Chongqing Science and Technology Bureau technology innovation and application development special under Grant cstc2021jscx-gksb0116. *(Corresponding author: Qingling Zhang.)*

Zhipan Wang, Di Liu, and Qingling Zhang are with the Shenzhen Key Laboratory of Intelligent Microsatellite Constellation, Shenzhen Campus of Sun Yat-sen University, Shenzhen 528406, China (e-mail: wangzhp25@mail2.sysu.edu.cn; liud73@mail2.sysu.edu.cn; zhangqling@mail.sysu.edu.cn).

Zhongwu Wang is with the Remote Sensing Application Center, Land Satellite Remote Sensing Application Center MNR, Beijing 637141, China (e-mail: wangzhongwu0514@outlook.com).

Xiang Liao is with the Chongqing Pioneer Satellite Technology Co., Ltd., Chongqing 401420, China (e-mail: liaox@mgcsat.com).

Digital Object Identifier 10.1109/JSTARS.2023.3269784

For change detection tasks in real-world scenarios, two key issues must be considered. The first one is, in haze-covered cloudy tropical and subtropical regions, optical remote sensing images are easily affected by clouds and haze, and it is difficult to obtain high-quality cloud-free images in some areas even for several months. Thus, researchers usually use cloud masking or cloud removal methods to improve the quality of images. The main idea of the cloud masking method is that cloud-covered parts in an image are invalid areas, thus masking them and keeping the ideal pixels is a natural idea. At present, there are many cloud masking methods, such as the Fmask algorithm [16]. The cloud removal methods, especially in the area covered by thin clouds, believe that cloud interference only partially affects the recognition of ground objects. If thin clouds can be removed, then the model performance in such regions can be improved. Most of the traditional cloud and haze removal algorithms are inspired by the dark channel prior algorithm [17]. With the development of deep learning, the cloud and haze removal algorithm based on deep learning has gradually become mainstream, and the thin cloud removal effect is greatly improved compared with traditional methods. However, the deep learning method is highly dependent on high-quality samples and requires a large amount of computation. Both cloud masking and cloud removal methods require preprocessing of remote sensing images, which will reduce the computation efficiency and increase the complexity of the whole process.

The other problem is the large spectral differences along the seamline regions in mosaicking images. The width of a single high-resolution remote sensing image is usually small; therefore, it is necessary to mosaic multiple images into a large image for large-scale change detection. As we know, the spectral difference between different images may be enormous due to seasonal rotation, observation angle of satellites, etc., therefore, such a phenomenon will increase the commission and omission alarms in the final change detection result based on deep learning models. Although the problem of spectral difference can be eliminated to a certain extent by using the uniform color algorithm [18], it is still difficult to achieve a smooth transition like a natural image along seamlines. Furthermore, the computation capacity is also relatively high for change detection tasks, and too complicated preprocessing step will decrease the automation level of the whole change detection task.

We can rethink the essence of deep learning from another perspective. Deep learning simulates the way of thinking like a human brain, and haze only reduces the recognizable level of ground objects, but for human eyes, ground objects are still can be identified in remote sensing images. Therefore, we can imagine that if a deep learning model learns enough semantic information from the haze or thin clouds covered regions or seamline regions with large spectral differences, its generation ability will be improved. Thus, a simple but effective way is to generate sufficient new training samples from haze/thin cloud-covered images or image parts along seamlines with large spectral differences. In this article, a new data augmentation method for remote sensing change detection in real-world scenarios is proposed, which improves the detection accuracy of a deep learning model in areas covered by haze/thin clouds or large spectral differences along seamlines. Moreover, it also provides ideas for object detection or land cover classification in other real-world scenarios.

The main contributions of this article are as the following.

1) To the best of our knowledge, we are the first one who proposes a data augmentation method to improve change detection model accuracy in complex areas. Our data augmentation method contains mosaic simulation and haze image simulation, which are easy to implement, and can also be integrated into remote sensing image change detection algorithms.
2) Our data augmentation method is also suitable for other downstream tasks, such as object extraction in a haze coverage area, and it also provides a new method reference for other intelligent interpretation frameworks from the perspective of training data generation.

The rest of this article is organized as follows. Section II introduces the related works of data augmentation. In Section III, the details of the proposed method are presented. The change detection result based on various data augmentation methods is displayed in Section IV. In Section V, a further discussion of the proposed method is presented. Finally, Section VI concludes this article.

II. Related Work

Data augmentation is utilized to improve the generalization ability of deep learning models [17], such as image classification, object detection, semantic segmentation [19], and instance segmentation [21]. Data augmentation is not the only way to improve the generalization ability of deep learning models, other methods such as dropout [22], and batch norm [23] are proven effective. In addition, model architecture can also prevent model overfitting. However, although data augmentation is usually used to improve model performance from the perspective of data, it can also be integrated with model architecture, such as dropout [22], to further improve model performance and reduce the risk of model overfitting. Therefore, we do not discuss the influence of model structure in this article. Data augmentation methods can be divided into two main categories: basic image manipulations and deep learning methods [3].

1) *Basic image manipulations*: It mainly uses the basic image operation for data augmentation, such as geometric transformations [24], color space augmentation [25], kernel filter [26], random erasing [8], and mixing images [4].

Geometric transformations usually use image rotation, image flip, and other methods to accomplish, which is suitable for semantic segmentation tasks. Color space augmentation mainly uses image channel transformation, such as converting the original RGB three-channel image into a single-channel image, transforming the original RGB image into HSV, YUV, and other color spaces, or randomly exchanging RGB channels. Color space augmentation is relatively suitable for image classification tasks, but it is not completely suitable for pixel-level segmentation tasks, especially for remote sensing images. Because the random exchange of band information will change the original features of objects, such as forests in the false-color image is red, a random exchange between the visible band and the

near-infrared band will make forests and other objects lose their original spectral information, and then may affect the model performance.

Kernel filter [26] mainly utilizes image filtering to sharpen and blur training samples to generate new samples, such as the PatchShuffle method [26]. For example, researchers conducted experiments on the CIFAR-10 dataset by using the PatchShuffle method, and the error of prediction was reduced by 0.67%. Essentially, the convolutional neural network is also a filtering method, so the kernel filter data augmentation method is not very obvious for absolute accuracy improvement [17]. The Random erasing method mainly processes the training samples by randomly removing small blocks to avoid the model paying much attention in a local area, thereby reducing the risk of overfitting. For remote sensing change detection, the random erasing method cannot be used directly because it may remove the changing area and cause the model to learn wrong semantic information. Mixing images are widely used in object detection. For instance, the YOLOv4 in [20] proposed a new data augmentation method, which mosaics four images into one to effectively improve object detection accuracy. However, for pixel-level segmentation tasks, mixing images will lead to spatial information discontinuity, which may negatively affect model accuracy.

Generally speaking, the mathematical principle of basic image manipulations is clear and simple, and it can simulate the actual distribution of data to a certain extent. Therefore, most of the current remote sensing change detection models based on deep learning [27], [28], [29], [30], [31], [32] utilize basic image manipulations as the preprocessing step to improve model generalization ability.

2) *Deep learning methods*: It is similar to mixing images, but it uses a learned approach for data augmentation. Such methods mainly contain adversarial training [33], GANs [34], neural style transfer [35], and metalearning [36].

