IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 18, 2019, accepted June 10, 2019, date of publication June 13, 2019, date of current version June 28, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2922839

High-Resolution Remote Sensing Image Change Detection Combined With Pixel-Level and Object-Level

LU XU[1], WEIPENG JING[1], (Member, IEEE), HOUBING SONG[2], (Senior Member, IEEE), AND GUANGSHENG CHEN[1]

[1]College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China
[2]Department of Electrical, Computer, Software, and Systems Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA

Corresponding author: Houbing Song (songh4@erau.edu)

This work was supported in part by the National Natural Science Foundation of China under Grant 31770768, in part by the Natural Science Foundation of Heilongjiang Province of China under Grant F2017001, in part by the Heilongjiang Province Applied Technology Research and Development Program Major Project under Grant GA18B301, and in part by the China State Forestry Administration Forestry Industry Public Welfare Project under Grant 201504307.

ABSTRACT High-resolution remote sensing images are abundant in texture information, and the detection method of the change of pixel-level mainly analyzes the spectral information of the image, which has certain limitations. In this paper, a high-resolution remote sensing image change detection method combining pixel and object levels is proposed to solve the problem that many pepper and salt phenomenon and false detection in the change detection of pixel-level and object-level change detection method are cumbersome for image segmentation process. We integrate the multi-dimensional features of high-resolution remote sensing images and use random forest classifiers to classify to obtain the pixel-level change detection results. Then, we use the improved U-net network to semantically segment the post-phase remote sensing image to obtain the image object segmentation result. Finally, the consequences of pixel-level change detection and image object segmentation result are fused to obtain the image changing area and the unchanging area. The experimental results demonstrate that the algorithm has a higher accuracy rate and detection precision.

INDEX TERMS Change detection, random forest, remote sensing, semantic segmentation, U-net.

I. INTRODUCTION

In a rapidly changing modern society, urban expansion, land degradation, forest reduction, illegal land use and other phenomena occur all the time. In order to realize a modern smart city, it is necessary to grasp the current situation and change of land use [1], [2]. Remote sensing image change detection refers to the use of remote sensing images and related data in the same region at different periods, and compares them with image processing and mathematical models to analyze and judge the changes between the images to extract significant changes in the contrast images and generate the changing images, so as to obtain accurate ground object change information [3]. The change detection of remote sensing image can monitor the surface change, which has important and far-reaching significance for agricultural investigation, urban development research, forest resources monitoring and disaster relief work. However, if the method of manually plotting the changing areas in remote sensing images is adopted, it will consume a lot of manpower, material resources, and financial resources, and it will become less efficient.

In recent years, with the development of sensors and technology, the resolution of satellite remote sensing images has been constantly improved [4], [5]. Compared with remote sensing images with medium and low resolution, the spectral difference of similar ground objects in the remote sensing image with high resolution increased, which will show more details [6]. The traditional method of change detection based on spectral information reduces the separability between the changing and unchanging regions, and the difficulty of change feature extraction increases. Therefore, a more accurate change detection method is required.

According to the granularity unit of change detection analysis, the existing high-resolution remote sensing image change detection methods can be divided into pixel-based change detection and object-oriented change

The associate editor coordinating the review of this manuscript and approving it for publication was Zhen Ling.

detection [7], [8]. The pixel-based change detection takes pixels as processing unit, and generally adopts direct comparison or classification first and then comparison methods, including difference method [9], ratio method [10], classification detection method [11], principal component analysis [12], [13] and so on. Such methods are simple and easy to operate, and are widely used. But the processing only considers the characteristics of a single pixel and lacks the spectral features of the surrounding pixels and spatial feature information of adjacent pixels, and the detection accuracy is low. However, the details of ground objects in current high-resolution remote sensing image are clear and the target details are richer, Moreover, background features sometimes exhibit color and texture features similar to the target. It is difficult to obtain ideal change detection results only by using low-level features such as color, texture and edge in the image [14]. In order to improve the accuracy of change detection and better to avoid ''pepper and salt phenomenon'', the object-oriented change detection method was proposed and gradually improved. The object-based method processing unit as an object. Based on the image segmentation, the image is divided into several regions with homogeneity of shape and spectral properties, and then the change detection is carried out [15].

Traditional object-oriented change detection algorithms rarely consider the correlation between the neighborhood of the object. In the process of object-oriented change detection, image segmentation is a very important step, but the existing image segmentation algorithm still has certain limitations, which will cause the ''over-segmented'' or ''under-segmented'' of the object.

With the continuous development of deep learning, neural networks are gradually applied to the field of remote sensing. Semantic segmentation combines both tasks of image segmentation and target recognition to identify the category of each region and obtain an image with semantic labeling. Therefore, semantic segmentation has become a hot issue in the field of computer vision and pattern recognition. Traditional deep convolutional classification networks such as AlexNet [16] and VGGNet [17] usually contain full connected layer, so the size of the input image is required to be fixed. These network models have the disadvantages of high storage cost, low computational efficiency and the limited size of the sensing region.

