

Data Augmentation and Few-Shot Change Detection in Forest Remote Sensing

Songyu Zhu , Weipeng Jing , *Member, IEEE*, Peilun Kang, Mahmoud Emam , and Chao Li 

Abstract—Forest remote sensing change detection provides an important technical support for forest management decisions and analysis of forest disturbance factors. However, lack of data in specialized fields leads to the detection accuracy-improved difficulty. In this study, a forest remote sensing change detection model in the context of few-shot learning is proposed. The proposed model achieves end-to-end change detection algorithm for forest scenes from two perspectives of data augmentation and updated few-shot algorithm. First, forest fragments images are jointly generated by using feature extraction network and generative adversarial network. Then, forest fragments are blended into the original change detection dataset by Poisson blending method to achieve effective augmentation. Furthermore, the end-to-end change detection network is also updated using a few-shot learning. In addition, Meta-learning module is added to the slow feature analysis algorithm based on the multiattention mechanism to realize the detection effect improvements. The proposed model improves the F1 score from 86% to 91% on the two datasets. Moreover, it increases the F1 score by 6.52% on average.

Index Terms—Few-shot learning, forest change detection, remote sensing data augmentation, slow feature analysis.

I. INTRODUCTION

REMOTE sensing change detection relies on remote sensing technology to analyze the changes occurring at the same location in different periods, which greatly promotes the progress of Earth observation [1]. It can help human beings to quickly and accurately obtain information on changes in the Earth's surface features and guide human beings to scientifically develop and utilize the Earth's resources [2]. For example, the change in buildings can analyze the expansion of cities [3], the change in vegetation cover can infer the development status of the ecological environment [4], the change of roads can judge the development of urban transportation and economic

prosperity [5], etc. Remote sensing change detection in forest scenes can help in analyzing the deforestation and disasters in a certain area at a certain time, and therefore guide the forest management decisions in that area [6]. Due to long duration and slow change characteristics in forest scenes, automated change detection methods for forest remote sensing are highly required.

Deep learning-based approaches have greatly improved the efficiency of change detection [7], with significant advantages in automation and accuracy. Furthermore, deep learning networks have powerful feature extraction and classification capabilities [8], and also more sensitive for changing pixel features in change detection. In recent years, deep learning-based change detection algorithms have been attracted much attention [9], breaking through the bottleneck of traditional remote sensing technology and making great progress [10]. Generally, accurate deep learning detection models require a large and effective amount of data. The core idea behind deep learning is to learn the corresponding features with a large amount of data, summarize the specific laws for prediction, and then build a more complete prediction model.

However, many scenes in reality often lack effective training samples, such as change detection in forest scenes. The existing available forest remote sensing datasets are limited. The cost of remote sensing image acquisition is very high, and the labeling of forest information requires strong professionalism, which is often difficult to be done by nonprofessionals. In addition, unless affected by man-made disasters, forest cycle change is a very long process and vulnerable to seasonal influence, which makes it difficult to obtain effective forest change information images in a short time, which brings great difficulties to the automated forest change detection. From the current technical means, the few sample problems can be studied from two aspects: data augmentation and few-shot learning algorithms.

Data augmentation [11] expands the training dataset by increasing the amount of image samples, increases the training data, reduces the overfitting phenomenon of network training for deep learning models, and, thus, trains a network model with stronger generalization abilities. GoodFellow et al. [12] suggested that acquiring more valid data is more important than the augmentation of algorithms. Therefore, it is important to conduct research on forest image data augmentation at the data level.

On the other hand, few-shot learning algorithm improvements can solve machine learning problems with limited data [13]. Moreover, deep learning-based models with a small number of learning samples have a great potential interest in numerous remote sensing applications that lack data. Few-shot learning

Manuscript received 25 March 2023; revised 14 May 2023; accepted 7 June 2023. Date of publication 13 June 2023; date of current version 10 July 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 32171777 and Grant 32271865, and in part by the Science and Technology Program of Inner Mongolia Autonomous Region under Grant 2022YFSJ0037. (Corresponding authors: Weipeng Jing; Chao Li.)

Songyu Zhu is with the College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China, and also with the College of Electronics and Information Engineering, Harbin Vocational and Technical College, Harbin 150081, China (e-mail: zhusongyu@nefu.edu.cn).

Weipeng Jing, Peilun Kang, and Chao Li are with the College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China (e-mail: jwp@nefu.edu.cn; 374774222@nefu.edu.cn; lichaone-fuzyz@nefu.edu.cn).

Mahmoud Emam is with the Faculty of Artificial Intelligence, Menoufia University, Shibin El Kom 32511, Egypt (e-mail: memam@ai.menofia.edu.eg).

Digital Object Identifier 10.1109/JSTARS.2023.3285389

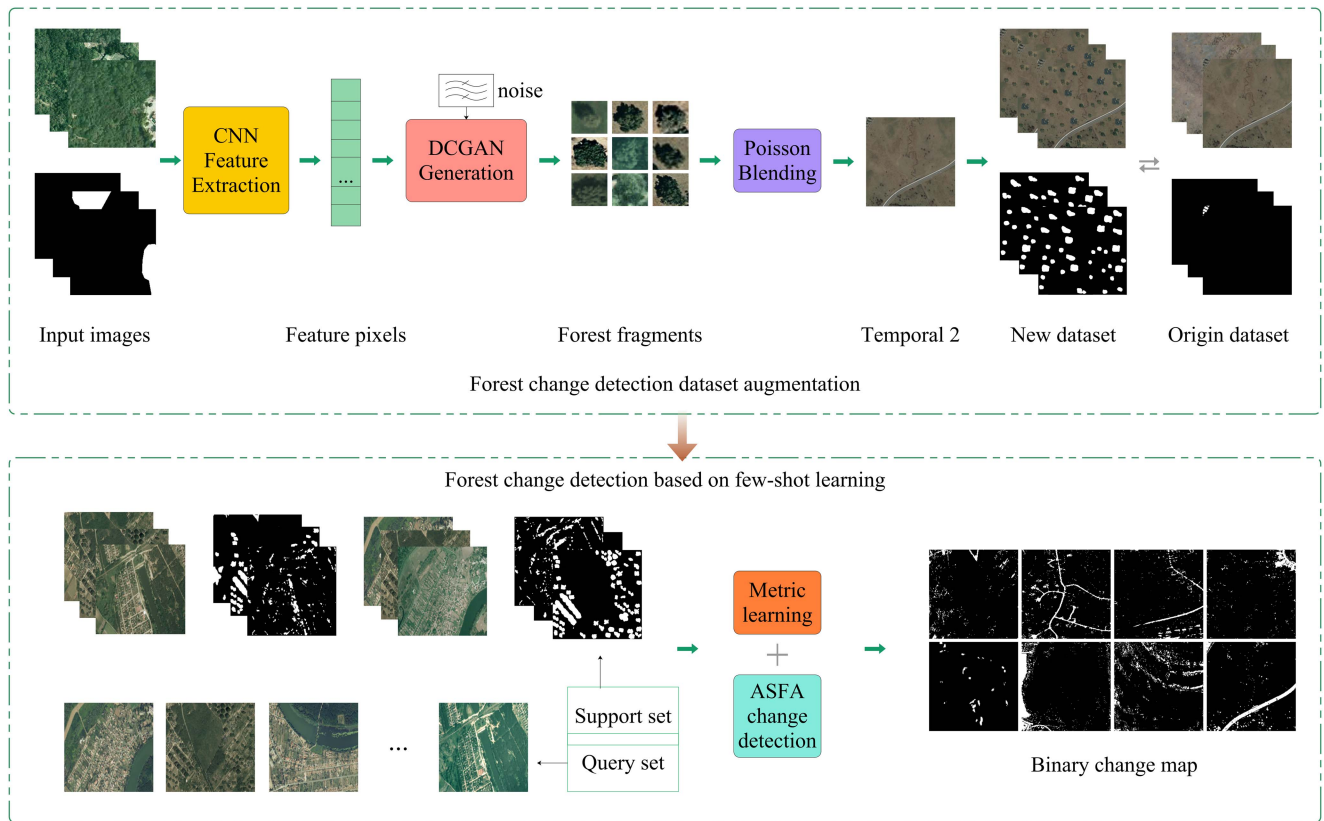


Fig. 1. Overall structure of FRSCD_DAFS. The above part is the process of data augmentation, and the below part is the process of change detection based on few-shot learning.

algorithms not only save the cost of acquiring data, but also can narrow the gap between artificial intelligence and human intelligence in providing the possibility of change detection for specific special scenarios, which is a topic that must be tackled in the development of deep learning [14].

Most of the current data augmentation algorithms are based on transformation methods of the image itself, such as cropping, rotating, flipping, stitching, or adding noise to expand the samples [15]. However, for forest scenes with little change information, such methods cannot fundamentally solve the problem of increasing the effective information for forest changes. Therefore, thinking in terms of enhancing effective change information and balancing the ratio of positive and negative samples in the image, adding effective forest change samples to the original dataset is the key problem to solve the lack of forest change detection data. Few-shot learning algorithm can be understood as letting the model learn to learn, which is also a manifestation of metalearning. Many scholars are currently conducting research under the benchmark of N-way K-shot [16]. However, for the forest remote sensing change detection problem, few researchers start to use both data augmentation and few-shot algorithm improvement at the same time. Therefore, we employ these two aspects deeply and construct a forest remote sensing change detection model based on data augmentation and few-shot learning algorithm.

In this article, a forest remote sensing change detection model based on data augmentation and few-shot learning is proposed,

namely FRSCD_DAFS. First, we apply convolution neural network (CNN) to extract deep features of forest samples. Second, the deep convolution generative adversarial network (DCGAN) is used to generate forest sample fragments images. Third, the obtained forest fragments are synthesized to the second temporal background in the original forest change detection dataset by Poisson blending to achieve effective change sample expansion of the dataset. Finally, the change detection algorithm is also improved based on few-shot learning without changing the number of images in the dataset. Effective training data and algorithms are provided for the change detection model to achieve the few-shot forest remote sensing change detection task. The overall structure of the proposed model is shown in Fig. 1. The main contributions of this study can be summarized as follows.

- 1) A complete system of forest remote sensing change detection model FRSCD_DAFS under few-shot learning is constructed, which combines data augmentation and few-shot learning algorithm to fill the gap of data and algorithm co-augmentation under forest scenario and improve the automation efficiency of forest remote sensing change detection.
- 2) A method of forest remote sensing image augmentation is proposed to increase the number of samples for forest change information, balance the number of positive and negative samples in change detection, and achieve a real and effective information expansion of the dataset with the same number of images in the dataset.

- 3) A change detection algorithm based on few-shot learning is proposed to jointly compute interclass features and intraclass features, and combine migration learning and data augmentation methods to improve the change detection accuracy of end-to-end forest remote sensing images.

II. RELATED WORK

A. Forest Change Detection Methods

Change detection methods in forest scenes have not been fully studied yet and occupy only a small part of many kinds for remote sensing change detection. The traditional detection methods are mainly relied on vegetation indices for detection [17], in addition to machine learning-based detection methods [18]. Due to the severe scarcity of data, there are also few specialized forest scene change detection models among the deep learning detection methods. Therefore, remote sensing change detection for forest scenes still needs to be focused on data and models for research and updating.