Adversarial training uses two or more deep learning models for adversarial training, which can effectively reduce commission errors. GAN can produce new training samples from noise by learning general knowledge from a large number of training samples. The principle of the neural style transfer is ingenious, it can keep the main content of training samples unchanged but only change the style form to generate new training samples to improve the performance of a model. Metalearning improves the generalization ability of a model by adding one-dimensional or multidimensional data and performing mixed training with original samples.

As a whole, data augmentation methods based on deep learning are effective in image classification but they are also not completely suitable for remote sensing change detection, because they may change the pixel distribution of an input image, which may cause a label image to mismatch the training images. Furthermore, methods of this kind are also very time-consuming and difficult to converge [17].

Let us rethink the original purpose of data augmentation, its essence is to simulate the true distribution of training data, and improve the performance of a model on the test dataset [17]. Optical remote sensing images are often contaminated by haze and large spectral differences along seamlines regions.

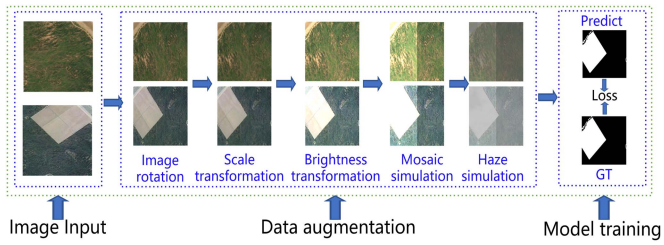

Fig. 1. Proposed data augmentation method for change detection.

Therefore, from the perspective of training data, we can effectively improve the model generalization ability and detection accuracy by new training samples that are generated by simulating haze cover and seamlines.

III. Proposed Method

The flowchart of the proposed method is displayed in Fig. 1. We can see that each two time-phase change detection training samples will be augmented by several data augmentation methods, such as image rotation, scale transformation, and haze simulation. Then, the augmented samples are mixed up with the original samples, and finally, all of them will be sent to the deep learning model for training to improve the model generalization ability.

A. Mosaic Simulation

To improve the detection accuracy of the deep change detection model along seamlines regions, we have developed a data augmentation method, named Mosaic Simulation. As we discussed previously, the pixel spectral difference on both sides of the seamline from two adjacent images is often enormous in a large mosaiced image. Taking advantage of this fact, we can use random subregion spectral transformations to accomplish mosaic simulation.

First, for an arbitrary training sample, we randomly generate a seamline in the horizontal or vertical direction, and the seamline is determined as follows:

$$\text{line} = \text{random}(0,\ \text{imgsize}) \tag{1}$$

where *random* represents generating a random value between *0* and *imgsize*. imgsize stands for the size of input images. The generation of random lines is divided into two cases: the first one is to select rows for spectral transformation, and the other one is to select columns for spectral transformation. For any input training sample, the aforementioned two methods are randomly selected due to realistic multiple images mosaic situations; the mosaicking operation will not be stitched in only one direction, thus we randomly select the horizontal subregions or the vertical subregion to perform the spectral transformation. The spectral transformation is as follows:

$$\text{Transform Value} = \text{scale} \times [0,\ \text{pixel Value}] \tag{2}$$

where *Transformvalue* denotes random linear transformation in the interval [*0, pixelValue*], and the *scale* represents the spectral stretching ratio, which significantly controls the brightness of

Algorithm 1: Mosaic Simulation.

Input: bi-temporal images $D_{\text{sample\_}T1}$, $D_{\text{sample\_}T2}$
Output: D_{aug}
1: Initialize D_{sample}, random select from $D_{\text{sample\_}T1}$, $D_{\text{sample\_}T2}$
2: // Iterate each sample in D_{sample}
3: **for** D_{aug} in D_{sample}:
4: Generate mosaic line by (1)
5: Subrange pixel value transform by (2)
6: Image saving: new training sample, D_{aug}
7: **end for**

Algorithm 2: Haze Image Simulation.

Input: Thin cloud image dataset $D_{\text{thinCloud}}$
Input: Sample image D_{sample}
Output: D_{aug}
1: Initialize $D_{\text{thinCloud}}$ and D_{sample}. D_{sample} is generated by random selected from $D_{\text{sample\_}T1}$, $D_{\text{sample\_}T2}$
2: // Iterate each sample in D_{sample}
3: **for** D_{aug} in D_{sample}
4: Get *t(x)* image based on (8)
5: Generate a new sample based on (3)
6: Image saving: new training sample, D_{aug}
7: **end for**

the subregion. With such a spectral transformation method, the seamline effect can be simply simulated. In a whole training dataset, the new training samples are generated with the mosaic simulation method as follows.

B. Haze Image Simulation

How to simply simulate the effect of haze or thin cloud coverage on a cloud-free sample image? Inspired by the dark channel prior algorithm [37], [38], [39], [40], [41], [42], we have developed a simple haze image simulation method. First, let us revisit the simulation equations for haze images

$$I(x) = J(x)\,t(x) + A\,(1 - t(x)) \tag{3}$$

where *I(x)* denotes a hazy image, *J(x)* denotes a haze-free image, *t(x)* denotes the transmittance image, *A* is the atmospheric light value, and *x* is the single pixel in the whole image [37].

Our ultimate goal is to simulate the real sample image as a hazy image. According to the above formula, once we know the transmittance image *t(x)* and atmospheric light value *A*, a real sample image can be simulated as the effect of haze coverage. To simulate the coverage of haze and mists on a real image, we obtained 15 scenes of GF1-PMS and GF2-PMS true color images in multiple geographic regions, such as in the southeast coast, central, northwest, northeast, and southwest of China. All high-resolution remote sensing images are downloaded from China Resource Satellite Center (http://www.cresda.com/CN/), and 400 image blocks with 512 × 512 pixels of various real haze-covered scenes are manually cropped from the image as a hazy image dataset. It is worth noting that this haze image dataset do not contain thick clouds because the spectral values of the ground objects under thick clouds have been completely lost.

In hazy or thin cloud-covered areas, the morphological feature of the ground object is similar, thus 400 hazy or thin cloud-covered images are enough to represent all kinds of scenes. For an arbitrary input training sample, we first randomly select an image *I(x)* from the hazy image dataset to construct the transmittance image *t(x)* and atmospheric light value *A*, and then use the dark channel prior algorithm to separate the haze image *J(x)* from the input image *I(x)*. Finally, the separated haze image *J(x)* is added to the haze-free training sample image; therefore, a sample image covered by haze can be regenerated.

Based on [37], we can get atmospheric light value *A*. As described in [37], the pixel value of each channel is first sorted in ascending order, and then the average value of the top 1% pixels is selected as the atmospheric light value *A*. Once the atmospheric light value *A* is known, we can transform (3) into (4) as follows:

$$\frac{I^c(x)}{A^c} = t(x)\frac{J^c(x)}{A^c} + 1 - t(x). \tag{4}$$

In (4), *C* represents the band number of an image. For a true-color image, *C* is 3. Assuming that the local area of the transmittance image *t(x)* is invariant, we continue to transform (4) into (5)

$$\begin{aligned}&\min_{y\in\Omega(x)}\left(\min_c \frac{I^c(x)}{A^c}\right)\\ &\quad = t(x)\min_{y\in\Omega(x)}\left(\min_c \frac{J^c(x)}{A^c}\right) + 1 - t(x).\end{aligned} \tag{5}$$

According to the dark channel prior theory, the value of the dark channel image in the area without haze coverage is approximately equal to 0, which satisfies the following equation:

$$J_{\text{dark}} = \min_{y\in\Omega(x)}\left(\min_c \frac{J^c(x)}{A^c}\right) = 0. \tag{6}$$

We substitute (6) into (5), then the estimation equation of transmittance image *t(x)* can be expressed as

$$t(x) = 1 - \min_{y\in\Omega(x)}\left(\min_c \frac{I^c(x)}{A^c}\right). \tag{7}$$

To keep the image with certain cloudiness, we introduce an identity parameter *w* ($0 < w < 1$), then the transmittance image *t(x)* calculation formula is expressed as

$$t(x) = 1 - w \times \min_{y\in\Omega(x)}\left(\min_c \frac{I^c(x)}{A^c}\right). \tag{8}$$

In the training dataset, the new training samples are generated with haze image simulation as follows.