In 2014, Jonathan Long *et al.* [18] proposed a fully convolutional network. The FCN uses the framework of end-to-end segmentation method, which takes into account both local information and global information so that the segmentation can be performed on images of any size to achieve semantic segmentation results for pixel-level annotation. U-net is an extension of FCN and a fully convolutional network with good expansibility at present [19]. The shallow layer of the network is used to solve the problem of pixel positioning, while the deeper layer is used to solve the problem of pixel classification. By combining low-level feature mapping and constructing high-level complex features, the image segmentation problem is solved. The use of deep learning can greatly improve people's ability to process high-resolution remote sensing images, change the disadvantages of traditional remote sensing data processing, such as time-consuming and low efficiency, and realize an intelligent, large-scale and real-time remote sensing information production.

In this paper, we propose a method combining pixel-level change detection and object-level change detection based on deep learning semantic segmentation. The remainder of this paper is organized as follows. Section II introduces relevant research on recent change detection methods. Section III shows the change detection method of merging pixel-level and object-level. Section IV describes the experimental results and analysis of the change detection experiment. Finally, conclusions are presented in Section V.

II. RELATED WORK

In the early pixel-level change detection, spectral and pixel information of high-resolution remote sensing images are mostly used to detect changes by extracting image difference maps or features of image shape and texture [20]. Dian *et al.* [21] uses the mean-shift segmentation algorithm to obtain the gray feature information of geographic objects with different phases. Combined with the change vector analysis, the maximum mathematical expectation algorithm is used to automatically extract the change region. Gong *et al.* [22] proposes a neighborhood-based ratio method to generate difference maps. Although the algorithm is superior to the traditional method of generating the difference map, the change detection result is severely affected by noise and the denoising effect is not ideal.

Recently, object-oriented change detection methods have been constantly improved. Although they improve the accuracy and reduce the noise compared with the pixel-oriented methods, there are still shortcomings. A scale-driven object-oriented change detection model is proposed to analyze the scale uncertainty existing in the segmentation results, so as to weaken the impact of segmentation uncertainty on the change detection results [23]. This method has the uncertainty of segmentation scale, which is easy to introduce noise in the process of change detection and reduce the reliability of the change detection results.

The object-oriented change detection methods can also take advantage of various features of the images, such as texture features, shape features, and context features. An object-oriented multi-feature adaptive fusion change detection method is proposed, which combines texture features such as contrast, correlation and homogeneity, as well as shape features such as edge, compactness, density and shape coefficient [24]. The false detection rate and missed detection rate is reduced and the accuracy of detection is improved through constructing a neural network to merge the features. Wang *et al.* [25] uses the object-based variation vector analysis method, correlation coefficient method, etc., to comprehensively utilize the various features of the object to participate in the analysis, which can improve the detection accuracy of the change compared with using only one

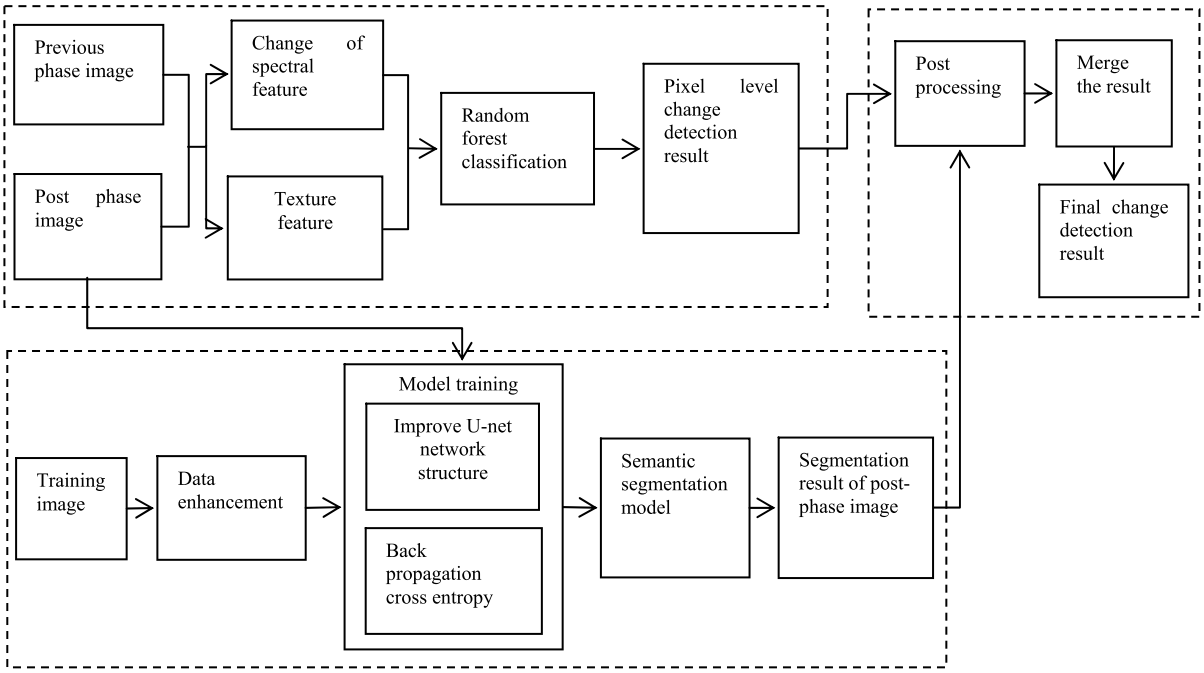

FIGURE 1. The schema of proposed approach.

single feature. However, the above methods have a significant dependence on several aspects such as the quality of feature selection and feature weight distribution.