B. Remote Sensing Data Augmentation

The change detection model based on deep learning requires learning pixel features of the forest remote sensing images and the change difference features between two temporal images through a large amount of training data [19]. However, limited by the slow change of the forest scenes, less change information is acquired in the images, and even less effective data are transferred to the model for training after image cropping, which greatly reduce the generalization ability of the model, resulting in limited applicability of the model to only test environments with substantially similar training data. Therefore, the focus of effective data augmentation is to enhance the change information in the two temporal images, enhance the sample number of change instances in the training data, enhance the number of change pixels, increase the proportion of change information in the images, and avoid occurrence of “invalid training.”

To achieve effective data augmentation, we use DCGAN network to generate change instances that need to be added to the images [20]. Hence, to generate real forest instances that provide change forest samples for the dataset. The generated change instances are called forest fragments, and the forest fragment images are mixed into the temporal background images for change detection, which increases the number of change samples and achieves effective positive sample augmentation.

C. Change Detection Based on Deep Learning

Deep learning-based change detection methods are broadly divided into two main categories: one is classification method followed by detection [21], [22], [23], [24], and the other is end-to-end direct detection [25], [26]. With feature extraction capabilities of convolutional neural networks, the method of first classifying image pixels and then realizing pixel as difference is widely used [27], [28], [29]. In order to improve detection efficiency, the end-to-end detection method, which is also known as one-step change detection, has gradually attracted many researchers' attention [30], image transformation methods are

used as a mainstream [31], [32], [33], [34]. For special application scenarios with few limited data samples, the few-shot learning concept gradually occupies the hot spot, so it is an inevitable trend of development to study forest remote sensing change detection based on few-shot learning algorithms.

To solve remote sensing change detection problem, Jing et al. [35] introduced a change detection algorithm based on multiattention mechanism. First, a bag-of-words model with automatic clustering was used to preclassify the images, and the classification number was obtained as a channel parameter for the subsequent training network. Second, feature extraction was performed using a residual network, while a 3-D attention mechanism was added to compute and update full-time for the space, time, and channel of remote sensing images, respectively, to obtain more accurate and complex feature information. Finally, the high-dimensional features were passed into the slow feature analysis network, the feature values of pixels were then sorted and compared, and the extraction of change information was realized according to the theorem and threshold calculation. In this article, we construct a feature extraction network and a feature projection network, and propose a change detection algorithm based on few-shot learning, which performs well in forest remote sensing change detection.

III. METHODOLOGY

The accurate prediction effect of the deep learning model relies on a large amount of training data. Consequently, improving number and utilization rate of the samples is the key for solving the problem of few samples remote sensing change detection. The target object of change detection in forest scenes is forest communities, which are affected by human or natural factors with large changes, showing regional changes and concentrated changes under the interpretation of remote sensing images, while the changes occurring due to their growth factors are slow, and the changes are not easy to monitor, resulting in an uneven distribution of positive and negative samples in the real forest datasets. Therefore, in the naturally acquired forest remote sensing images, the detection network learns limited features and the effective training information accounts with a small percentage, which affects the accuracy of the model. Furthermore, to improve the effectiveness of training information, the coverage of positive samples to the whole remote sensing image should be increased. In addition, forest image features should be fully extracted by means of deep learning characteristics to effectively detect forest scene changes in the context of few-shot data.

This section will illustrate the proposed method of few-shot data augmentation in conjunction with the practical application of forest scene change detection, using the adjusted DCGAN network to generate new sample data to form forest image fragments. Then, we fuse the fragments with the temporal images in the forest change detection dataset to reconstruct a change detection dataset with large number of valid samples, which is a good data foundation for subsequent change detection tasks. The overall work is divided into the following three main steps.

- 1) *Forest image fragment generation*: The main focus of data augmentation is to increase the diversity of forest image

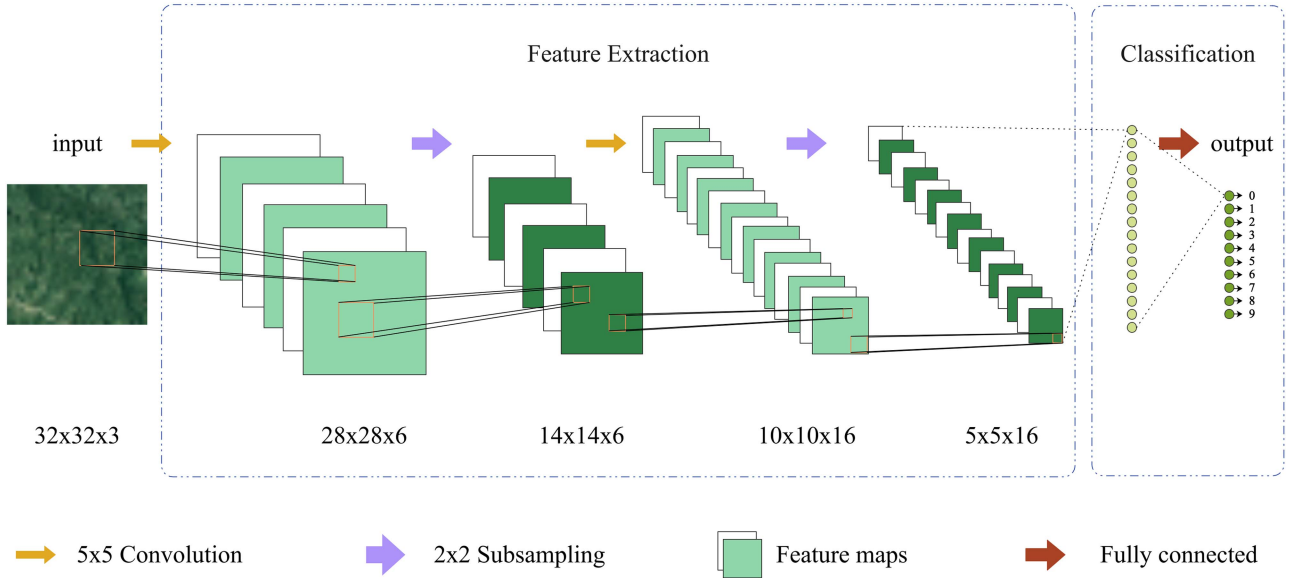


Fig. 2. Structure of forest feature extraction network. The left is the generator network and the right is the discriminator network.

forward samples. We refer to the generated new forest forward samples as forest fragments. Considering the generalization ability of the final change detection model, the diversity of forest image fragments needs to be evaluated. We employ an improved DCGAN network to counteract the generation of forest fragments with diversity and render their color edges by using the improved color transfer (CT) algorithm to enable better fusion reconstruction with the real dataset and avoid inharmonious.

- 2) *Reconstruction of fragments and temporal images*: The generated forest fragments enhance the number of positive samples in change detection, and by fusing these fragments with the real change detection dataset, the coverage of positive samples can be improved and the number of valid information is extended. We use Poisson blending for fusing forest fragments and background data to construct new few-shot dataset, which is suitable for forest scene change detection tasks.
- 3) *Change detection network*: The change detection in forest scenes is one of the important applications for Earth observation. We tend to improve the change detection algorithm proposed by Jing et al. [35]. We improve the slow feature analysis method by using few-shot learning and using the newly generated change detection dataset for change prediction and analysis.

A. Forest Image Fragment Generation

The generation of forest image fragments is achieved by a feature extraction network and a GAN. We use a CNN network to achieve diverse pixel feature extraction of the forest [36], and the extracted features are projected and passed to the discriminative network in the DCGAN for forest fragmentation generation feedback as decision discriminative information [37]. In addition, to increase the diversity of forest fragments and the

robustness of the model, the constraint of forest fragment area is added to make the model converge in a controlled range.

1) *Forest Image Feature Extraction*: In order to avoid the gradient explosion phenomenon of GANs [38], a convolution-based feature extraction network is added to the generative image first. The essence of convolutional neural network is to fit a more complex function by layer simple activation function, and adjust the training parameters by the feedback information during the fitting process, so that the cost function is minimized and the optimal strategy is obtained. Convolutional neural networks have a strong ability of representation learning and translation invariant for the classification of pixels [39]. In the forest feature extraction network, the normalized 3-D pixel data, i.e., 2-bit pixel points on the plane and RGB channels, are first passed to the input layer. Then, the feature maps are processed by convolutional, pooling, and fully connected layers. Finally, the feature classification of pixels is obtained by the logic function of the output layer. The structure of the forest feature extraction network is shown in Fig. 2.

The convolution part is calculated, as shown in (1). The summation part is equivalent to the primary cross-correlation. Here, b is the deviation amount, Z^l and Z^{l+1} denote the input and output of the convolution in the l and $l+1$ layers, respectively, also known as the feature map, L_{l+1} is the size of Z^{l+1} , and it assumes that the feature maps are of the same length and width, $Z(i, j)$ and K , respectively, denote the pixels of the feature map and the number of channels of the feature map, and finally, f , s_0 , and p are the parameters of the convolution, which are corresponding to the kernel size, the stride, and the padding layers, respectively.

$$\begin{aligned}
 Z^{l+1}(i, j) &= [Z^l \otimes w^{l+1}](i, j) + b \\
 &= \sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [Z_k^l(s_0i + x, s_0j + y) w_k^{l+1}(x, y)] + b
 \end{aligned}$$

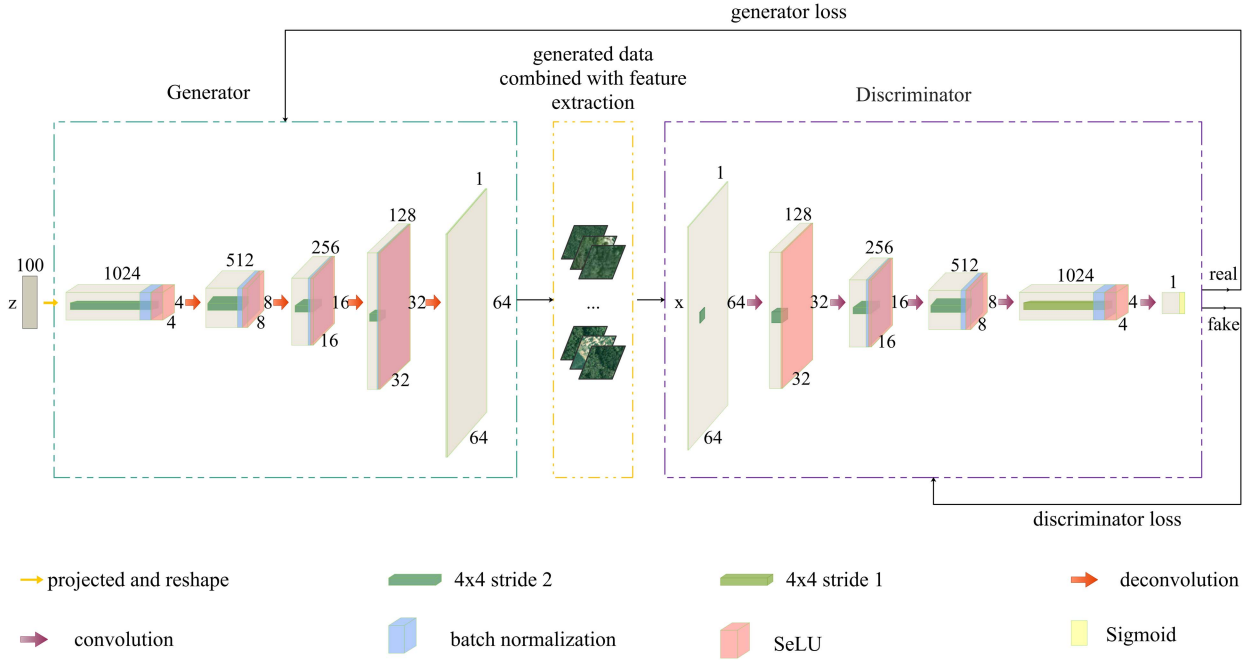


Fig. 3. Structure of the improved generate network.

$$(i, j) \in \{0, 1, \dots, L_{l+1}\}$$

$$l_{l+1} = \frac{L_l + 2p - f}{s_0} + 1. \quad (1)$$

2) *Forest Fragment Generator*: A GAN is a powerful network for generating real images, which can generate new images outside the training dataset and similar to the distribution of the training set with the help of random noise as an input.