Based on mosaic simulation and haze image simulation algorithms, a new training sample is generated, and the effect is shown in Fig. 2. In Fig. 2, we can see that the haze coverage effect in a real-world scene is simulated to a certain extent. If the newly generated samples and the original samples are mixed for training, will the model accuracy be effectively improved in the haze cover scene? We will discuss and analyze the effect in detail in Section IV.

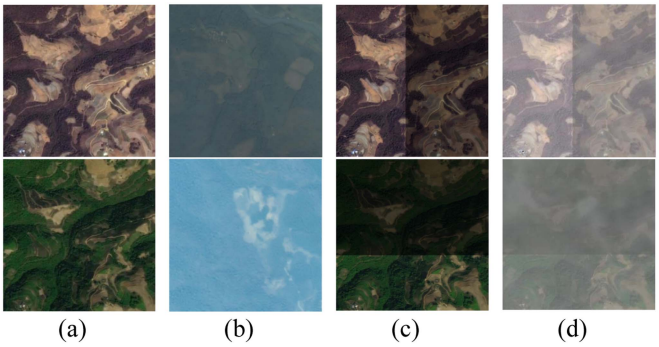

Fig. 2. Data augmentation. (a) T1 and T2 images. (b) Cloud template image with different thicknesses. (c) Data augment by mosaic simulation. (d) Data augment by haze image simulation.

C. Accuracy Assessment

We used Recall, Precision, F1, and over accuracy (OA) for accuracy assessment [43]. The Recall indicator reflects the omission error of positive samples, The higher the value, the lower the omission of positive samples. The Precision indicator reflects the commission error of positive samples, The higher the value, the lower the commission error of positive samples. The F1 metric is a combination of Recall and Precision metrics, which visually evaluates the comprehensive performance of the algorithm, while the OA metric is a standard classification accuracy evaluation, which visually reflects the detection accuracy of positive and negative samples. The mathematical calculation formula for the above indicators is as follows:

$$\begin{aligned} \text{Recall} &= \frac{\text{TP}}{\text{TP}+\text{FN}} \\ \text{Precision} &= \frac{\text{TP}}{\text{TP}+\text{FP}} \\ \text{F1} &= \frac{2\times\text{Recall}\times\text{Precision}}{\text{Recall}+\text{Precision}} \\ \text{OA} &= \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \end{aligned} \tag{9}$$

where TP indicates that the predicted pixel result is a positive sample and the ground truth pixel is also a positive sample. TN indicates that the predicted pixel is a negative sample and the ground truth pixel is also a negative sample. FP indicates that the pixel in the ground truth image is a positive sample but is predicted as a negative sample. FN indicates that the pixel in the ground truth image is a negative sample but is predicted as a positive sample.

IV. Experimental Details

To verify the effectiveness of the proposed data augmentation method, a large-scale real-world scene of bitemporal high-resolution images in Sipsongpanna was conducted (see Fig. 3), and we will evaluate the effectiveness of the proposed data augmentation method from the qualitative and quantitative perspective.

Sipsongpanna is located in Yunnan Province, China; it belongs to the tropical monsoon climate, with mostly primary forests and excellent natural conditions, and the famous Sipsongpanna Primeval Forest Park is located here. According to the historical remote sensing image coverage experience, it is difficult to obtain high-quality optical remote sensing images here, thus to generate the high-quality base map at each time phase, multiple data sources are used, including 2m GF1-PMS, 1m GF2-PMS, and 2m GF6-PMS. In this study, all high-resolution remote sensing images are acquired from http://www.cresda.com/CN/ and are preprocessed by geometric correction based on ArcGIS10.6 software. Only true-color bands are preserved for the final change detection task. The acquisition time of bitemporal images is in the summer season of 2020 and 2021, respectively. The final image size of each time phase in 2020 and 2021 is 67 827 × 88 524 × 3 pixels.

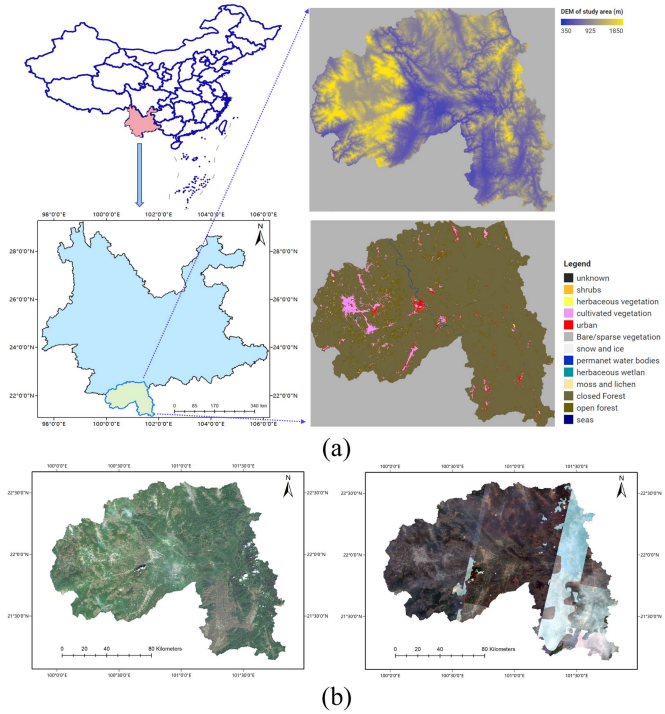

Fig. 3. Study area. (a) DEM and landcover map of Sipsongpanna. (b) Mosaicking image of Sipsongpanna, the left, and right maps are 2020 and 2021 images, respectively.

In most cases, Sipsongpanna is usually covered with haze or thick clouds. The bitemporal images are displayed in Fig. 3. We can see that the mosaicking image of 2021 is stitched by multiple images of different time phases, the spectrum around the seamline area is quite diverse, and the lower-right region in the whole image is covered with haze and thick clouds. To evaluate the deep learning model performance, the ground truth image is generated by human visual interpretation.