Due to the shortcomings of both pixel-based change detection and object-oriented segmentation, the fusion of the two methods has become a research hotspot in recent years. For example, Cao *et al.* [26] extracts pixel-based variation features and object-based variation features, and carries out change detection in high-resolution remote sensing images by means of horizontal set evolution and support vector machine. Tang *et al.* [27] uses spectral and texture information to build a single Gaussian model, and then used pixel-level detection results as seed regions. Simultaneously growing on the image area before and after the change at the same time, and the growth region is taken as a union to segment the change object. Xiao *et al.* [28] takes the super pixel as the basic unit of change detection, uses the entropy rate-based method to segment the image. Extracts the change intensity image and slope difference image in the spectral space and slope space, and use the weighted fusion of these two kinds of difference information according to a certain weight to obtain the final change and unchanged area. From the perspective of entropy rate superpixel segmentation algorithm, it only considers the spectral information of remote sensing image, and has a certain disadvantage to the segmentation quality of the image.

Deep learning model can be used to improve the quality of high-resolution remote sensing image segmentation. Girshick *et al.* [29] proposes a segmentation method of region-based CNN (R-CNN), which selects multiple candidate image regions from the image as input samples to be learned in CNN. Compared with traditional image segmentation methods, it can implement semantic segmentation tasks better. However, this method only realizes the image-level classification and cannot directly recognize the semantics of each pixel in the image. Shu *et al.* [30] tries to find the possible object center point of the detection object first, then use the model combined with R-FCN and ResNet-101 to segment the image according to the center point. Kampffmeyer *et al.* [31] conducts semantic segmentation of remote sensing images in urban areas, modeling uncertain information, and focuses on solving the problem of inaccurate segmentation caused by low priority of small objects. Semantic segmentation methods based on deep learning often require large-scale or even super large-scale data sets, days of training time, and precise manual marking.

To solve the above problems, we propose a method combining pixel-level change detection and object-level change detection based on semantic segmentation. There are two innovations in this paper. In the pixel-level change detection part, we use the spectral features and texture features of the image to construct the feature vector. The random forest algorithm is used to improve the accuracy of pixel level change detection. In the object-oriented segmentation part, we propose an improved semantic segmentation network structure of high-resolution remote sensing images based on the u-net network for semantic segmentation of images.

III. METHOD OF THIS PAPER

The flow chart of the change detection algorithm proposed in this paper is shown in Figure 1, which mainly includes the following three steps: (1) Perform spectral feature differences and texture feature extraction on two-phase images, and merge the two features. Perform random forest classification

on the merged features to obtain initial pixel-level detection results. (2) Use the improved U-net network to carry out semantic segmentation on the post-phase image; (3) Merge the pixel-level detection result with the semantic segmentation result, and the detection result graph is obtained after removing isolated and noise points.

A. PIXEL-LEVEL CHANGE DETECTION

1) SPECTRAL FEATURES DIFFERENCE

The spectral feature is a common image feature in remote sensing image analysis. In high-resolution remote sensing image, besides the panchromatic band, color information of four bands including red, green, blue, and near-infrared is also included. The changing intensity of remote sensing images with multiple bands can be calculated according to the difference of each band. Assuming that x1 and x2 are two pixels of the position corresponding to the image S1 and the image S2, respectively, the spectral feature vectors of the two images are respectively composed of their pixel values p1 and p2 in the respective bands. By calculating the intensity change of each pixel, the spectral intensity change figure C of two remote sensing images can be obtained. The calculation formula is as follows:

$$C = \sqrt{\sum_{i=1}^{n} (p_{2i} - p_{1i})^2} \tag{1}$$

where n is the number of bands of two remote sensing images. If the changing intensity of a pixel in spectral difference figure C is greater than the change threshold value δ, it indicates that the positions of pixel x1 and x2 are changed, otherwise there is no change.

2) EXTRACTION OF TEXTURE FEATURE

The texture feature is a worldwide feature that depicts the repeated local patterns in the image and their arrangement rules. It describes the surface properties of the scene corresponding to the image or image region and is closely related to the dynamic detection of pixel gray values. As a kind of statistical feature, texture features are often invariant in rotation and have a strong resistance to noise. The integration of texture features into the change detection method can significantly improve the change detection accuracy of remote sensing images [32]. Texture features have been extensively applied in image retrieval and image classification. The Gray-level Co-occurrence Matrix (GLCM) method is a widely used extraction method with strong adaptability and robustness. Therefore, in this paper, the gray level co-occurrence matrix is utilized to extract texture features of remote sensing images for change detection.