GAN networks train both generators and discriminators [40], [41], which are very suitable for few-shot forest change detection scenarios where training data are limited and lacking the network performance. Among many variant networks, the performance of DCGAN has been proven to be stable and excellent [42]. Therefore, we select the DCGAN network as the backbone network for generating forest fragments images.

To avoid the repeated blurring samples of the traditional DCGAN network model, we adjust the activation functions of the generator and discriminator by choosing SeLu function [43] for computation. SeLu function retains the computational part of the input value less than 0 when completing nonlinear activation of the convolutional layer and preserves the richer pixel features. The SeLu activation function can automatically correct the cases of different deviations occurring in the neural network, and the network model can self-adapt to the changes in the input and improve the overall robustness of the model.

The expression of the SeLu function is shown in (2). When the parameters λ and α are set to $\lambda \approx 1.0507$ and $\alpha \approx 1.6733$ [44], the distribution of the output of each layer will converge to the standard normal distribution under the condition that the model weights obey the standard normal distribution. This adaptive property can avoid the problem of gradient disappearance or

explosion when generating forest fragments.

$$f(x) = \lambda \begin{cases} \alpha (e^x - 1), & x \leq 0 \\ x, & x > 0. \end{cases} \quad (2)$$

Since we have extracted features from the forest samples in the source dataset before the generative model, it provides more detailed features for the discriminator's discrimination in the GAN network, which makes the relationship between the generator and the discriminator more exciting and rigorous, forcing the generator to generate more realistic images and putting higher quality requirements on the whole generative model. The structure of the improved DCGAN model is shown in Fig. 3.

The random noise is passed into the generative network, and the feature map is obtained through projection and reshape operations in the fully connected layer, followed by the multilayer deconvolution algorithm upsampling to generate the final image $G(Z)$. In the discrimination network, the generated image is first entered, downsampled by convolution, processed in the fully connected layer, and then fed into the SeLu function to discriminate the probability of the image from being true or false. When the probability approaches 0.5, the forest fragment image is considered to be up to standard.

The loss functions of each of the generation and discrimination are shown in (3) and (4). Here, m is the number of samples, $D(x)$ represents the probability of real images, $G(Z)$ represents the false images generated by the generator, and $D(G(Z))$ represents the probability of the generator generating false images.

$$\frac{1}{m} \sum_{i=1}^m [\ln(1 - D(G(Z^{(i)})))] \quad (3)$$

$$\frac{1}{m} \sum_{i=1}^m [\ln D(X^{(i)}) + \ln(1 - D(G(Z^{(i)})))] \quad (4)$$

3) *Color and Style Transfer*: To have diversity in the generated forest fragment images, the selection of the source dataset takes into account the color change of the forest with the seasons, so the color of the background image needs to be considered when synthesizing the new dataset. This part will converge the color and edge of the generated forest fragments toward the pixels of the background image, making the newly generated image looks more realistic and complete.

We employ the CT method as introduced in [45] and [46]. Furthermore, we consider the forest fragment as the target image and the background image as the source image. Consequently, we intend to transform the color and style of the forest fragment into a fragment that is consistent with the background image. Since the original algorithm computes the transformation matrix for the whole target image, while in our scenario, the target of the matrix computation is set within the forest fragment, that is, only the forest sample is computed and hence the image is decoded. In addition, when we train the GAN, the labels of the aged forest fragments are generated correspondingly. Therefore, the actual range of the fragments can be easily controlled.

The CT algorithm takes a statistical view of the pixel values of the image as 3-D random variables and the image as a collection of vectors. The mean of the covariance of the 3-D components is calculated, and the covariance matrix is decomposed to obtain the rotation matrix. The covariance is used to calculate the correlation between the 3-D components. Finally, the pixel data of the target image are spatially translated to obtain the target forest fragments with similar colors.

The process of blending is shown in (5). Here, I is the color of the final result, w_{ij} are the weights calculated based on the distance between the pixel data and the observation center, and a_i are the weights that can be adjusted [45]. The I_j labels the data of the intermediate image or the response to a separate target sample, and the parameters m and n are the number of samples in the source and target images, respectively. In the experiments, m and n have the same value.

$$\begin{aligned} I &= (R, G, B)^T \\ I_j &= (R_j, G_j, B_j)^T \\ I &= \sum_{i=1}^m \sum_{j=1}^n a_i \cdot w_{ij} \cdot I_j \\ \sum_{i=1}^m a_i &= 1, a_i \geq 0. \end{aligned} \quad (5)$$

B. Reconstruction of Fragments and Temporal Images

The change detection dataset contains two temporal image sets, and its detection results without considering semantic information are not subject to the image order of the image set, so the change detection model has detection symmetry during the training. In this part of the reconstruction process, we fuse the generated forest fragments into the latter-period

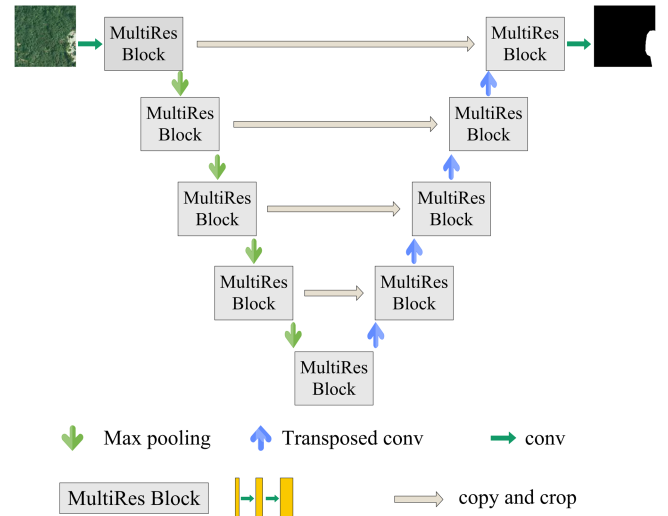


Fig. 4. Structure of MultiRes-Unet.

images of the real change detection dataset by the Poisson hybrid algorithm. Before fusing the fragments it is necessary to determine where the fragments appear to prevent them from obscuring the entities on the original dataset and to complete the fusion of the fragments at suitable locations. The detection effect of few-shot image augmentation is verified according to the constraint of the fused fragment area.

1) *Composition Position Selection*: The whole idea is to be sure that the generated forest fragments are randomly distributed on the background image homogeneously. In that case, the fragments appear in any corner of the image. Here, we combined the function of random and uniform to determine the location range. The training unit, which was cut to unified size, could still obtain more positive training samples.

Another important principle in the selection of the location for the forest fragment synthesis is not to obscure the original physical objects on the background image. To achieve this purpose. We can annotate the background image and generate a label map by semantic segmentation model. When sampling the positions in the image uniformly, the labeled positions are avoided and the updates are recorded until all positions on the image are sampled. We use the mature and simple semantic segmentation model MultiResUnet [47] for label annotation, and its model structure is shown in Fig. 4.

The MultiRes block replaces the block in the Unit model with the MultiRes Block [47]. The MultiRes Block reduces the semantic gap between encoder and decoder features by passing the encoded features through a series of convolutional layers. Residual connections are also introduced and the number of convolutional layer filters within the block is controlled.

2) *Image Composition*: At this point, we obtain the appropriate positions of the forest fragments in the foreground and the background images, and then, we fuse them utilizing Poisson blending [48], which is a computational method based on equation solving and has an excellent performance in image fusion. By optimizing the equations, the boundaries of the forest fragments and the background images are organically blended

to maintain similar brightness and gradient for image synthesis. In the 3-D Cartesian coordinate system, the Poisson equation is expressed as follows:

$$f(x, y, z) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \varphi(x, y, z) \quad (6)$$

where ∂ is the partial derivative, and f and φ are functions of real or complex values on manifolds. The left-hand side operator is called the Laplace operator, and if the right-hand side of the equation is 0, i.e., the regional dispersion is 0, this equation becomes a chi-squared equation called the Laplace equation.

The gradient of asymptotics can be slowed down by restricting the Laplace operator [49], [50]. The scalar function on the forest fragment region can be uniquely determined by the boundary values and the internal Laplace condition, which means that the Poisson equation must have a unique solution. The gradient problem can be solved by interpolation methods and function derivation on the three-channel image as a way to guarantee consistent brightness between the forest fragments and the background image and to make the generated image more natural.

The amount of forest fragments added to the background image is bounded by the coverage area of the pixels of the forest fragments. Experiments have shown that the best detection is achieved when the coverage area of the forest fragments reaches 50% of the whole, i.e., when the positive and negative sample avenues are balanced.

C. Change Detection Network

According to the characteristics of the constructed forest scene change detection dataset, we use the slow feature analysis change detection algorithm with few-shot learning theory combined with multiattention mechanism to update the original change detection model. Based on the few-shot data, the slow change pixels have smaller feature values according to the slow feature analysis theory, and the change pixels with higher feature values are obtained by combining with the threshold algorithm. The feature classification extraction and change detection are realized simultaneously to build an end-to-end detection system.

1) *Few-Shot Learning*: Allowing neural networks to get accurate classification results by learning only a small number of training samples is a challenging task for many deep learning models. Few-shot learning enables the use of prior knowledge to quickly generalize the models with new tasks where sample information is limited or scarce.

There are two sets in the few-shot learning structure [51], which are support set and query set. The support set hosts a small number of new labeled category samples, which are used by the pretrained model for generalization, whereas query set holds samples of new categories and known data, and the model needs to be generalized using prior knowledge and information learned in the support set [52], [53]. The number of new categories is denoted by N , and the labeled samples available in each category are denoted by K , constituting an “ N -way K -shot” learning scheme. When the number of samples is small, the model often needs to resort to metalearning or transfer learning to complete the training. When N increases, the difficulty of the classification

task increases and the accuracy of the prediction decreases, whereas if K is increased, the expected risk will be reduced and better results will be obtained for metric learning [54]. In addition, when the value of K is small, it is more suitable for the implementation of metric learning methods [55].

In this article, we apply the idea of few-shot learning to the change detection of forest scenes. According to the scenes in our dataset, we set N to 4 and K to 20 in the experiments. Meanwhile, we add feature extraction module and feature projection module to the CNN model, which are corresponding to the interclass features and intraclass features of a small amount of sample data in change detection, respectively, to fully exploit the performance of the migration learning [56], [57], [58], [59]. Following the scheme of few-shot learning, the sample dataset is divided into labeled support dataset and unlabeled query dataset. The generalization ability of the model is improved by extracting deep features and other additional features. Furthermore, metric learning is performed on the two sets to calculate the distance similarity between similar points in the postprojection space. By comparing the similarity, the approximate classes in the query dataset are obtained. Then, the optimized feature information is passed into the slow feature analysis network with multiattention mechanism for change detection. The structure diagram for few-shot learning is shown in Fig. 5.