A. Dataset and Study Area

First, we produce a deforestation detection dataset to train a deep learning model. The spatial resolution of this dataset is 2 m (the main data sources are GF1-2m, GF2-1m, and GF6-2m. GF2 is resampled to 2 m using the nearest neighbor sampling

TABLE I
PARAMETER SETTING OF DIFFERENT DATA AUGMENT METHODS

Method	Implementation	Parameter setting
Image rotation	The direction of the objects in remote sensing images is irregular from different angles, so we set the rotation angle to be between [45, 300] degrees; it can represent the morphological characteristics of objects at various angles. In the specific operation, each sample is rotated, and then the image is interpolated to a size of 512×512 pixels with the nearest neighbor method, thus the image size of the newly generated training samples is unchanged.	The range of image rotation: 45°–300°
Scale transformation	The difference of scale, intuitively manifested is spatial resolution. To describe the characteristics of ground objects at different scales, we use the nearest neighbor interpolation method to scale up and scale down training samples. In high-resolution remote sensing training samples, the scale parameter is set to [0.4, 3.2], which indicates the process of an object from blurring to clear, respectively. This scale setting can represent the scale effect of most objects at different spatial resolutions.	The range of scale transformation: 0.4–3.2
Brightness transformation	Under different angles and lighting conditions, the spectral differences of the same area photographed by satellite are very obvious. To describe the spectral difference, we set the change rate of DN value to [0.6, 2.4], representing the process of objects from dark to bright, respectively.	The range of brightness transformation: 0.6–2.4
Mosaic simulation	For a single training sample, randomly select a row or column for initialization, and the selected point index range is [0, 512], which means that the seamline can be generated arbitrarily from the starting position to the edge, and then generates new samples with mosaic simulation.	The stretched ratio of DN value: 0.6–2.4
Haze image simulation	For each training sample, we randomly choose a haze image from our haze image dataset (400 images) as the template image, then, the new haze-simulated image is generated.	Haze image simulation: Randomly choose a haze image from the haze image dataset as a template to generate a new sample.

method), and the production time of the former time-phase image is 2019, and the latter time-phase image is 2020. The sample collection area is mainly located in the Yangtze River Economic Belt of China, including 11 provinces, such as Hunan Province, Hubei Province, Jiangxi Province, Jiangsu Province, Guizhou Province, and Guangxi Province. The quantity of samples is 8002 true-color images, and the image size of each sample is 512×512 pixels.

B. Change Detection Method and Implementation Details

Most of the existing deep learning-based change detection methods are based on fully convolutional neural networks from the pixel-level segmentation aspect to obtain binary change results [43]. From the perspective of the deep learning model architecture, the encoding–decoding structure is the mainstream model structure, it is simple and elegant, and it also can easily implement the end-to-end change detection task. To fairly compare the accuracy improvement of the new data augmentation method, SNUNet in [43] is used as our baseline change detection method. SNUNet is a dense feature connection model with a siamese structure, its main contribution is to improve the segmentation accuracy in small targets and boundary areas by designing an ensemble channel attention module.

Our operating system is Win10 with AMD 5600x CPU and 64GB RAM, and the GPU is a TeslaV100 with 32GB memory. The deep learning library is Pytorch1.8.1, Adam is used as the model optimizer, the initial learning rate is 0.0001, and the total training epoch is 60. For every 10 epochs, the initial learning rate will drop to 1/10 of the original learning rate. Due to the limitation of GPU memory and computational efficiency, the batch size is set to 22. Once the model is trained, the model inference is carried out with half-precision to save GPU memory and improve processing efficiency.

From the characteristics of remote sensing images, we select several corresponding data augmentation methods for comparison and then analyze the impact of different data augmentation methods on the accuracy of change detection. In remote sensing images, scale, and spectrum are the most prominent features. The ground objects on remote sensing images with different spatial resolutions are quite different, such phenomenon is also the most significant difference between remote sensing images and natural images. Both scale transformation and image rotation will change the original position information of the input image, which has a positive impact on improving the generalization ability of the change detection model.

To simulate the spectral differences of ground objects under different lighting conditions, the commonly used means is brightness transformation. Therefore, we adopt three methods to simulate spectral difference, and they are image rotation, scale transformation, and brightness transformation [3]. The implementation details of different comparison algorithms are given in Table I.

To compare the performance of data augmentation methods fairly, we have built three comparison methods (see Table II).

1) *Method 1*: Image data augmentation with random rotation, scale transformation, and brightness transformation on a training sample in sequence, then mix the new samples with the original samples for the next step of model training.
2) *Method 2*: Image data augmentation with random rotation, scale transformation, brightness transformation, and

TABLE II
COMPARISON METHODS

Data augment methods	Image rotation	Scale transformation	Brightness transformation	**Mosaic simulation**	**Haze image simulation**
Method 1	√	√	√		
Method 2	√	√	√	√	
Ours	√	√	√	√	√

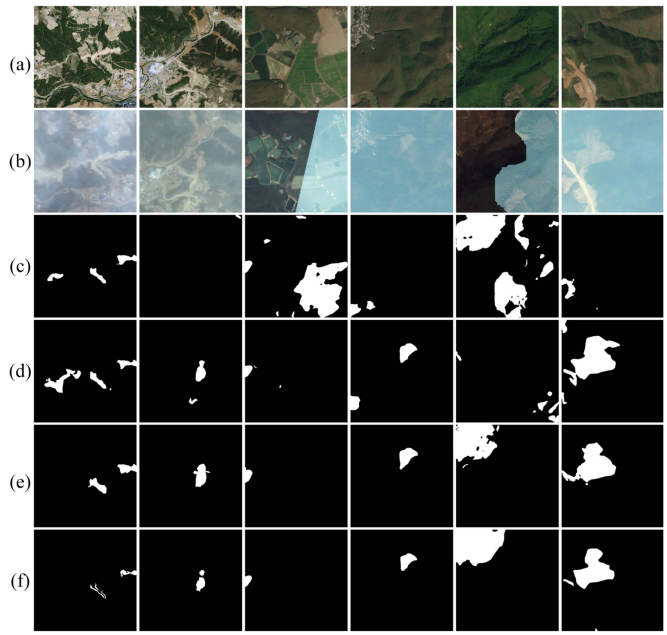

Fig. 4. Deforestation detection in haze-cover scenarios. (a) T1 image. (b) T2 image. (c) Change detection with data augment Method 1. (d) Change detection with data augment Method 2. (e) Change detection with our data augment method. (f) Ground truth.

mosaic simulation on a training sample in sequence, then mix the new samples with original samples for the next step of model training.

3) *Ours*: Image data augmentation with random rotation, scale transformation, brightness transformation, mosaic simulation, and haze image simulation on a training sample in sequence, the next step is the same as Method1 and Method2.

C. Qualitative and Quantitation Comparisons

1) Deforestation Detection in a Complex Scenario: To verify the change detection accuracy of the proposed data augmentation method in a complex scenario, we selected 1024 images (each image size is 512 × 512 pixels) from the Yangtze River Economic Belt of China for accuracy evaluation. All remote sensing images were obtained from the China Resources Satellite Center (http://www.cresda.com/CN/), and the resolution of the base map image is 2 m, the former time phase image is in 2020, and the latter time phase image is in 2021. To evaluate the accuracy efficiently and reasonably, several change detection results are shown in Fig. 4.

In Fig. 4, our change detection result is visually better than other methods. Data augmentation with Method 1 does not use mosaic simulation or haze image simulation, thus the trained change detection model cannot detect the slight deforestation change in the haze cover region, and there is also a little false detection result in the edge and seamline area. In addition to the means used in Method 1, mosaic simulation is adopted in Method 2. From the visual perspective, the trained change detection model can detect deforestation change in the seamline area but it still produces a few false detected results in the haze coverage area. Generally speaking, Method 2 is better than Method 1. In a word, from the qualitative perspective, mosaic simulation and haze image simulation are used in our data augmentation method at the same time; as we can see from Fig. 4, we have obtained the best deforestation detection results, especially in the haze coverage and seamline regions.