The gray level co-occurrence matrix was originally proposed by Robert M., which is a matrix function of pixel distance and angle. It reflects the comprehensive information of the image in the direction, interval, change range and speed by calculating the correlation between the grayscale of two points with a certain distance and a certain direction in the image. The gray level co-occurrence matrix can derive 14 kinds of gray level co-occurrence matrix features by calculating the texture direction, texture scale, window moving distance and other parameters [33]. Due to a large amount of redundancy between these features, we select four less correlated features with small correlation to extract image texture features, namely Mean, Homogeneity, Contrast, Correlation. The calculation formulas are as follows:

$$Mean = \sum_{i,j=1}^{N} i \times (P_{i,j}) \tag{2}$$

$$Hom = \sum_{i,j=1}^{N} \frac{P_{i,j}}{1 + (i-j)^2} \tag{3}$$

$$Con = \sum_{i=1}^{N} \sum_{j=1}^{N} (i,j)^2 P_{i,j} \tag{4}$$

$$Correlation = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} (ij) P_{i,j} - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{5}$$

where N is the number of gray levels of the image, i and j are the row and column elements of the matrix elements respectively. $P_{i,j}$is the element of GLCM, among which $\mu_x$and$\mu_y$are the mean value, $\sigma_x$and $\sigma_y$are the standard deviations.

Texture scale is one of the crucial parameters in extracting texture features. Appropriate texture scale can more accurately display the texture features of the target and improve the accuracy of remote sensing image change detection. If the window is too small, the calculation of texture feature extraction is too large, while if the window is too large, the image texture feature extraction is not obvious, resulting in low precision. We choose 5×5 window to extract GLCM texture features.

3) RANDOM FOREST CHANGE DETECTION METHOD

Pixel-based remote sensing change detection is to select the features of remote sensing images and judge whether the pixel points have changed through these features. Therefore, it can be regarded as a classification problem. Random forest (RF) is a non-parametric classification method driven by data. Multiple decision trees are integrated as a classifier, and each decision tree is a classifier. Random forest integrates all classification voting results and takes the category with the most votes as the final output result. Compared with other machine learning algorithms, the random forest has the advantages of high accuracy, few parameters, and strong robustness by learning and training with a given sample without prior knowledge [34].

In this paper, spectral features and texture features are combined with the random forest model to propose a pixel-based remote sensing image change detection algorithm using the random forest. The main processes of the pixel-based change detection method proposed in this paper include:

(1) Extract spectral difference features and GLCM texture features of corresponding pixels in two-phase high-resolution remote sensing images and construct the feature vector.

(2) N training samples are selected according to the principle of sample selection. The Bootstrap self-help method is used to randomly extract k sample subsets. Each sample subset has N samples, and the unsampled samples are used for prediction to evaluate errors.

(3) Determine sample feature dimension M, randomly select M features from it, and establish a decision tree model $T_k(x)$for k sample subsets.

(4) Use the decision tree model to classify the classification feature vector.

(5) Take all decision tree voting results, use the majority voting method for each pixel to obtain the final classification result and obtain the pixel level change detection result.

When samples are extracted by the Bootstrap method, the unsampled samples are called OOB (out-of-bag) data. According to the convergence of the random forest classification algorithm, with the increase of the number of sample subsets k, the OOB gradually decreases until it converges to a constant. At this time, the model constructed has the strongest generalization ability and the highest classification accuracy. Therefore, k in this paper takes the number of sample subsets when OOB begins to converge.

B. SEMANTIC SEGMENTATION BASED ON U-NET

Semantic segmentation refers to the classification of the categories to each pixel on the image belongs. Recently proposed full convolutional neural network for semantic segmentation of images can be utilized to segment images of any size. The U-net network used in this paper belongs to the improvement of the full convolutional neural network. The training of the improved U-net model is divided into three steps: data enhancement, network model training, and result optimization. After the U-net model is trained, the posterior phase remote sensing image that needs to be change detected is divided semantically using this model to obtain the segmentation results of objects in the image.

1) DATA ENHANCEMENT

The dataset used for the u-net model training in this paper is the data provided by the CCF Big Data Competition. This data set is the high-resolution remote sensing image of a city in south China in 2015, which contains 5 large-size RGB remote sensing images with labels, ranging from 3000×3000 to 8000×8000. The dataset samples are labeled with 5 types of objects, namely vegetation (mark 1), architecture (mark 2), water (mark 3), road (mark 4) and others (mark 0). Cultivated land, forest land and grassland are all classified as vegetation.

Since the image of the original data set is 5 large-sized remote sensing images, the number of images is small and the sizes are different, therefore, the dataset should be enhanced. First, the original image and label are cut into 256×256 small images by random cutting. Then, the image after cutting is flipped up and down, flipped left and right and rotated by 90° to enhance the generalization ability of the model. Part of the cutting graph is shown in Figure 2. After cutting, the operation of adjusting the light intensity and increasing the noise is implemented on some images. Through the above operation, we obtain 100,000 256×256 images, of which 70% were used for training, 20% for verification and 10% for testing.

FIGURE 2. Part of the Cut sub graph.