The formula for the goal of metric learning is shown in (7), where W and B denote intraclass features and interclass features, respectively. The objective function sums the intraclass distances, interclass distances, and the relative distances of these two distances.

$$D = \lambda_1 \Lambda_1(d^W) + \lambda_2 \Lambda_2(d^B) + \lambda_3 \Lambda_3(d^W, d^B). \quad (7)$$

The backbone of the CNN feature extractor consists of four convolutional blocks and two maximum pooling layers, where each convolutional block consists of one convolutional operation, one batch normalization process, and the ReLU activation function [60]. The feature extraction module consists of four bottleneck layers, each containing three 3×3 convolutions, three batch standardization processes, and a ReLU function. The feature projection module consists of three bottleneck layers and one convolutional layer.

2) *Multiattention Slow Feature Analysis*: The slow feature analysis algorithm provides an end-to-end change detection method, and we add a multidimensional attention mechanism [61] to the base model to learn better feature expressions in space–time and channel to improve change detection. In this article, the change detection model proposed in [35] is further improved by using a few-shot learning approach for preclassification, followed by slow feature analysis change detection with an attention mechanism.

The optimized features are obtained by the classification of few-shot learning, and the number of categories continued to be calculated as the channel parameters. The size and channel of the image are reacquired according to the classified category number, and the channel attention weight, temporal attention weight, and spatial attention weight of the image are calculated, respectively, and the extracted complex features are used as the input of the slow feature analysis network.

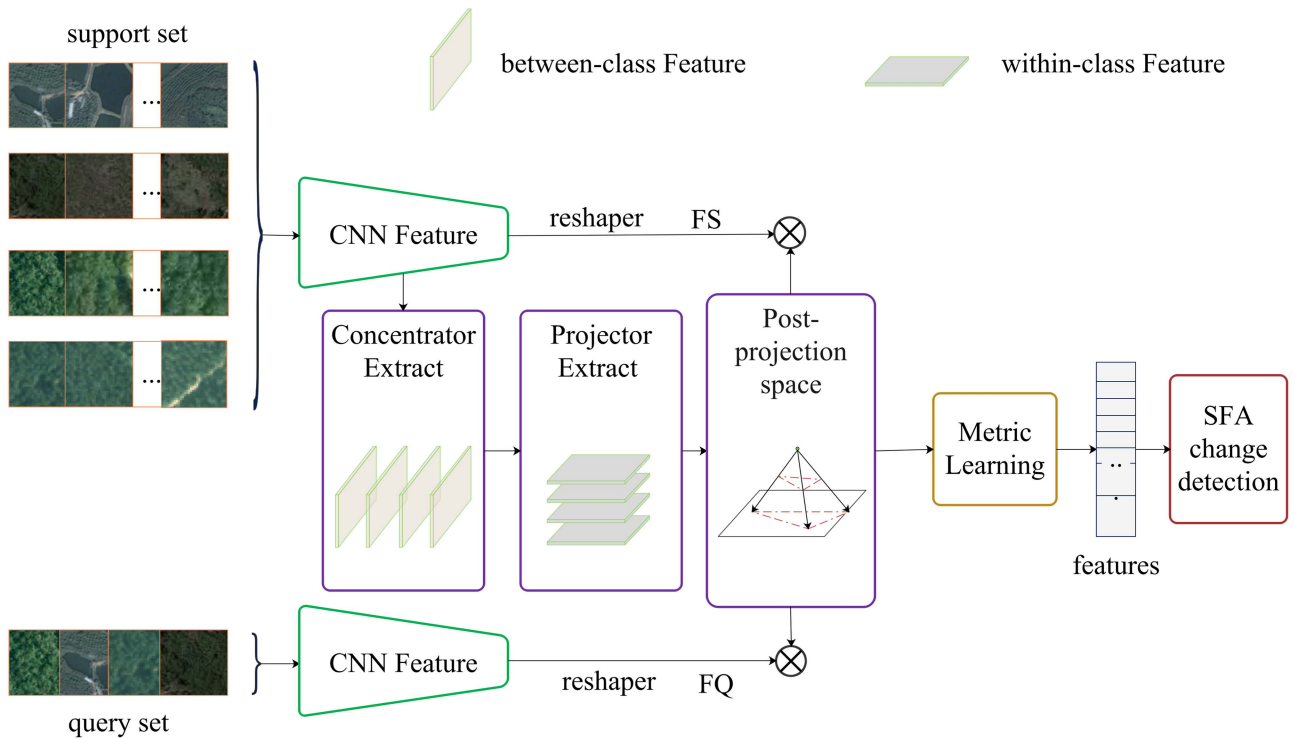


Fig. 5. Structure of few-shot learning.

The attention mechanism focuses more on a certain local feature in the global context and focuses on calculating the probability of the attention distribution for key vectors and calculating the weight value of local feature attention. In addition, the attention mechanism helps in selecting the important information to be passed to the network and selectively trains the information to reduce the network burden. The spatial and channel-based attention mechanisms can be integrated into any convolutional neural network for training to improve network performance [62]. By combining attention mechanisms in both spatial and temporal dimensions, attention weights can be computed more efficiently. In remote sensing images, the channel attention mechanism focuses on the likelihood of useful information, whereas the spatial attention mechanism focuses on the location of useful information [35].

According to the theory of slow feature analysis, pixels with small eigenvalues change slowly or even invariably. By suppressing such invariant components and minimizing their eigenvalues, the change components, which are the pixels that change, are highlighted in reverse. Finally, the change map is obtained by a threshold segmentation.

IV. EXPERIMENTS

We select four real remote sensing image datasets containing forest scene changes and classify the tasks according to scale and size of the images to form two change detection datasets and two forest debris generation datasets. The effectiveness of the proposed model is verified by comparing the evaluation indices of the experiment and the ablation study. The presentation diagrams of some images from the datasets are shown in Fig. 6.

A. Origin Dataset for Change Detection

The remote sensing change detection dataset with labels for forest scenes is very scarce, and the dataset with a larger proportion of forest scenes is selected as the experimental dataset to be used as much as possible in the existing public dataset. We select the SZTAKI [63] and LEVIR-CD [62] datasets as the original change detection datasets.

The SZTAKI AirChange Benchmark dataset [63] is an optical remote sensing aerial image of ground change over 23 years completed by Csaba Benedek and Tamás Szirányi (MTA SZTAKI). It contains 13 pairs of two-time images and change reference maps. The resolution of the images is 1.5 m and the size is 952×640 . The change information in the dataset includes new urban areas, large tree planting, barely arable land, and building changes.

LEVIR-CD is a new large-scale change detection dataset released in 2020 by Chen et al. [62] at the LEVIR lab of Beihang. All images are acquired by Google Earth. It contains 637 pairs of ground changes over 16 years in 20 areas of Texas, USA. The images have a resolution of 0.5 m and a size of 1024×1024 . The dataset is annotated with a large number of land use impacts of building growth, including changes in ground vegetation.

Both datasets contain vegetation cover and effects of various changes on forest vegetation. These two datasets are suitable as raw datasets for data augmentation of forest scene change detection.

B. Dataset for Forest Generator

The forest fragments are generated by a modified DCGAN network, and the training of the model requires a large amount

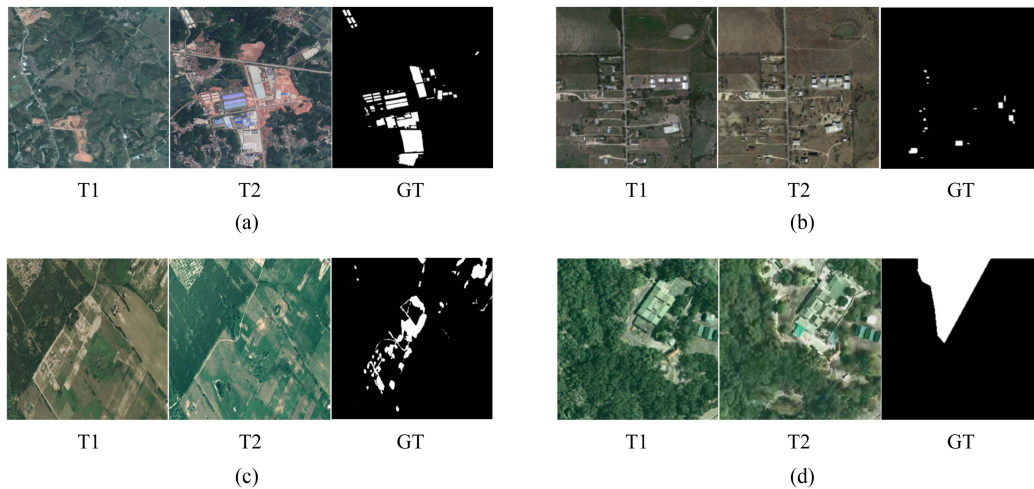


Fig. 6. Presentation diagrams of some images from the datasets. (a) SZTAKI. (b) LEVIR-CD. (c) CD_Data_GZ. (d) SYSU-CD.

of labeled forest scene data, but it is difficult to find datasets with individually labeled forest scenes in existing publicly available datasets. We redirect our focus to the change detection dataset with labels, and we find that: in the dataset containing forest changes, the prechanged images and labeled reference maps can form a set of labeled data with masks of the changed part of the forest scene. Therefore, we choose two change detection datasets (i.e., CD_Data_GZ [64] and SYSU-CD [65]) to reorganize the first moment images and labeled maps into a training set for the generation of forest fragments, respectively.

The CD_Data_GZ dataset [64] is a 20-pair high-resolution image of the suburbs of Guangzhou, China, acquired by Peng’s team from Google Earth. It contains images for changes from 2006 to 2019. The images have a resolution of 0.55 m and vary in size from 1006×1168 pixels for the smallest image to 493×5224 pixels for the largest. Changes in forest, farmland, and buildings are recorded in the images, while the reference map mainly shows the changes in buildings. A large number of images are transformed from forest scenes to building scenes, and their binary reference maps are corresponding to forest changes, providing training datasets for forest fragment generation.

The dataset SYSU-CD was established by a team from Sun Yat-sen University that captured the development of Hong Kong between 2007 and 2014. It includes a total of 800 pairs of images for urban architecture, road expansion, and vegetation changes. A total of 20 000 image pairs with a resolution of 0.5 m with size 256×256 are generated after flip-rotation. According to our task requirements, some images, such as marine construction and road expansion, are excluded by manual screening, and only data samples of forest changes are selected among them. A total of 7500 pairs are employed (4500/1500/1500 pairs for training/validation/testing, respectively).

C. Synthesis Training Samples

Since change detection is symmetric, that is, changing samples in either temporal phase 1 or temporal phase 2 has no effect

on the results of change detection, we choose to synthesize the forest fragments on the second temporal background of the change detection dataset. According to the size and spatial resolution of the dataset and other factors, we set the strategy of synthesizing forest fragments with background images. The forest fragments generated by CD_Data_GZ are synthesized onto the second temporal background of the SZTAKI dataset. The forest fragments generated by SYSU-CD are synthesized onto the second temporal background of LEVIR-CD.

The number of forest fragments to be synthesized is also a factor to be considered. Our aim is to make the positive and negative samples as balanced as possible to obtain more valid training data, so too few fragments or too many fragments will affect the final detection results. We set the constraint of forest fragment synthesis, i.e., the coverage percentage of forest fragments. We set up three sets of experiments, and since the change detection dataset inherently contains a small amount of forest changes, the coverage percentages of forest fragments are set to 15%, 35%, and 50%, respectively.

The SZTAKI dataset is cropped to 120 pairs of images (72/24/24 pairs for training/validation/testing, respectively). The LEVIR-CD dataset is cropped to 10 192 pairs of images (7120/1024/2048 pairs for training/validation/testing, respectively) according to the distribution of the original dataset. The CD_Data_GZ dataset is cropped to 3349 pairs (2009/670/670 pairs for training/validation/testing, respectively). Finally, the SYSU-CD dataset is cropped into a total of 7500 image pairs (4500/1500/1500 pairs for training/validation/testing, respectively). The specific dataset parameters are given in Table I.