TABLE III
ACCURACY EVALUATION

Data augment methods	Recall	Precision	F1	OA
Method 1	0.6762	0.9683	0.3981	0.7240
Method 2	0.7412	0.9732	0.4918	0.7823
Ours	**0.8103**	**0.9853**	**0.5732**	**0.8461**

The quantitative accuracy assessment result is displayed in Table III. It can be seen that compared with Method 1 or Method 2, our data augmentation method has achieved relatively higher accuracy, it also proves the effectiveness of mosaic simulation and haze image simulation data augmentation methods. Moreover, from the practical aspect, although mosaic simulation or haze image simulation increases the preprocessing time to a certain extent, it provides a cheap and effective solution to improve the model generalization ability in low-quality scenarios.

2) Deforestation Detection in a Real-World Scenario: The deforestation detection results in Sipsongpanna with different data augmentation methods, as shown in Fig. 5.

As a whole, our method effectively reduces the commission alarms in the low-right region. The deforestation area of Method 1 is 25.15 km^2, the deforestation area of Method 2 is 15.92 km^2, our deforestation area is 12.11 km^2, and the GT area is 9.12 km^2. The deforestation area of our method is closer to the GT. To display the result clearly, we enlarged the subregions A, B, and C in Fig. 6.

In Fig. 6, we can see that the bitemporal spectral difference in subregion A is obvious, and deep learning model trained with data augmentation Method 1 does not detect deforestation change very well; in addition, there are more commission alarms in the seamline area compared with other two methods due to excessive spectral differences. Change detection results with data augmentation method 2 can effectively reduce

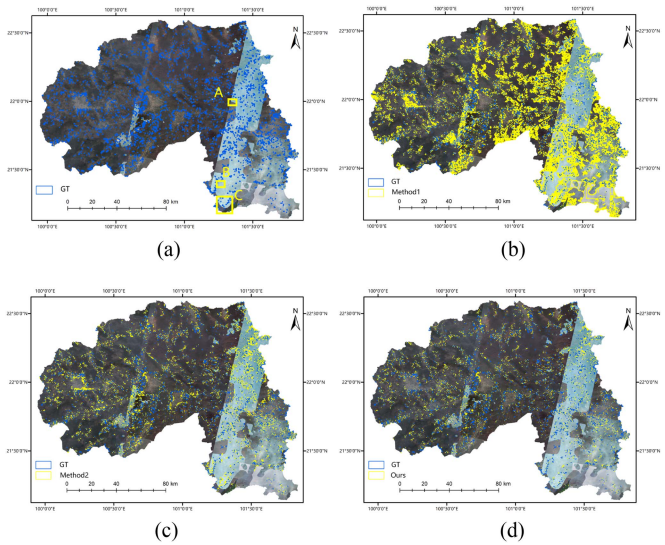

Fig. 5. Deforestation detection result. (a) GT. (b) Method1 result. (c) Method2 result. (d) Our result. (the blue color represents ground truth, and the yellow color represents deforestation result of different algorithms.).

TABLE IV
ACCURACY EVALUATION

Data augment methods	Recall	Precision	F1	OA
Method 1	0.4897	0.3240	0.3118	0.9671
Method 2	0.6014	0.4017	0.4284	0.9811
Ours	**0.6815**	**0.4381**	**0.4891**	**0.9851**

commission alarms in the seamline area; however, there are still some false detection results in the haze coverage area. With the help of mosaic simulation and haze image simulation, our method can effectively reduce the commission alarms; it significantly improved the model generation ability in the seamline regions.

In subregion B, this region is mainly used to test the effect of model detection under relatively thick cloud cover conditions. Despite the image in this region being covered with thick haze, the morphological and textural features of ground objects are still discernible. As can be seen, the change detection result with data augmentation Method 1 has a relatively high omission alarm, and deforestation change cannot be well perceived in most cases. For the deforestation detection result with data augmentation Method 2, the visual effect is improved in some regions but there are still a few omission alarms in them. Other deforestation detection results, such as in subregion C, also prove the effectiveness of our data augmentation method. Although there are a few missed detection pixels in our result, the overall accuracy improvement is obvious compared with the other two data augmentation methods.

The quantitative accuracy evaluation of deforestation detection results in Sipsongpanna is given in Table IV. From Table IV, we can see that our data augmentation method achieves the highest quantitative accuracy. Thus, it supports our hypothesis

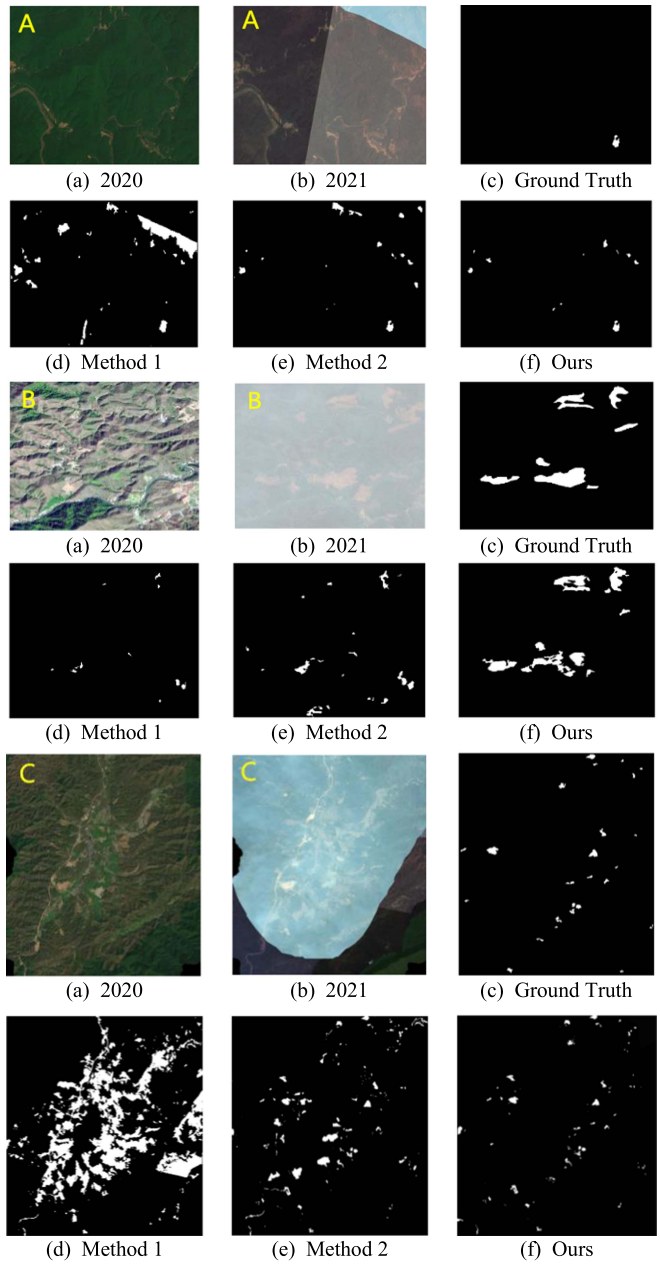

Fig. 6. Change detection result of different data augment methods. (a)–(c) Subregions image of Sipsongpanna in 2020 and 2021, Ground truth of deforestation detection, respectively. (d) Change detection by Method 1. (e) Deforestation detection by Method 2. (f) Deforestation detection by Method 3.

that the model performance in complex scenes will be improved by using mosaic simulation and haze image simulation.

V. DISCUSSION

Why the proposed data augmentation method can effectively improve the accuracy of the deep learning model in complex scenes? To answer this question, we explain it from the perspective of features extracted by deep learning, and we also extend the discussion of the application scope of our new data augmentation methods.