2) U-NET NETWORK MODEL TRAINING

The U-net network adopts a network structure including convolutional layers, downsampling layers and upsampling layers, which is U-shaped as a whole. The original U-Net network consisted of 19 convolutional layers, 4 downsampling layers and 4 upsampling layers, and Relu is used as the activation function. The downsampling layer is used to gradually display the environmental information, while the upsampling process is combined with the downsampling layer information and the upsampled input information to restore the detail information, and restores the image precision gradually. In order to improve the accuracy of network segmentation, we improve the U-net network structure. The improved network structure is shown in Figure 3. Compared with the original U-net model, the number of convolution layers is increased to extract features, and the batch normalization layer and dropout layer are added. The number of downsampling layers and upsampling layers is unchanged, and there is no full connection layer. The input image of the network is the cut subgraph of the training data after data enhancement, and the output image is the same size as the input image. The value of each pixel point represents the predicted category value of this point.

The left half of the network is a typical convolutional neural network structure, which extracts feature through convolution and downsampling operations. In this part, there are four sets of convolution operations and four down-sampling operations, among which each set of convolution operations adopts three consecutive convolution operations, one more convolution operation than the original structure. The size of the convolution kernel is 3×3. The step size is set at one, and zero padding is used. Each set of convolution operations is dimensioned by a downsampling operation and the Relu function is used as the activation function. The depth of the original U-net convolution gradually increases from

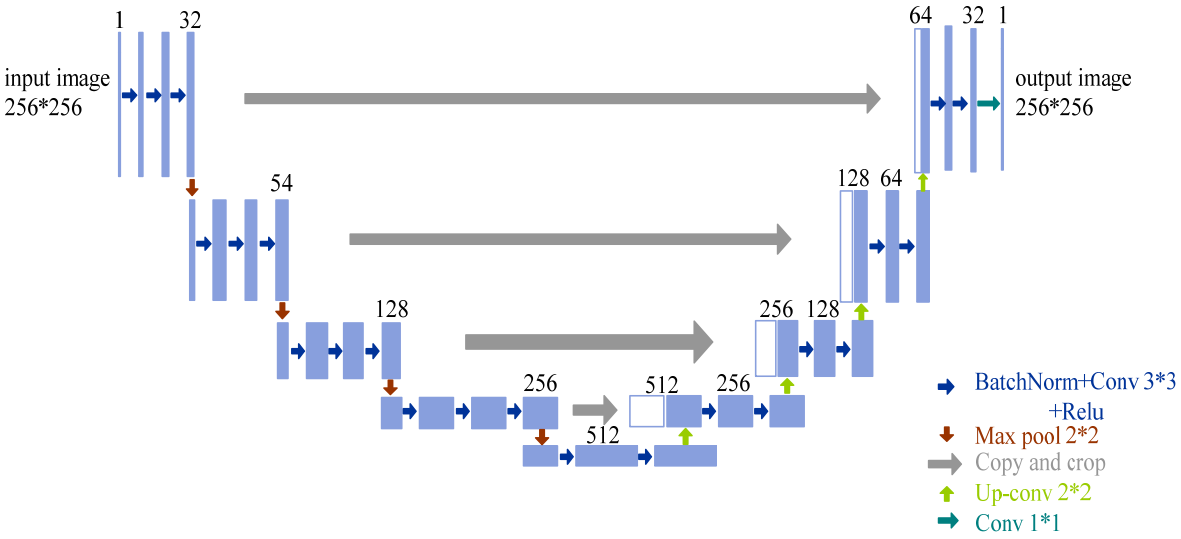

FIGURE 3. Schematic diagram of the network structure.

64 to 1024. In this paper, the number of convolution kernels is increased from 32 to 512, which makes the network easier to converge and the segmentation precision is higher.

The right half of the network consists of 4 sets of deconvolution operations and 4 upsampling operations. First, the features obtained by the up-sampling operation of the upper layer are merged with the same number of channels corresponding to the left side of the network, and then deconvolution is performed twice. The size of the convolution kernel used is also 3×3, unit step size, and zero padding is used. In addition, the dropout layer is added to randomly hide the 50% node weight, which can prevent over-fitting to some extent. The final layer of the network is a 1×1 convolutional layer that converts the feature vector into the number of classification results required and uses softmax as the classification function. The original U-net network does not have a batchnorm layer, and there will be problems such as slow learning speed, learning effect depending on the initial data distribution, and gradient explosion in the backpropagation process. In order to solve the above problems and accelerate the convergence speed of the model, the model proposed in this paper adds batch standardization layer before convolutional layer and deconvolution layer and normalizes the features of each layer. The batchnorm process formula is as follows:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{6}$$

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_B)^2 \tag{7}$$

$$\overline{x_i} = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \tag{8}$$

$$y_i = \gamma \overline{x_i} + \beta \tag{9}$$

where the meaning of parameters is shown in table 1. The normalization process destroys the feature distribution to a certain extent, so a reverse operation is required to extend and translate the normalized data.

In the back propagation of the neural network, a loss function is needed to represent the error between the real value and the predicted value, so as to optimize the network weight during the training process. The loss function in this paper uses a cross entropy function with the following formula:

$$loss = -\sum_{n} y \ln a + (1 - y)\ln(1 - a) \tag{10}$$

where y is the real value of the sample, and a is the actual output result. The optimizer used in this article is AdamOptimizer, and the learning rate is set to 0.001.