D. Experiment Setup

We set up eight experimental strategies as ablation experiments to verify the function and importance of each component in the model separately. The four important components in the model include: feature extraction network, forest fragment synthesis network, few-shot learning network, and multiattention

TABLE I
DESCRIPTION OF DATASET PARAMETERS

Dataset	pairs	resolution(m)	size	patches	training / validation / test
SZTAKI	13	1.5	952 × 640	256 × 256	72 / 24 / 24
LEVIR-CD	637	0.5	1024 × 1024	256 × 256	7120 / 1024 / 2048
CD_Data_GZ	20	0.55	diversity	256 × 256	2009 / 670 / 670
SYSU-CD(forest)	7500	0.5	256 × 256	256 × 256	4500 / 1500 / 1500

mechanism of slow feature change analysis network. In setting up the experimental strategies, the first three components are arranged and combined to be validated on two datasets, and the model with the best results is selected for comparison with other methods.

To verify the performance of the proposed model and the change detection effect, we choose the mainstream end-to-end detection methods based on image transformation for comparative evaluation, including CVA, PCA, MAD, SFA, and DSFA. As representative methods of unsupervised end-to-end change detection, they are verified to be effective in an amount of applications relied on image transformation theory. The values of forest debris coverage are also verified by setting three sets of experiments with forest debris coverage up to 15%, 35%, and 50%, respectively. Each set of experiments is realized by two datasets.

For all experiments, we use the Pytorch framework. The incoming data are 256 × 256 pixel images, which are trained by momentum stochastic gradient descent [66]. In these models, the weights and bias matrices of each layer are randomly initialized, the initial learning rate is set to 0.01, and the decay coefficient is set to 0.9 [67]. After completing 100 training sessions, the learning rate is reduced to 0. The momentum and weight decay are set to 0.99 and 0.0005, respectively [68]. The batch size is set to 8. Finally, we save the best model as the final result.

E. Evaluation and Metrics

To evaluate the performance of the proposed model quantitatively, we select the following representative metrics [35], including Mean_IoU (MIoU), precision rate, recall rate and F1 score, as shown in the following:

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \left[\frac{TP}{FN + FP + TP} \right] \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (11)$$

where TP and FP refer to true positives and false positives, and TN and FN refer to true negatives and false negatives, respectively. MIoU reflects the degree of coincidence between the predicted change map and the reference one. The higher the MIoU value, the more accurate the model prediction. The precision rate represents the samples predicted to be positive,

whereas the recall rate considers the positive examples predicted correctly. F1 score is a comprehensive evaluation index, which is the harmonic average of Precision and Recall. F1 score considers both precision and recall, and it measures the model comprehensively when they are contradictory.

F. Experimental Results and Analysis

The proposed few-shot forest remote sensing change detection model contains four main components: forest feature extraction network—CNN, fragment generation and synthesis network—GP, few-shot learning network—FSL, and the improved multiattention mechanism change detection—ASFA. In order to test and verify the function and importance of each component, the relevant components are arranged and combined, and eight experimental strategies are set. Among them, the constraint of fragment synthesis data augmentation is set to 50%, and the experiments are conducted on two datasets (LEVIR-CD and SZTAKI). The experimental results are given in Tables II and III. The best performances are highlighted with bold face.

The validation of the two datasets leads to similar conclusions. The first set of experiments is applied to the ASFA model only, and the values of its evaluation metrics are used as reference values for the other seven sets of experiments. In the second set of experiments, a feature extraction network is added, and the forest features are directly entered to the change detection model. The results show that the training has an overfitting problem with unsatisfactory accuracy. The third group of experiments adds a forest fragment generation and synthesis network, and the results show that the data augmentation can improve the overall detection accuracy. The fourth group of experiments uses only the detection strategy of few-shot learning, and the results show that the improved detection algorithm improves the overall model, but not as significantly as data augmentation. In the fifth group of experiments, the forest fragmentation generation network is improved by feature extraction, and the synthetically enhanced data provides an effective information to the detection model and obtains a large improvement in accuracy. The sixth set of experiments extracts the forest features by convolutional networks and passes them to the few-shot change detection, which also has the risk of overfitting as the second set of experiments. The seventh set of experiments combines data augmentation and few-shot algorithms to achieve better detection accuracy. The eighth group of experiments contains all the components to form a perfect forest change positive sample augmentation network and few-shot change detection algorithm, which improves the F1 index significantly under the condition of balancing positive and negative samples, and also

TABLE II
RESULTS OF THE EXPERIMENTS COMPONENTS ON LEVIR-CD DATASET

No	CNN	GP	FSL	ASFA	MIoU	Precision	Recall	F1
1	0	0	0	1	0.75436	0.81722	0.90748	0.85998
2	1	0	0	1	0.31842	0.35479	0.75649	0.48303
3	0	1	0	1	0.77801	0.84588	0.90652	0.87515
4	0	0	1	1	0.76943	0.79189	0.96445	0.86969
5	1	1	0	1	0.78587	0.84625	0.91677	0.88010
6	1	0	1	1	0.34689	0.38138	0.79322	0.51510
7	0	1	1	1	0.79681	0.83671	0.94353	0.88692
8	1	1	1	1	0.83552	0.87267	0.95152	0.91039

TABLE III
RESULTS OF THE EXPERIMENTS COMPONENTS ON SZTAKI DATASET

No	CNN	GP	FSL	ASFA	MIoU	Precision	Recall	F1
1	0	0	0	1	0.75064	0.77741	0.95614	0.85756
2	1	0	0	1	0.33893	0.37202	0.79213	0.50627
3	0	1	0	1	0.78078	0.83432	0.92405	0.87689
4	0	0	1	1	0.77372	0.79616	0.96485	0.87243
5	1	1	0	1	0.80030	0.84865	0.93355	0.88908
6	1	0	1	1	0.51062	0.55420	0.86653	0.67604
7	0	1	1	1	0.80917	0.85214	0.94133	0.89452
8	1	1	1	1	0.85027	0.87576	0.96690	0.91908

improves the detection performance. Overall, the contribution of GP components in the ablation experiment is more significant from the results. F1 score increases by 2.01% on average after data augmentation. The results also prove that four components need to cooperate and win-win to obtain higher performance.

The experimental results show the model performance of the four important components work together to achieve data-enhanced change detection, and demonstrate that the proposed data-enhanced few-shot multiattention slow feature analysis change detection (FRSCD_DAFS) has better detection results on the forest scene dataset. In the following, the algorithm comparison with other image transforms is developed using two datasets, and the effect of forest fragment coverage percentage on the results is verified. The validation results of the comparison experiments using the two datasets (LEVIR-CD and SZTAKI) are given in Tables IV and V. Furthermore, The change map on the two datasets are shown in Figs. 7 and 8.

To verify the effect of forest fragment cover percentage on the results, the experiments are selected to cover 15%, 35%, and 50% for the two datasets. The comparison graph of the dataset with data augmentation is shown in Fig. 9. The experimental results are given in Tables VI and VII.

Generally, the best results are obtained at a forest fragment coverage of 50%, which is compatible with the theoretical results. In addition, we test the performance of the data augmentation model applied to several image transformation detection

TABLE IV
RESULT OF CHANGE DETECTION EXPERIMENT ON LEVIR-CD DATASET

Method	MIoU	Precision	Recall	F1
CVA	0.67347	0.76708	0.84666	0.80491
PCA	0.70283	0.78858	0.86610	0.82552
MAD	0.71513	0.79308	0.87925	0.83394
SFA	0.72652	0.80208	0.88530	0.84164
DSFA	0.74321	0.81158	0.89826	0.85273
ASFA_FS	0.76943	0.79189	0.96446	0.86969

TABLE V
RESULT OF CHANGE DETECTION EXPERIMENT ON SZTAKI DATASET

Method	MIoU	Precision	Recall	F1
CVA	0.65848	0.77767	0.81126	0.79411
PCA	0.67157	0.79044	0.81711	0.80355
MAD	0.65883	0.78006	0.08092	0.79437
SFA	0.70031	0.80045	0.84850	0.82378
DSFA	0.71651	0.81811	0.85235	0.83488
ASFA_FS	0.77372	0.79616	0.96485	0.87242

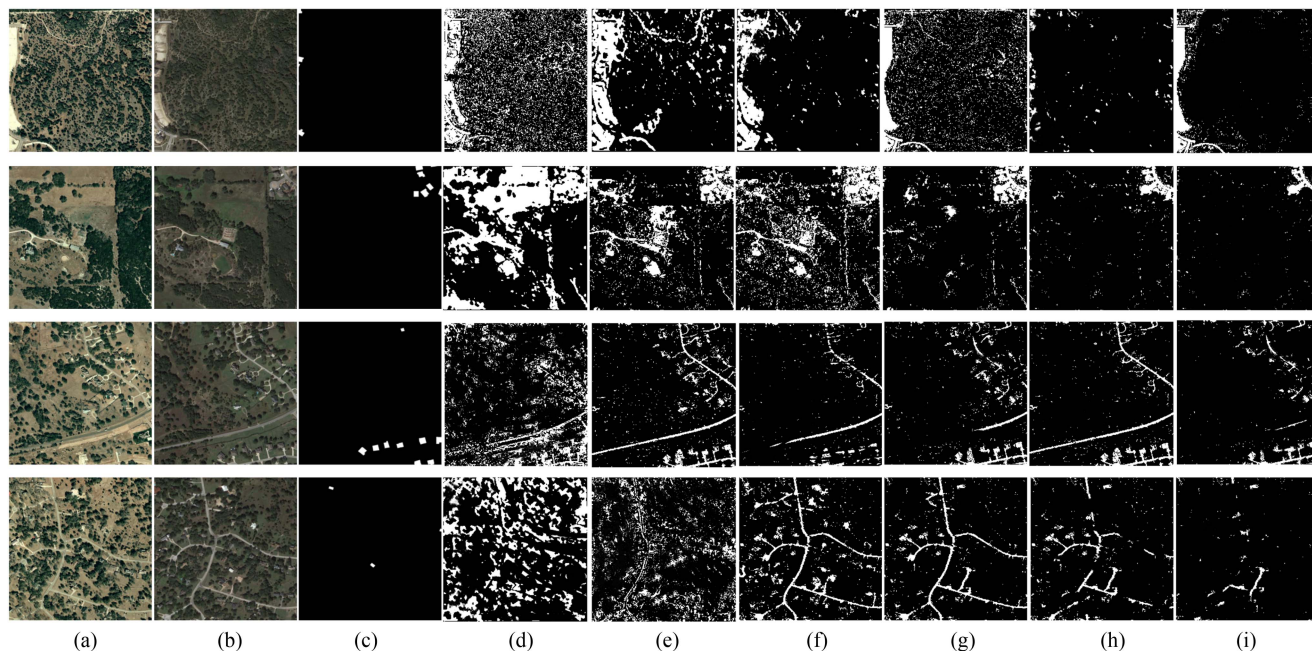


Fig. 7. Change map of change detection on LEVIR-CD dataset.