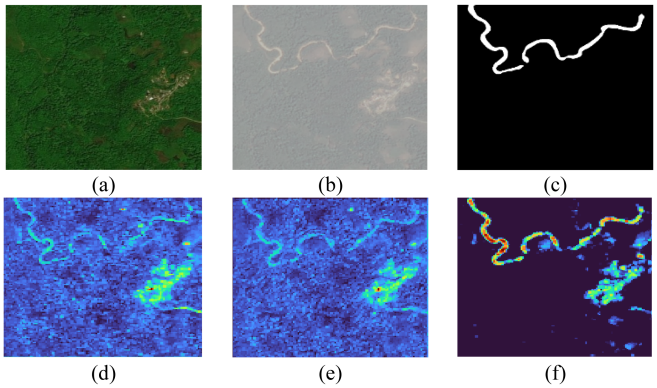

Fig. 7. Visualization of the feature map. (a) T1 image. (b) T2 image. (c) GT. (d) Feature map of the model trained by data augment method 1. (e) Feature map of a model trained by data augment method 2. (f) Feature map of the model trained by the proposed method.

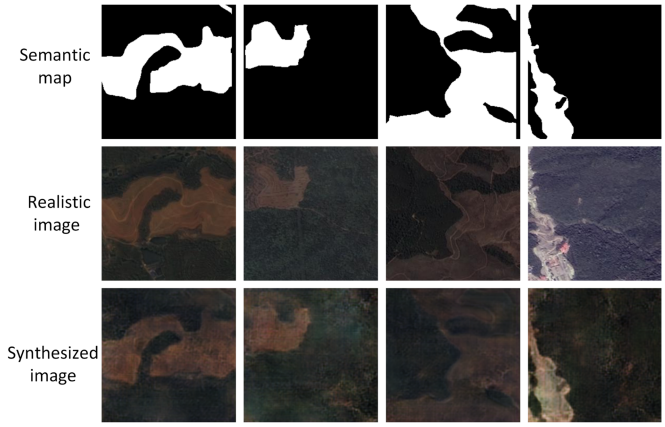

Fig. 8. Newly generated training samples based on the GauGAN model.

TABLE V
ACCURACY EVALUATION

Data augment methods	Recall	Precision	F1	OA
GauGAN	0.7908	0.9800	0.5123	0.7991
Ours	**0.8103**	**0.9853**	**0.5732**	**0.8461**

A. *Influence on Deep Learning Model Extraction Features*

To analyze the reasons for the improvement of model accuracy by different data augmentation methods, we dissect it from the perspective of deep neural network feature extraction. Through the visualization of intermediate features [44], the differences between various models trained with different data augmentation methods are displayed in Fig. 7.

In Fig. 7, the feature maps of different models correspond to the convolutional kernel with the maximum weight of the layer, which indicates the most effective feature extraction result of the layer, and the network layers corresponding to different models are the penultimate output of the decoding layer of the SNUNet network, which is sufficient to represent the high-level semantic features of the model. It is clear to see that the trained model based on our data augmentation method has a strong feature extraction ability as it can perceive the change region very effectively. The trained model with the data augmentation Method 1 has the weakest feature extraction ability, while the effect of the model trained with mosaic simulation is somewhat improved, but the feature extraction results are still not yet obvious. Fundamentally, the deep learning model is powerful enough to learn general semantic information from a large training sample. The previous change detection models do not take into account that, in real-world change detection scenarios, the probability of optical remote sensing images being covered by haze or clouds is high, but the existing training dataset usually only contains good quality samples; naturally, if we direct training change detection model in this dataset, it will result in deep learning model that cannot learn enough semantic information; thus, once the trained model is applied to real scenes covered by haze or thin clouds, the change detection effect is usually unsatisfactory.

Our data augmentation method can improve the deep learning model performance in haze cover and seamline scenes. In principle, a simple haze simulation is used to augment the samples so that the deep learning model can learn enough semantic information to improve the generalization ability of the deep learning model. The implementation principle of the mosaic simulation data augmentation method is using spectral transformation to implement, this means is very simple and can easily simulate the effect of a multi-temporal image mosaic, thus improving model performance in excessive spectral differences regions.

B. *Compared With the GAN Data Augmentation Method*

As we know, GAN models are also very popular for data augmentation. Therefore, we also selected a classical sample generator, whose name is GauGAN for further comparison [45]. GauGAN is a generator model, which is composed of several residual blocks, and the building change detection experiment demonstrates that it can generate new high-quality samples [45]. The source code of GauGAN can be accessed at https://github.com/justchenhao/IAug_CDNet. Several newly generated deforestation samples are displayed in Fig. 8.

Each semantic map (dataset in Section IV-A) in the original training samples produced a new synthesized image based on GauGAN, then these newly generated images would be augmented by several operations, such as Image rotation, Scale transformation, and Brightness transformation; finally, these new images mixed up with the original sample images for the model training step. The quantitative accuracy (dataset in Section IV-C) between GaunGAN and the proposed data augmentation method was given in Table V. We can see that the GauGAN method also achieved a relatively satisfactory change detection result. However, there are several limitations in data augmentation methods based on GAN, such as the model training progress being very time-consuming.

C. *Experiments on the Haze-Free Scenario*

How does the deep learning model perform in haze-free scenarios? We also conduct an experiment to compare. It is worth

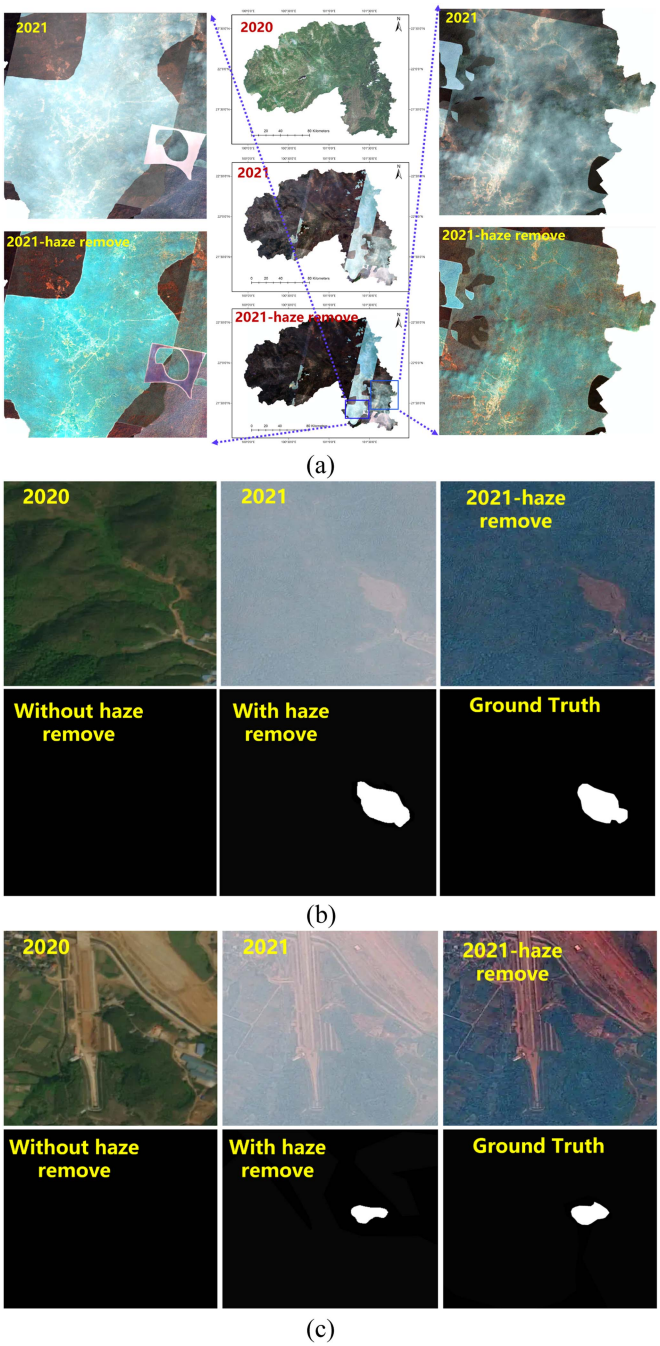

Fig. 9. Change detection in the haze-free scenario. (a) Haze remove images. (b) and (c) Change detection result based on Method 2 in haze cover scenario and haze-free scenario.