3) RESULT OPTIMIZATION

If only one network model is employed for training, the problem of selection preference will appear. Inspired by the voting majority of random forest, we use the integrated learning method to fuse the results of multiple networks to determine the final segmentation results. The results of FCN, the original U-net and the network in this paper are fused, and the voting strategy is still adopted to predict the category of each pixel. The category with the highest number of votes among the three network prediction results is the category of this pixel. Comprehensive consideration of the segmentation results of the three models can avoid the phenomenon of noise, remove the pixels with obvious classification errors, and improve the prediction ability.

C. OBTAIN THE RESULT OF CHANGE DETECTION

After the improved U-net network training, semantic segmentation is carried out for the last phase image that needed change detection. And then the pixel-level change detection

TABLE 1. The meaning of the parameters.

Parameter	Meaning
m	batch size
μ_B	mean value
σ_B^2	partial variance
ε	constant set for maintaining numerical stability
$\overline{x_i}$	normalized value
γ	Extended parameter
β	translation parameter
y_i	batch standardization result

result is fused with the semantic segmentation result. Finally, the false detection result is eliminated by combining the statistical principle to obtain the final change detection result. In this paper, the method in literature [14] is used for result fusion, and the principle is as follows:

$$label = \begin{cases} 1, & (\sum_{i\in\Omega} p_i) \Big/ Area > T \\ 0 & , else \end{cases} \tag{11}$$

where label represents the label of a pixel in the object Ω after the result fusion, p_i represents the prediction label of the pixel i in the object Ω, its value is 1 or 0. The area represents the pixel area of the object Ω, and T is the adaptive fusion threshold. The calculation formula of T is as follows:

$$T = \frac{1}{2 - Area/MaxArea} \tag{12}$$

where MaxArea represents the pixel area value of the object of the largest area in the image. There are still a few isolated points, cavities and other noise in the obtained results. In this paper, methods such as hole filling and the deletion of the small area false alarm area are used to optimize the prediction results so that the change detection results are more accurate.

D. ACCURACY EVALUATION

In order to evaluate the change detection accuracy of the proposed method, the confusion matrix is used to calculate the accuracy of the change detection result. The overall accuracy OA, false alarm rate FA and missed detection rate MA calculated by the confusion matrix are the commonly used evaluation indexes for the change detection accuracy. The overall accuracy represents the proportion of the correctly detected changing region and unchanged region in the real changing region and unchanged region, and the calculation formula is as follows:

$$OA = \frac{Cc + Uu}{T} \tag{13}$$

where Cc represents the number of pixels that the changing pixels are correctly detected as changes, Uu represents the number of pixels that the unchanged pixels are correctly detected as unchanged, and T is the total number of pixels of the detected image. The overall accuracy can reflect the detection accuracy as a whole. The false alarm rate indicates the proportion of unchanged region detected as the changed region in the real unchanged region. The calculation formula can be expressed as:

$$FA = \frac{Cu}{Cc + Cu} \tag{14}$$

where Cu represents the number of pixels in which unchanged pixels are detected as changed pixels, and virtual detection can be evaluated by the false alarm rate. The missed detection rate indicates the proportion of the unmeasured area in the real change area to the real change area. The calculation formula is as follows:

$$MA = \frac{Uc}{Cc + Uc} \tag{15}$$

where Uc is the number of pixels in which changing pixels are detected as unchanging pixels, and the missed detection rate can reflect the degree of missed detection of the change detection.

IV. EXPERIMENT AND ANALYSIS

In order to verify the effectiveness of the proposed method, we use two sets of high-resolution remote sensing images including different phases to conduct experiments to evaluate the accuracy of the change detection. Compared and analyzed with the detection accuracy of two traditional pixel-level change detection methods and this paper's pixel-level change detection method + the original U-net network, so as to verify the effect of this paper's method.

A. DATASETS DESCRIPTION

In this paper, two sets of data are used for the change detection experiment. The two sets of experimental datasets have been registered in advance, and the relative radiation correction and other operations have been performed on them.

The first set of data is high-resolution remote sensing images of a certain place obtained in 2006 and 2007. The image size is 725×941, and the image resolution is 2.5/m. The main land categories include buildings, farmland, and other vegetation and roads. As shown in Figure 4, where Figure 4 (a) is the image acquired in 2006, Figure 4(b) is the image acquired in 2007, Figure 4(c) is the ground truth, where white is the changing area and black is the area without change.

The second set of data is the high-resolution remote sensing image of a certain place obtained in 2000 and 2005, with the image size of 640×952 and the image resolution of 1.5/m.