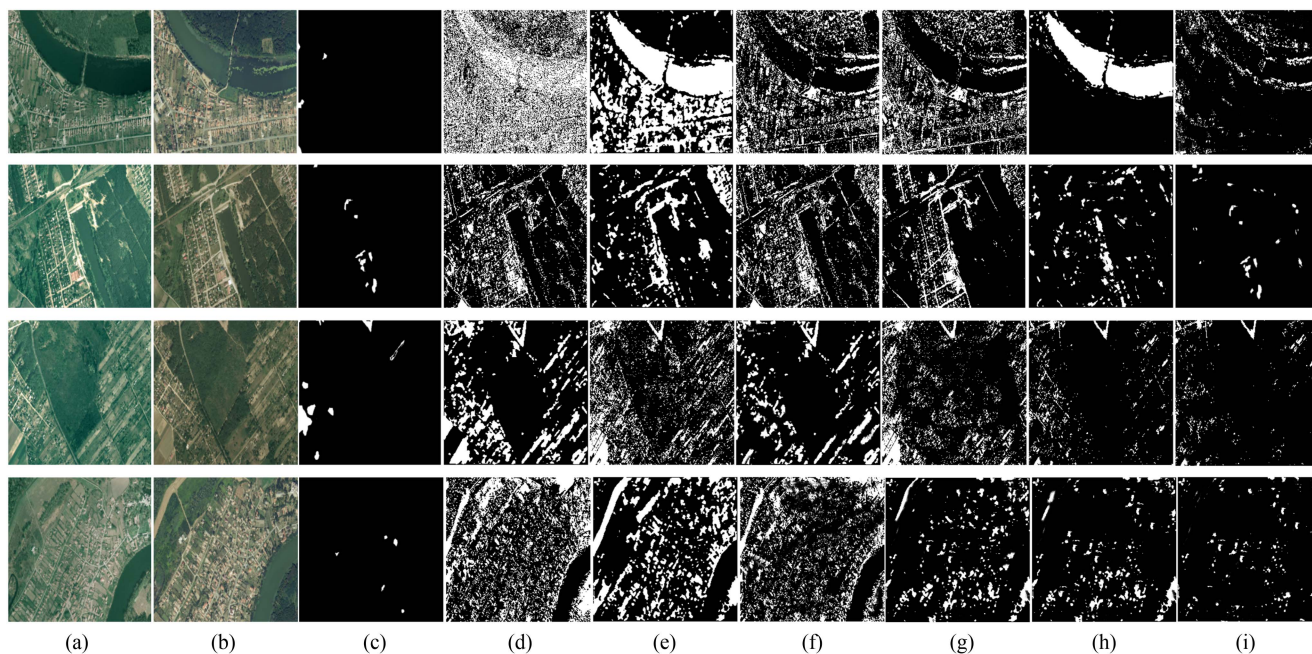


Fig. 8. Change map of change detection on SZTAKI dataset.

TABLE VI
RESULT OF EXPERIMENT WITH DIFFERENT AUGMENTED COVERAGE ON
LEVIR-CD DATASET

Method	MIoU	Precision	Recall	F1
FRSCD_DAFS_15	0.77669	0.86931	0.87937	0.87431
FRSCD_DAFS_35	0.81045	0.89150	0.89914	0.89531
FRSCD_DAFS_50	0.83552	0.87267	0.95152	0.91039

TABLE VII
RESULT OF EXPERIMENT WITH DIFFERENT AUGMENTED COVERAGE ON
SZTAKI DATASET

Method	MIoU	Precision	Recall	F1
FRSCD_DAFS_15	0.78162	0.89564	0.85994	0.87743
FRSCD_DAFS_35	0.83406	0.89203	0.92771	0.90952
FRSCD_DAFS_50	0.85027	0.87576	0.96690	0.91908

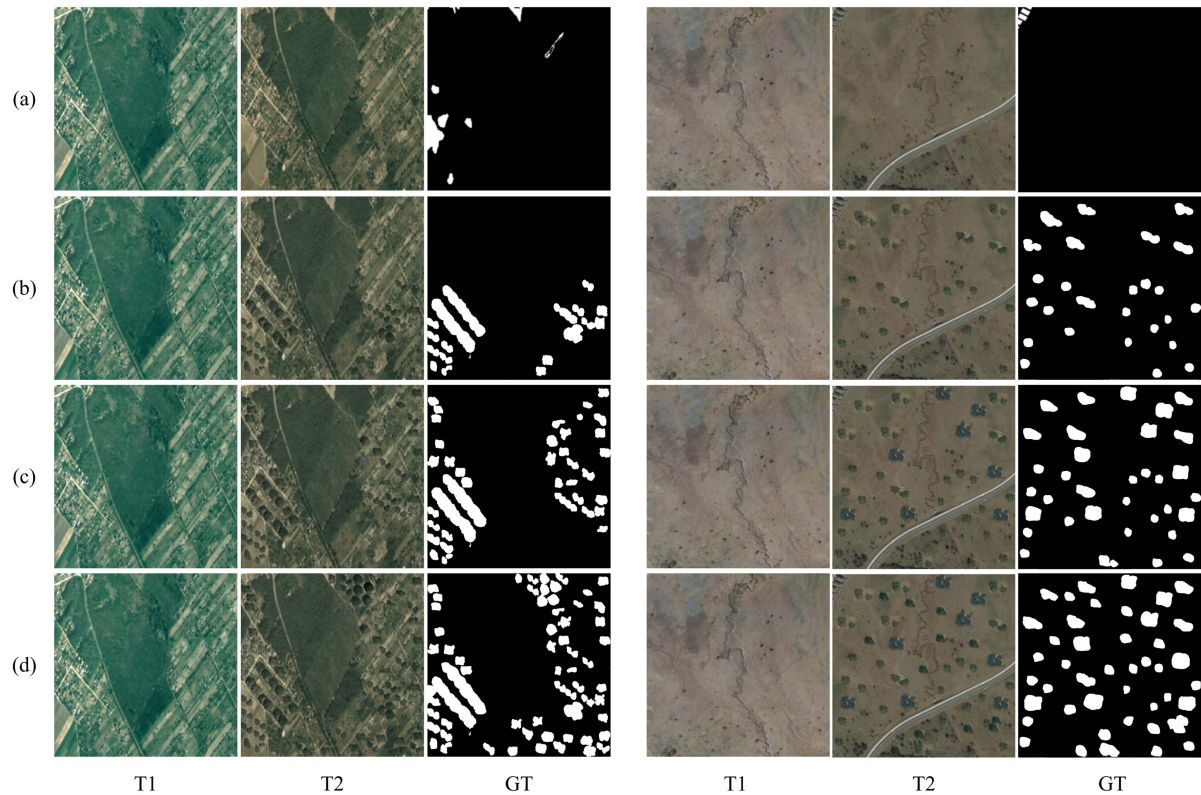


Fig. 9. Dataset after augmentation. (a) Origin data. (b) 15% forest fragment coverage. (c) 35% forest fragment coverage. (d) 50% forest fragment coverage. The left one displays the dataset of SZTAKI, and the right displays the dataset of LEVIR-CD.

TABLE VIII
RESULT OF CHANGE DETECTION WITH AUGMENTED DATASET ON LEVIR-CD

Method	MIoU	Precision	Recall	F1
CVA	0.67347	0.76708	0.84666	0.80491
CVA_DA	0.68433	0.74049	0.90023	0.81258
PCA	0.70283	0.78858	0.86610	0.82552
PCA_DA	0.70510	0.76130	0.90523	0.82705
MAD	0.71513	0.79308	0.87925	0.83394
MAD_DA	0.72423	0.77510	0.91691	0.84164
SFA	0.72652	0.80208	0.88530	0.84164
SFA_DA	0.76036	0.80534	0.93157	0.86387
DSFA	0.74321	0.81158	0.89826	0.85273
DSFA_DA	0.79940	0.84763	0.93355	0.88852

TABLE IX
RESULT OF CHANGE DETECTION WITH AUGMENTED DATASET ON SZTAKI

Method	MIoU	Precision	Recall	F1
CVA	0.65848	0.77767	0.81126	0.79411
CVA_DA	0.68436	0.74002	0.90097	0.81260
PCA	0.67157	0.79044	0.81711	0.80355
PCA_DA	0.69265	0.74590	0.90657	0.81842
MAD	0.65883	0.78006	0.08092	0.79437
MAD_DA	0.71481	0.76433	0.91691	0.83369
SFA	0.70031	0.80045	0.84850	0.82378
SFA_DA	0.75330	0.80001	0.92807	0.85929
DSFA	0.71651	0.81811	0.85235	0.83488
DSFA_DA	0.78956	0.83524	0.93522	0.88240

algorithms with a forest fragment coverage parameter of 50%. The results demonstrate that our proposed image augmentation method can be transferred to other change detection models and improve the detection results, as given in Tables VIII and IX. Table VIII gives the result of change detection with augmented dataset on LEVIR-CD dataset. Table IX gives the result of change detection with augmented dataset on SZTAKI dataset.

V. DISCUSSION

We proposed a forest remote sensing change detection model based on data augmentation and few-shot learning. In this section, we will illustrate the uniqueness of our work and the select basis of the method and data; moreover, the limitations and future work are discussed in detail.

Compared with other end-to-end change detection methods, we introduce the idea of few-shot learning and use metric learning to calculate the distance. The detection model is improved at

the level of algorithm and data, respectively. The uniqueness of this data augmentation is mainly reflected in the direct increase of the changed samples without enhancing the number of images. Different from data transformation enhancement methods, remote sensing images still keep real in authenticity and integrity in this way. It is somewhat surprising that the model without GP component runs a poor results. The experiment provides a new insight into the relationship between the data augmentation and the feature extraction.

Data selection provides important support for the training and performance of deep learning algorithms. Appropriate datasets can improve the accuracy and the ability of generalization, repeatability of the algorithm. In this model of forest change detection, we chose the appropriate data with a simple background and the main object of detection. The four datasets we selected are diverse, including different categories, angles, and lighting conditions. This can help the algorithm to better generalize to new scenes. In optical images, RGB data are a standard format for color images that have a relatively higher spatial resolution. We could accurately analyze image elements due to the comprehensive representation of the colors. Combined with the four standard datasets we selected, RGB data are considered to be very suitable for change detection in this scene.

However, our proposed model still has some limitations. As we all know, hyperspectral images contain more spectral and channel information, which is more advantageous in discriminating the seasonal changes of the forest. A feature extraction network is added in the process of generating forest fragments, which improves the detection effect but slightly increases the complexity of the model. Depending on the previously mentioned limitations, our future research can be divided into two directions. First, we may further optimize the forest change detection model, study a lighter model, and generalize it to other detection field scenes. Second, we will study the effects of illumination and shadows on images, extend the model's time-series detection capability, and study the long time-series forest change detection algorithm based on Landsat images, which is more instructive in practical applications.

VI. CONCLUSION

This study provides a complete change detection model for forest remote sensing under few-shot conditions and achieves the improvement of change detection accuracy from both data augmentation and few-shot learning algorithms together. A forest feature extraction network, a forest fragment generation network, and a seamless image synthesis method are proposed in this article. The effective change information of real forest scenes is increased without increasing the number of images, and the data expansion at the positive sample level is achieved. At the same time, a dual-network change detection algorithm for few-shot learning is proposed, and the slow feature analysis network is further improved to achieve end-to-end forest change detection. More importantly, the model we proposed can migrate to other forest and grass scenarios due to the rich remote sensing elements contained in training datasets. It is experimentally verified that this model increases the F1 score by 6.52% on

average and has an excellent performance in a two-pronged approach.

ACKNOWLEDGMENT

This article was produced by the IEEE Publication Technology Group, Piscataway, NJ, USA.