noting that data augmentation Method 2 in Table I was selected to train a change detection model. To remove the haze or thin cloud in Fig. 3, we used the dark channel prior algorithm to generate a haze-free scenario [37]. The original image, haze-free image, and several change detection results are displayed in Fig. 9. The quantitative accuracy assessment is given in Table VI.

In Fig. 9, we can see that change detection results on haze removal scenarios has been improved, and the accuracy assessment in Table VI also demonstrates this point. This auxiliary

TABLE VI
ACCURACY EVALUATION

Method	Recall	Precision	F1	OA
Method 2 (haze cover scenario)	0.6014	0.4017	0.4284	0.9811
Method 2 (haze-free scenario)	**0.6314**	**0.4191**	**0.4291**	**0.9821**

TABLE VII
ACCURACY EVALUATION

Data augment methods	Recall	Precision	F1	OA
Method 1	0.7221	0.9883	0.7469	0.8227
Ours (haze simulation)	**0.8856**	0.9867	**0.7623**	**0.9307**

experiment gives us some useful inspiration that developing excellent haze or fog removal algorithms for optical high-resolution images is important to improve the change detection model performance in these complex scenarios. However, in this article, we are not aiming to design a new or state of art haze removal algorithm, we want to improve model performance only based on training samples, and the final experiment proved that it was effective.

D. Experiments on Other Datasets

In order to test the universality of the proposed data augmentation methods, we also conducted a road extraction experiment to verify. The road extraction task does not require bitemporal images, only one time-phase image is needed, which provides a good reference for other pixel-level segmentation tasks, such as building extraction [46], [47].

D-Linknet model was selected as our baseline model [48]. The training dataset was downloaded from the official website (http://deepglobe.org/leaderboard.html). To verify the impact of different data augmentation methods on the accuracy, we use the Method 1 data augmentation method as a comparison, which was mentioned in Section IV. In order to compare the extraction difference between different data augmentation methods under haze cover conditions, we also used the haze image simulation method to generate new test samples. The newly generated haze cover images and road extraction results are displayed in Fig. 10.

In the visual comparison aspect, the trained road extraction model based on our data augmentation method is superior to the traditional data augmentation methods, such as spectral transformation and scale transformation. Interestingly, the road extraction model obtained by using our data augmentation method is effective in reducing the omission alarms to a certain extent, such as in the haze coverage areas. Therefore, it also proves that the haze image simulation method can be very effective in improving the generalization ability of deep learning models in haze-covered regions by allowing the models to pay more attention to this complex region.

We also used Recall, Precision, F1, and OA to evaluate the quantitative accuracy, as given in Table VII. Compared with

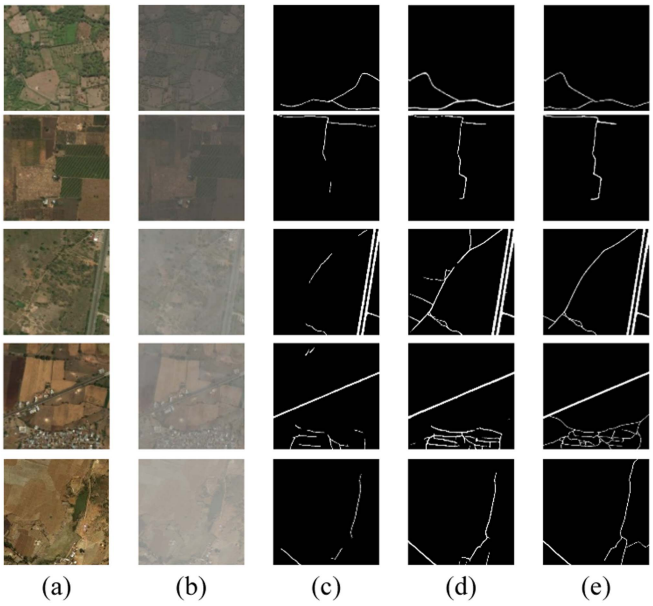

Fig. 10. Road extraction. (a) Original image. (b) Haze image simulation. (c) Road extraction with data augmentation method1. (d) Road extraction result with the proposed data augmentation method (only used haze simulation). (e) Ground truth.

Method 1, our Recall indicator is relatively higher, and the road extraction result is also consistent with the realistic extraction situation.

E. Limitation

However, the proposed data augmentation method still has several drawbacks: it is only suitable for haze or thin cloud-covered areas, and there are still a few detection errors in the thick clouds cover regions because in these regions, the ground objects are invisible to conduct change detection tasks.

VI. Conclusion

Unlike the previous change detection methods on high-resolution remote sensing images, which are mostly improved from the model architecture aspect, we conducted an in-depth analysis of the training dataset and found that model accuracy is poor in the haze coverage and seamline regions. From this perspective, we propose a new data augmentation method, namely, mosaic simulation and haze image simulation, and the subsequent experiments in real-world scenes have proved the effectiveness of this method.

The main contribution of this study is that we do not explore the innovation in the architecture of the deep learning model but constructs a simple and effective data augmentation method to improve the model performance in the haze-covered and seamline regions, which provides a cheap solution to these complex scenes. In addition, our new data augmentation method is not only adapted to the change detection task but also provides an algorithm reference for other deep learning tasks, such as building change detection in haze-covered regions or object detection under specific meteorological conditions.