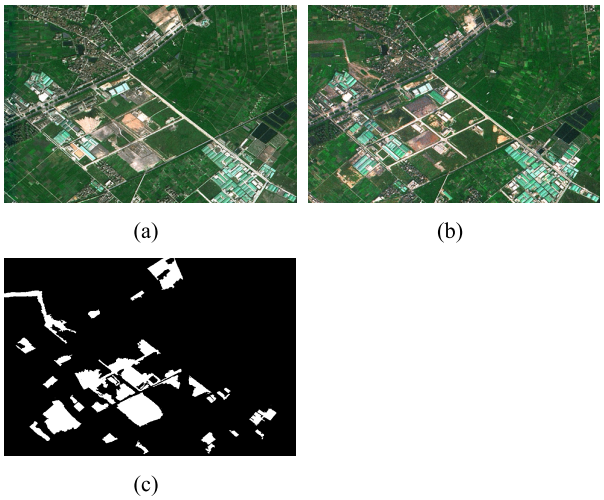

FIGURE 4. The first test dataset. (a) Remote sensing image in 2006; (b) Remote sensing image in 2007; (c) Ground truth for reference.

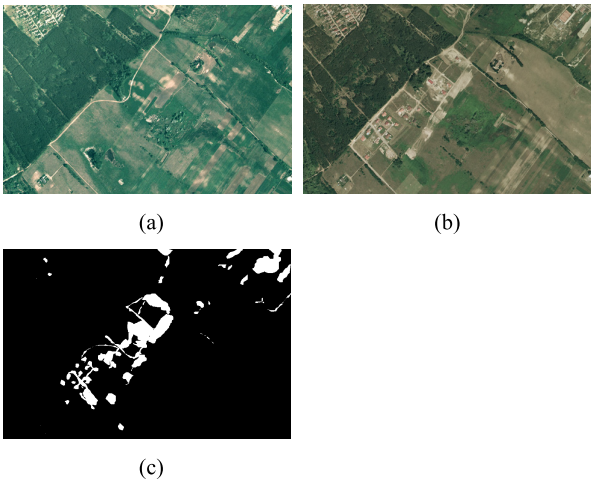

FIGURE 5. The second test dataset. (a) Remote sensing image in 2000; (b) Remote sensing image in 2005; (c) Ground truth for reference.

The main features include buildings and roads. As shown in Figure 5, where Figure 5(a) is the image acquired in 2000, Figure 5(b) is the image acquired in 2005, and Figure 5(c) is the change reference figure.

B. COMPARISON AND ANALYSIS OF CHANGE DETECTION RESULTS

The method in this paper is compared with the method of difference method, principal component analysis and the method of combining the pixel-level change detection method proposed in this paper and the original U-net network.

Figure 6 and Figure 7 are the results of the change detection results of the two groups of experiments, respectively. Among them, figure (a) (b) (c) (d) show the detection result of the difference method, principal component analysis and the method of combining the pixel-level change detection method proposed in this paper and the original U-net network. Compared with the first two methods of change detection, the method in this paper has higher precision and improves

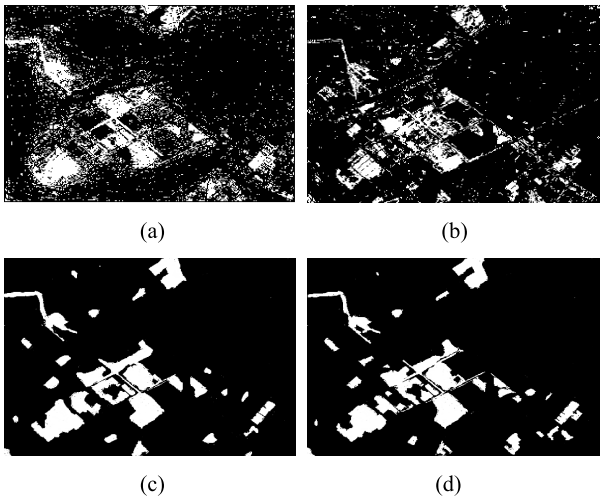

FIGURE 6. The change detection results of the first test dataset. (a) Difference method; (b) Principal component analysis method; (c) Pixel-level method in this paper + original U-net; (d) Method proposed in this paper.

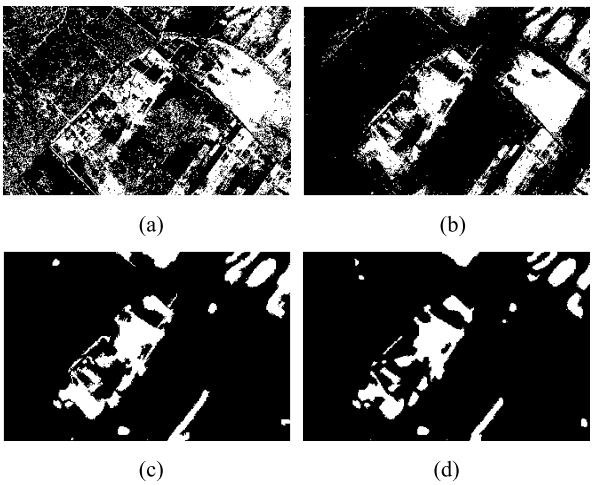

FIGURE 7. The change detection results of the second test dataset. (a) Difference method; (b) Principal component analysis method; (c) Pixel-level method in this paper + original U-net; (d) Method proposed in this paper.

the pepper and salt phenomenon in the pixel-level detection method. The changing area is more continuous, which reduces the pseudo-changes such as isolated noise. As for the improvement of the U-net network, the result of the method in this paper is more accurate than the original U-net network in the aspect of edge processing, and the changing region segmentation is more complete. The results of the two experiments show that the method proposed in this paper can improve the detection accuracy.