REFERENCES

- [1] C. Persello et al., "Deep learning and Earth observation to support the sustainable development goals: Current approaches, open challenges, and future opportunities," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 172–200, Jun. 2022, doi: [10.1109/MGRS.2021.3136100](https://doi.org/10.1109/MGRS.2021.3136100).
- [2] M. Liu, Z. Chai, H. Deng, and R. Liu, "A CNN-Transformer network with multiscale context aggregation for fine-grained cropland change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 4297–4306, May 2022, doi: [10.1109/JSTARS.2022.3177235](https://doi.org/10.1109/JSTARS.2022.3177235).
- [3] H. Zheng et al., "HFA-Net: High frequency attention siamese network for building change detection in VHR remote sensing images," *Pattern Recognit.*, vol. 129, 2022, Art. no. 108717, doi: [10.1016/j.patcog.2022.108717](https://doi.org/10.1016/j.patcog.2022.108717).
- [4] S. Kumar, S. Arya, and K. Jain, "A SWIR-based vegetation index for change detection in land cover using multi-temporal landsat satellite dataset," *Int. J. Inf. Technol.*, pp. 2035–2048, 2022, doi: [10.1007/s41870-021-00797-6](https://doi.org/10.1007/s41870-021-00797-6).
- [5] Q. Li, R. Zhong, X. Du, and Y. Du, "TransUNetCD: A hybrid transformer network for change detection in optical remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Apr. 2022, Art. no. 5622519, doi: [10.1109/TGRS.2022.3169479](https://doi.org/10.1109/TGRS.2022.3169479).
- [6] L. Bullock et al., "Timeliness in forest change monitoring: A new assessment framework demonstrated using Sentinel-1 and a continuous change detection algorithm," *Remote Sens. Environ.*, vol. 276, 2022, Art. no. 113043, doi: [10.1016/j.rse.2022.113043](https://doi.org/10.1016/j.rse.2022.113043).
- [7] 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, doi: [10.3390/rs14040871](https://doi.org/10.3390/rs14040871).
- [8] W. Chen et al., "GCSANet: A global context spatial attention deep learning network for remote sensing scene classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 1150–1162, Jan. 2022, doi: [10.1109/JSTARS.2022.3141826](https://doi.org/10.1109/JSTARS.2022.3141826).
- [9] H. Jiang et al., "A survey on deep learning-based change detection from high-resolution remote sensing images," *Remote Sens.*, vol. 14, no. 7, 2022, Art. no. 1552, doi: [10.3390/rs14071552](https://doi.org/10.3390/rs14071552).
- [10] Y. Li, O. Song, and Y. Zhang, "Combining deep learning and ontology reasoning for remote sensing image semantic segmentation," *Knowl.-Based Syst.*, vol. 243, 2022, Art. no. 108469, doi: [10.1016/j.knsys.2022.108469](https://doi.org/10.1016/j.knsys.2022.108469).
- [11] M. Xu et al., "A comprehensive survey of image augmentation techniques for deep learning," *Pattern Recognit.*, vol. 137, 2023, Art. no. 109347, doi: [10.1016/j.patcog.2023.109347](https://doi.org/10.1016/j.patcog.2023.109347).
- [12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, doi: [10.1007/s10710-017-9314-z](https://doi.org/10.1007/s10710-017-9314-z).
- [13] M. Köhler, M. Eisenbach, and H. M. Gross, "Few-shot object detection: A survey," *Assoc. Comput. Machinery Comput. Surv.*, vol. 54 no. 11, Art. no. 242, pp 1–37, doi: [10.1145/3519022](https://doi.org/10.1145/3519022).
- [14] M. Hermann, S. Saha, and X. X. Zhu, "Filtering specialized change in a few-shot setting," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 1185–1196, Jan. 2023, doi: [10.1109/JSTARS.2022.3231915](https://doi.org/10.1109/JSTARS.2022.3231915).
- [15] H. Xu et al., "An enhanced framework of generative adversarial networks (EF-GANs) for environmental microorganism image augmentation with limited rotation-invariant training data," *IEEE Access*, vol. 8, pp. 187455–187469, 2020, doi: [10.1109/ACCESS.2020.3031059](https://doi.org/10.1109/ACCESS.2020.3031059).
- [16] Y. Cui et al., "Coarse-to-fine pseudo supervision guided meta-task optimization for few-shot object classification," *Pattern Recognit.*, vol. 122, 2022, Art. no. 108296, doi: [10.1016/j.patcog.2021.108296](https://doi.org/10.1016/j.patcog.2021.108296).
- [17] C. Wang et al., "Comparison of vegetation phenology derived from solar-induced chlorophyll fluorescence and enhanced vegetation index, and their relationship with climatic limitations," *Remote Sens.*, vol. 14, no. 13, 2022, Art. no. 3018, doi: [10.3390/rs14133018](https://doi.org/10.3390/rs14133018).
- [18] A. Goswami et al., "Change detection in remote sensing image data comparing algebraic and machine learning methods," *Electronics*, vol. 11, no. 3, 2022, Art. no. 431, doi: [10.3390/electronics11030431](https://doi.org/10.3390/electronics11030431).
- [19] 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, doi: [10.1016/j.isprsjrs.2021.12.005](https://doi.org/10.1016/j.isprsjrs.2021.12.005).

- [20] Q. Wu, Y. Chen, and J. Meng, "DCGAN-Based data augmentation for tomato leaf disease identification," *IEEE Access*, vol. 8, pp. 98716–98728, 2020, doi: [10.1109/ACCESS.2020.2997001](https://doi.org/10.1109/ACCESS.2020.2997001).
- [21] H. Duan, X. Dong, S. You, and S. Han, "A deep learning denoising framework based on FFDNet for SAR image change detection," in *Proc. IEEE 11th Int. Conf. Electron. Inf. Emerg. Commun.*, 2021, pp. 1–4, doi: [10.1109/ICEIECS1955.2021.9463810](https://doi.org/10.1109/ICEIECS1955.2021.9463810).
- [22] W. Zhang and H. Fan, "Application of isolated forest algorithm in deep learning change detection of high resolution remote sensing image," in *Proc. IEEE Int. Conf. Artif. Intell. Comput. Appl. (ICAICA)*, 2020, pp. 753–756, doi: [10.1109/ICAICA50127.2020.9181873](https://doi.org/10.1109/ICAICA50127.2020.9181873).
- [23] M. Esfandiari, G. Abdi, S. Jabari, and V. S. P. Lolla, "Building change detection in off-nadir images using deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 1347–1350, doi: [10.1109/IGARSS47720.2021.9553172](https://doi.org/10.1109/IGARSS47720.2021.9553172).
- [24] S. Li and L. Huo, "Remote sensing image change detection based on fully convolutional network with pyramid attention," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 4352–4355, doi: [10.1109/IGARSS47720.2021.9554522](https://doi.org/10.1109/IGARSS47720.2021.9554522).
- [25] Y. Yang, H. Gu, Y. Han, and H. Li, "An end-to-end deep learning change detection framework for remote sensing images," in *Proc. IGARSS-IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 652–655, doi: [10.1109/IGARSS39084.2020.9324076](https://doi.org/10.1109/IGARSS39084.2020.9324076).
- [26] C. Wu, B. Du, and L. Zhang, "Fully convolutional change detection framework with generative adversarial network for unsupervised, weakly supervised and regional supervised change detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, Jan. 18, 2023 doi: [10.1109/TPAMI.2023.2327896](https://doi.org/10.1109/TPAMI.2023.2327896).
- [27] Y. Peng, B. Cui, H. Yin, Y. Zhang, and P. Du, "Automatic SAR change detection based on visual saliency and multi-hierarchical fuzzy clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 7755–7769, Aug. 2022, doi: [10.1109/JSTARS.2022.3199017](https://doi.org/10.1109/JSTARS.2022.3199017).
- [28] J. Huang, Q. Fu, X. Wang, and Y. Ji, "Remote sensing building change detection based on improved U-Net," in *Proc. 3rd Int. Conf. Big Data, Artif. Intell. Internet Things Eng.*, 2022, pp. 772–775, doi: [10.1109/ICBAIE56435.2022.9985853](https://doi.org/10.1109/ICBAIE56435.2022.9985853).
- [29] C. Zhang et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 166, pp. 183–200, 2020, doi: [10.1016/j.isprsjprs.2020.06.003](https://doi.org/10.1016/j.isprsjprs.2020.06.003).
- [30] B. Du, L. Ru, C. Wu, and L. Zhang, "Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9976–9992, Dec. 2019, doi: [10.1109/TGRS.2019.2930682](https://doi.org/10.1109/TGRS.2019.2930682).
- [31] P. Du et al., "An improved change detection approach using tri-temporal logic-verified change vector analysis," *ISPRS J. Photogrammetry Remote Sens.*, vol. 161, pp. 278–293, 2020, doi: [10.1016/j.isprsjprs.2020.01.026](https://doi.org/10.1016/j.isprsjprs.2020.01.026).
- [32] C. Wu, H. Chen, B. Du, and L. Zhang, "Unsupervised change detection in multitemporal VHR images based on deep kernel PCA convolutional mapping network," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 12084–12098, Nov. 2022, doi: [10.1109/TCYB.2021.3086884](https://doi.org/10.1109/TCYB.2021.3086884).
- [33] S. Liu, Q. Du, X. Tong, A. Samat, and L. Bruzzone, "Unsupervised change detection in multispectral remote sensing images via spectral-spatial band expansion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3578–3587, Sep. 2019, doi: [10.1109/JSTARS.2019.2929514](https://doi.org/10.1109/JSTARS.2019.2929514).
- [34] L. Wanab, L. Maab, J. Guoad, M. Liuac, and D. Y. Ab, "Slow feature analysis based on convolutional neural network for SAR image change detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 3805–3808, doi: [10.1109/IGARSS47720.2021.9553912](https://doi.org/10.1109/IGARSS47720.2021.9553912).
- [35] W. Jing et al., "Remote sensing change detection based on unsupervised multi-attention slow feature analysis," *Remote Sens.*, vol. 14, no. 12, 2022, Art. no. 2834, doi: [10.3390/rs14122834](https://doi.org/10.3390/rs14122834).
- [36] T. Kattenborn et al., "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogrammetry Remote Sens.*, vol. 173, pp. 24–49, 2021, doi: [10.1016/j.isprsjprs.2020.12.010](https://doi.org/10.1016/j.isprsjprs.2020.12.010).
- [37] A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 1, 2021, Art. no. 100004, doi: [10.1016/j.jjime.2020.100004](https://doi.org/10.1016/j.jjime.2020.100004).
- [38] Z. Cai et al., "Generative adversarial networks: A survey toward private and secure applications," *ACM Comput. Surv.*, vol. 54, no. 6, pp. 1–38, 2021, doi: [10.1145/3459992](https://doi.org/10.1145/3459992).
- [39] A. Ajit, K. Acharya, and A. Samanta, "A review of convolutional neural networks," in *Proc. Int. Conf. Emerg. Trends Inf. Technol. Eng.*, 2020, pp. 1–5, doi: [10.1109/ic-ETITE47903.2020.049](https://doi.org/10.1109/ic-ETITE47903.2020.049).
- [40] Y. Wang et al., "An intelligent method for predicting the pressure coefficient curve of airfoil-based conditional generative adversarial networks," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 23, 2021, doi: [10.1109/TNNLS.2021.3111911](https://doi.org/10.1109/TNNLS.2021.3111911).
- [41] Q. Zhang, J. Yang, X. Zhang, and T. Cao, "Generating adversarial examples in audio classification with generative adversarial network," in *Proc. 7th Int. Conf. Image. Vis. Comput.*, 2022, pp. 848–853, doi: [10.1109/ICIVC55077.2022.9886154](https://doi.org/10.1109/ICIVC55077.2022.9886154).
- [42] B. Hariharan et al., "Hybrid deep convolutional generative adversarial networks (DCGANs) and style generative adversarial network (STYLEGANS) algorithms to improve image quality," in *Proc. 3rd Int. Conf. Electron. Sustain. Commun. Syst.*, 2022, pp. 1182–1186, doi: [10.1109/ICESCS4411.2022.9885611](https://doi.org/10.1109/ICESCS4411.2022.9885611).
- [43] A. Dureja and P. Pahwa, "Analysis of non-linear activation functions for classification tasks using convolutional neural networks," *Recent Patents Comput. Sci.*, vol. 12, no. 3, pp. 156–161, 2019, doi: [10.2174/2213275911666181025143029](https://doi.org/10.2174/2213275911666181025143029).
- [44] K. Pratama and D. K. Kang, "Trainable activation function with differentiable negative side and adaptable rectified point," *Appl. Intell.*, vol. 51, no. 3, pp. 1784–1801, 2021, doi: [10.1007/s10489-020-01885-z](https://doi.org/10.1007/s10489-020-01885-z).
- [45] X. Xiao and L. Ma, "Color transfer in correlated color space," in *Proc. ACM Int. Conf. Virtual Reality Continuum Its Appl.*, 2006, pp. 305–309, doi: [10.1145/1128923.1128974](https://doi.org/10.1145/1128923.1128974).
- [46] C. Gu, X. Lu, and C. Zhang, "Example-based color transfer with Gaussian mixture modeling," *Pattern Recognit.*, vol. 129, 2022, Art. no. 108716, doi: [10.1016/j.patcog.2022.108716](https://doi.org/10.1016/j.patcog.2022.108716).
- [47] N. Ibtihaz and M. S. Rahman, "MultiResUNet: Rethinking the U-net architecture for multimodal biomedical image segmentation," *Neural Netw.*, vol. 121, pp. 74–87, 2020, doi: [10.1016/j.neunet.2019.08.025](https://doi.org/10.1016/j.neunet.2019.08.025).
- [48] H. Wang et al., "Anomaly segmentation in retinal images with poisson-blending data augmentation," *Med. Image Anal.*, vol. 81, 2022, Art. no. 102534, doi: [10.1016/j.media.2022.102534](https://doi.org/10.1016/j.media.2022.102534).
- [49] R. R. Hussein, Y. I. Hamodi, and A. S. Roaa, "Retinex theory for color image enhancement: A systematic review," *Int. J. Elect. Comput. Eng.*, vol. 9, no. 6, 2019, Art. no. 5560, doi: [10.11591/ijece.v9i6.pp5560-5569](https://doi.org/10.11591/ijece.v9i6.pp5560-5569).
- [50] Y. Li et al., "Reilly-type inequality for the ϕ -Laplace operator on semislant submanifolds of Sasakian space forms," *J. Inequalities Appl.*, vol. 2022, no. 1, 2022, Art. no. 102, doi: [10.1186/s13660-022-02838-5](https://doi.org/10.1186/s13660-022-02838-5).
- [51] X. Sun, B. Wang, Z. Wang, H. Li, H. Li, and K. Fu, "Research progress on few-shot learning for remote sensing image interpretation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2387–2402, Jan. 2021, doi: [10.1109/JSTARS.2021.3052869](https://doi.org/10.1109/JSTARS.2021.3052869).
- [52] Q. Yang, X. Yang, and X. Ji, "NAM Net: Meta-network with normalization-based attention for few-shot learning," in *Proc. 2nd Int. Conf. Computer Sci., Electron. Inf. Eng. Intell. Control Technol.*, 2022, pp. 473–476, doi: [10.1109/CEIS7409.2022.9950152](https://doi.org/10.1109/CEIS7409.2022.9950152).
- [53] F. Zhou, L. Zhang, W. Wei, Z. Bai, and Y. Zhang, "Meta transfer learning for few-shot hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 3681–3684, doi: [10.1109/IGARSS47720.2021.9553981](https://doi.org/10.1109/IGARSS47720.2021.9553981).
- [54] X. Chao and Y. Li, "Semisupervised few-shot remote sensing image classification based on KNN distance entropy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 8798–8805, Oct. 2022, doi: [10.1109/JSTARS.2022.3213749](https://doi.org/10.1109/JSTARS.2022.3213749).
- [55] P. Zhu, M. Gu, W. Li, C. Zhang, and Q. Hu, "Progressive point to set metric learning for semi-supervised few-shot classification," in *Proc. IEEE Int. Conf. Image Process.*, 2020, pp. 196–200, doi: [10.1109/ICIP40778.2020.9191261](https://doi.org/10.1109/ICIP40778.2020.9191261).
- [56] S. Deng et al., "Emotion class-wise aware loss for image emotion classification," in *Proc. Artif. Intell., 1st CAAI Int. Conf.*, 2021, pp. 553–564, doi: [10.1007/978-3-030-93046-2_47](https://doi.org/10.1007/978-3-030-93046-2_47).
- [57] D. Kim, K. Ko, D. K. Han, and H. Ko, "Discriminatory and orthogonal feature learning for noise robust keyword spotting," *IEEE Signal Process. Lett.*, vol. 29, pp. 1913–1917, Sep. 2022, doi: [10.1109/LSP.2022.3203911](https://doi.org/10.1109/LSP.2022.3203911).
- [58] Z. Zhang, J. Hao, C. Pan, and G. Ji, "Oriented feature augmentation for few-shot object detection in remote sensing images," in *Proc. IEEE Int. Conf. Comput. Sci.*, 2021, pp. 359–366, doi: [10.1109/CEIS2496.2021.9574548](https://doi.org/10.1109/CEIS2496.2021.9574548).
- [59] C. Liu et al., "Two-channel feature extraction convolutional neural network for facial expression recognition," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 24, no. 6, pp. 792–801, 2020, doi: [10.20965/jaciii.2020.p0792](https://doi.org/10.20965/jaciii.2020.p0792).
- [60] K. Adem, "Impact of activation functions and number of layers on detection of exudates using circular hough transform and convolutional neural networks," *Expert Syst. Appl.*, vol. 203, 2022, Art. no. 117583, doi: [10.1016/j.eswa.2022.117583](https://doi.org/10.1016/j.eswa.2022.117583).