References

[1] T. Liu, L. Yang, and D. Lunga, "Change detection using deep learning approach with object-based image analysis," *Remote Sens. Environ.*, vol. 256, 2021, Art. no. 112308.
[2] Q. Shi, M. Liu, S. Li, X. Liu, F. Wang, and L. Zhang, "A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art no. 5604816, doi: 10.1109/TGRS.2021.3085870.
[3] C. Shorten and T. Khoshgoftaar, "A survey on image data augmentation for deep Learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
[4] H. Zhang et al., "Mixup: Beyond empirical risk management," in *Proc. 6th Int. Conf. Learn. Representations*, 2018.
[5] S.-A. Rebuffi et al., "Data augmentation can improve robustness," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 29935–29948.
[6] C. Gong et al., "Keepaugment: A simple information-preserving data augmentation approach," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021.
[7] F. J. Moreno-Barea, F. Strazzera, J. M. Jerez, D. Urda, and L. Franco, "Forward noise adjustment scheme for data augmentation," in *Proc. IEEE Symp. Ser. Comput. Intell.*, 2018, pp. 728–734.
[8] Z. Zhong et al., "Random erasing data augmentation," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 7, pp. 13001–13008.
[9] W. Li et al., "Data augmentation for hyperspectral image classification with deep CNN," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 593–597, 2018.
[10] G. Ghiasi et al., "Simple copy-paste is a strong data augmentation method for instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 2918–2928.
[11] M. Frid-Adar et al., "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.
[12] J. Wang and L. Perez, "The effectiveness of data augmentation in image classification using deep learning," *Convolutional Neural Netw. Vis. Recognit.*, vol. 11, no. 2017, pp. 1–8, 2017.
[13] E. D. Cubuk, B. Zoph, D. Mané, V. Vasudevan, and Q. V. Le, "AutoAugment: Learning augmentation strategies from Data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 113–123.
[14] J. Zhang et al., "Pise: Person image synthesis and editing with decoupled gan," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021.
[15] X. Li et al., "A deep translation (GAN) based change detection network for optical and SAR remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 179, pp. 14–34, 2021.
[16] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images," *Remote Sens. Environ.*, vol. 159, pp. 269–277, 2015.
[17] M. B. Pereira and J. A. dos Santos, "ChessMix: Spatial context data augmentation for remote sensing semantic segmentation," in *Proc. 34th SIBGRAPI Conf. Graph., Patterns Images*, 2021, pp. 278–285.
[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 25.
[19] H. Zhang et al., "Context encoding for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7151–7160.
[20] R. Gai, N. Chen, and H. Yuan, "A detection algorithm for cherry fruits based on the improved YOLO-v4 model," *Neural Comput. Appl.*, pp. 1–12, 2021.
[21] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-time instance segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 9157–9166.
[22] P. Baldi and P. J. Sadowski, "Understanding dropout," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, Paper 26.
[23] S. Santurkar et al., "How does batch normalization help optimization?," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, Paper 31.
[24] E. D. Cubuk et al., "Autoaugment: Learning augmentation strategies from data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019.
[25] Y. Xiao et al., "A new color augmentation method for deep learning segmentation of histological images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag.*, 2019, pp. 886–890.
[26] X. Li, Y. Peng, and M. Xu, "Patch-shuffle-based semi-supervised segmentation of bone computed tomography via consistent learning," *Biomed. Signal Process. Control*, vol. 80, 2023, Art. no. 104239.
[27] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Change detection. net: A new change detection benchmark dataset," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, 2012, pp. 1–8.

[28] A. Shafique et al., "Deep learning-based change detection in remote sensing images: A review," *Remote Sens.*, vol. 14, no. 4, 2022, Art. no. 871.
[29] Q. Zhu et al., "Land-use/land-cover change detection based on a Siamese global learning framework for high spatial resolution remote sensing imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 184, pp. 63–78, 2022.
[30] Z. Zheng et al., "ChangeMask: Deep multi-task encoder-transformer-decoder architecture for semantic change detection," *ISPRS J. Photogrammetry Remote Sens.*, vol. 183, pp. 228–239, 2022.
[31] G. Zhao and Y. Peng, "Semisupervised SAR image change detection based on a siamese variational autoencoder," *Inf. Process. Manage.*, vol. 59, no. 1, 2022, Art. no. 102726.
[32] X. Deng, Y. Zhang, and H. Qi, "Towards optimal HVAC control in non-stationary building environments combining active change detection and deep reinforcement learning," *Build. Environ.*, vol. 211, 2022, Art. no. 108680.
[33] M. Zajac et al., "Adversarial framing for image and video classification," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 10077–10078, 2019.
[34] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, 2018.
[35] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, and M. Song, "Neural style transfer: A review," *IEEE Trans. Visual. Comput. Graph.*, vol. 26, no. 11, pp. 3365–3385, Nov. 2020.
[36] H. Timothy et al., "Meta-learning in neural networks: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5149–5169, 2021.
[37] J. Long, Z. Shi, W. Tang, and C. Zhang, "Single remote sensing image dehazing," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 59–63, Jan. 2014.
[38] H. A. N. Jie and X. I. E. Yong, "Image dodging algorithm for GF-1 satellite WFV imagery," *Acta Geodaetica et Cartographica Sinica*, vol. 45, no. 12, 2016, Art. no. 1423.
[39] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
[40] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
[41] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Deblurring images via dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 10, pp. 2315–2328, Oct. 2018.
[42] M. Qin, F. Xie, W. Li, Z. Shi, and H. Zhang, "Dehazing for multispectral remote sensing images based on a convolutional neural network with the residual architecture," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1645–1655, May 2018.
[43] S. Fang, K. Li, J. Shao, and Z. Li, "SNUNet-CD: A densely connected siamese network for change detection of VHR Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 8007805.
[44] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 618–626.
[45] H. Chen, W. Li, and Z. Shi, "Adversarial instance augmentation for building change detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5603216.
[46] Y. Liu, C. Pang, Z. Zhan, X. Zhang, and X. Yang, "Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 5, pp. 811–815, May 2021.
[47] H. Fang, P. Du, and X. Wang, "A novel unsupervised binary change detection method for VHR optical remote sensing imagery over urban areas," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 108, 2022, Art. no. 102749.
[48] L. Zhou, C. Zhang, and M. Wu, "D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 182–186.

Zhipan Wang was born in Hengyang, China. He received the M.Sc. degree in photogrammetry and remote sensing from Southwest Jiaotong University, Chengdu, China, in 2017. He is currently working toward the Ph.D. degree in photogrammetry and remote sensing with Sun Yat-sen University, Shenzhen, China.

He was a senior program developer for Land Satellite Remote Sensing Application Center, MNR of China. His current research interest focuses on the change detection with deep learning method.

Di Liu received the M.E. degree in surveying and mapping engineering from the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China, in 2019. He is currently working toward the Ph.D. degree in aerospace science and technology with the School of Aeronautics and Astronautics, Sun Yat-sen University, Shenzhen, China.

His research interests include nightlight remote sensing, synthetic aperture radar, and artificial intelligence applications of multisource remote sensing.

Zhongwu Wang was born in Honghu, China. He received the Ph.D. degree in cartography and geography information system from the institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2009.

He is currently a Vice Researcher with Land Satellite Remote Sensing Application Center, MNR of China. His currently interests include land remote sensing, image fusion, land use and land cover monitoring, and deep learning segmentation.

Xiang Liao was born in Kaixian, Chongqing, China. He received the M.Sc. degree in aerospace science and technology from the Harbin University of Technology, Harbin, China, in 2015.

His current research interests include remote sensing cloud computing, InSAR deformation detection, and its application in southwest China.

Qingling Zhang received the Ph.D. degree in geography from Boston University, Boston, MA, USA, in 2009.

From 2008 to 2014, he was a Research Scientist with the School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA. From 2014 to 2018, he has been a Professor with the Shenzhen Institutes of Advanced Technology, as well as the Xinjiang Institute of Ecology and Geography, and Chinese Academy of Sciences. Since 2018, he has been working as a Full Professor with the School of Aeronautics and Astronautics, Sun Yat-sen University, Shenzhen, China. He is an expert in urbanization, land cover land use change, and night-time light remote sensing. His research interests include crop monitoring and yield forecasting, global environmental change, urban remote sensing, land cover land use change, food security, GIS and spatial analysis, remote sensing data mining based on cloud computing, change detection, object detection, and image processing method for night-time images.

Dr. Zhang was a recipient of the One Hundred Person Project of the Chinese Academy of Sciences Award for Excellence in 2015.