- [61] M. Guo et al., "Attention mechanisms in computer vision: A survey," *Comput. Vis. Media*, vol. 8, no. 3, pp. 331–368, 2022, doi: [10.1007/s41095-022-0271-y](https://doi.org/10.1007/s41095-022-0271-y).
- [62] H. Chen and Z. Shi, "A spatial-temporal attention-based method and a new dataset for remote sensing image change detection," *Remote Sens.*, vol. 12, no. 10, 2020, Art. no. 1662, doi: [10.3390/rs12101662](https://doi.org/10.3390/rs12101662).
- [63] C. Benedek and T. Sziranyi, "Change detection in optical aerial images by a multilayer conditional mixed Markov model," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 10, pp. 3416–3430, Oct. 2009, doi: [10.1109/TGRS.2009.2022633](https://doi.org/10.1109/TGRS.2009.2022633).
- [64] D. Peng, L. Bruzzone, Y. Zhang, H. Guan, H. Ding, and X. Huang, "SemiCDNet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5891–5906, Jul. 2021, doi: [10.1109/TGRS.2020.3011913](https://doi.org/10.1109/TGRS.2020.3011913).
- [65] 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, Aug. 2022, Art. no. 5604816, doi: [10.1109/TGRS.2021.3085870](https://doi.org/10.1109/TGRS.2021.3085870).
- [66] C. Song, A. Pons, and K. Yen, "Optimizing stochastic gradient descent using the angle between gradients," in *Proc. IEEE Int. Conf. Big Data*, 2020, pp. 5269–5275, doi: [10.1109/BigData50022.2020.9378007](https://doi.org/10.1109/BigData50022.2020.9378007).
- [67] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: [10.1109/TPAMI.2017.2699184](https://doi.org/10.1109/TPAMI.2017.2699184).
- [68] 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, Mar. 2022, Art. no. 5603216, doi: [10.1109/TGRS.2021.3066802](https://doi.org/10.1109/TGRS.2021.3066802).



Songyu Zhu is currently working toward the Ph.D. degree in forestry information engineering with Northeast Forestry University, Harbin, China, since 2020.

She is also with Harbin Vocational and Technical College, Harbin, as a Lecturer. Her research interests include remote sensing change detection, deep learning, and smart forestry.



Weipeng Jing (Member, IEEE) received the Ph.D. degree in computer architecture from the Harbin Institute of Technology, Harbin, China, in 2016.

He is currently a Professor with Northeast Forestry University, Harbin. He has authored or coauthored more than 100 research articles in refereed journals and conference proceedings, such as *IEEE TRANSACTIONS AND GEOSCIENCE AND REMOTE SENSING*, *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, and *IEEE SENSOR JOURNAL*.

His research interests include modeling and scheduling for distributed computing systems, artificial intelligence, remote sensing image intelligent mining, and spatial data mining.

Dr. Jing is currently a Member of the ACM and the China Computer Federation (CCF). He was the Publication Chair for the International Conference of Pioneering Computer Scientists, Engineers and Educators (ICPCSEE) (2016–2023) data Technologies and Applications (BDTA) (2015 and 2017), Collaborate 2017, 9th EAI International Wireless Internet Conference (WICON 2016), IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud 2016), etc.



Peilun Kang is currently working toward the Ph.D. degree in forestry engineering with Northeast Forestry University, Harbin, China.

His research focuses on forestry remote sensing and deep learning with a particular expertise in utilizing deep learning technologies to solve problems in the field of forestry.



Mahmoud Emam was born in Egypt in 1985. He received B.S. and M.S. degrees in mathematics and computer science from Menoufia University, Shibin El Kom, Egypt, in 2006 and 2012, respectively, and the Ph.D. degree in computer science and technology from the Harbin Institute of Technology, Harbin, China, in 2017.

From 2013 to 2017, he was a Research Associate with the Research Institute of Information Countermeasure Techniques (ICT), School of Computer Science and Technology, Harbin Institute of Technology.

He is currently an Assistant Professor with the Faculty of Artificial Intelligence, Menoufia University. He is an author and coauthor of many publications and a referee of many referred international journals and conferences. His research interests include artificial intelligence, digital forensics, computer vision, machine learning, and pattern recognition.

Dr. Emam was the recipient of two best paper awards from the IEEE and Springer conferences, Best Foreigner Ph.D. Student Award, and the 2016 Outstanding International Ph.D. Student Excellence Award by the Chinese Scholarship Council, China, in 2016.



Chao Li received the Ph.D. degree in information and communication systems from the Harbin Institute of Technology, Harbin, China, in 2012.

She is currently the Deputy Director of artificial intelligence and master's tutor with Harbin Northeast Forestry University, Harbin. She has chaired and completed projects, such as the National Natural Science Foundation of China Youth Project. She has long conducted research on signal processing, feature extraction, and multimodal fusion. She was a Visiting Scholar with the Center for Forestry and Natural Resources, University of West Virginia, Morgantown, WV, USA.

She has authored or coauthored more than 20 SCI/EI retrieval papers with the first author and corresponding author.

Dr. Li is currently a Member of the China Computer Federation (CCF) and the Chinese Society of Forestry (CSF). She was the recipient of the Second Prize in Heilongjiang Provincial Science and Technology Progress Award.