In order to compare the experimental results more intuitively, the quantitative analysis results of the two groups of experiments are shown in table 2 and table 3 respectively. Through these two tables, it can be seen that the method in this paper has higher accuracy, and it reduces the false alarm rate and the missed detection rate of detection results. In table 3, the false alarm rate of the median difference method and

TABLE 2. Evaluation of change detection of the first dataset.

	Detecting Changed Pixels	Detecting Unchanged Pixels	OA%	FA%	MA%
Difference Method	84142	598083	69.81	75.46	45.21
Principal Component Analysis Method	72698	609527	89.39	40.28	39.25
Pixel Method in This Paper + Original U-net	80882	601343	90.87	36.59	26.78
Method Proposed in This Paper	78198	604027	92.72	29.73	26.49

TABLE 3. Evaluation of change detection of the second dataset.

	Detecting Changed Pixels	Detecting Unchanged Pixels	OA%	FA%	MA%
Difference Method	203295	405985	69.47	870.8	25.38
Principal Component Analysis Method	130147	479133	79.75	80.68	28.57
Pixel Method in This Paper + Original U-net	75174	534106	90.47	45.21	23.71
Method Proposed in This Paper	63472	545808	91.74	40.71	23.30

the principal component analysis method are very high. The reason is that part of data set 2 is the change from a green land to cultivated land, and these two methods analyze the difference of the band information of the image, so this part will be detected as the changing area. In this paper, both cultivated land and green land are marked as vegetation. When the image is segmented, this part is divided into one category. Since the labels are the same, they will not be segmented, so they are ultimately determined as unchanged areas.

The reason why false alarm and missed detection still exist in the method in this paper is that when the semantic segmentation is carried out, the segmentation of part of roads and cultivated land and wasteland is not accurate enough, resulting in the segmentation error of test objects. Although the MA of the method proposed in this paper is closer to method three, the FA is better than method three. It can be seen that the change detection performance of the algorithm in this paper is more superior.

V. CONCLUSION

In this paper, a remote sensing image change detection method combining pixel-level and object-level is proposed, which combines the advantages of the two methods and solves the problem of salt and pepper phenomenon existing in the traditional pixel-level method, and the changing area is more complete. In this paper, the advantage of simple and easy operation of the pixel-level method is utilized to obtain the preliminary change detection results by using the random forest algorithm, and then the post-phase image semantic segmentation result is merged with the pixel-level change detection result to obtain the change region. Experiments were carried out on two sets of high-resolution remote sensing image datasets. The results demonstrate that the proposed method has a high detection accuracy. However, the method proposed in this paper still has some shortcomings and can only detect the changing area. The next step is to continue to improve the U-net model based on the work of this paper to improve the accuracy and achieve the ''from-to'' detection of the changed object.

REFERENCES

[1] H. Song, R. Srinivasan, T. Sookoor, and S. Jeschke, *Smart Cities: Foundations, Principles, and Applications*. Hoboken, NJ, USA: Wiley, 2017, pp. 1–906.

[2] W. Zou, W. Jing, G. Chen, Y. Lu, and H. Song, ''A survey of big data analytics for smart forestry,'' *IEEE Access*, vol. 7, pp. 46621–46636, 2019. doi: 10.1109/ACCESS.2019.2907999.

[3] Z. Zhao, Y. Meng, A. Yue, Q. Huang, Y. Kong, Y. Yuan, X. Liu, L. Lin, and M. Zhang, ''Review of remotely sensed time series data for change detection,'' *J. Remote Sens.*, vol. 20, no. 5, pp. 1110–1125, Jun. 2016. doi: 10.11834/jrs.20166170.

[4] B. Jiang, J. Yang, Z. Lv, and H. Song, ''Wearable vision assistance system based on binocular sensors for visually impaired users,'' *IEEE Internet Things J.*, to be published.

[5] B. Jiang, J. Yang, H. Xu, H. Song, and G. Zheng, ''Multimedia data throughput maximization in Internet-of-Things system based on optimization of cache-enabled UAV,'' *IEEE Internet Things J.*, to be published.

[6] T. Leichtle, C. Geiß, M. T. Lakes, and H. TaubenbÃűck, ''Unsupervised change detection in VHR remote sensing imagery—An object-based clustering approach in a dynamic urban environment,'' *Int. J. Appl. Earth Observ. Geoinf.*, vol. 54, pp. 15–27, Feb. 2017.

[7] G. F. Tong, L. Yong, D. Weili, and Y. Xiaoyang, ''Review of remote sensing image change detection,'' *J. Image Graph.*, vol. 20, no. 12, pp. 1561–1571, 2015. doi: 10.11834/jig.20151201.

[8] A. P. Tewkesbury, A. J. Comber, N. J. Tate, A. Lamb, and P. F. Fisher, ''A critical synthesis of remotely sensed optical image change detection techniques,'' *Remote Sens. Environ.*, vol. 160, pp. 1–14, Apr. 2015.

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

[10] M. Gong, Z. Zhou, and J. Ma, ''Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering,'' *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.

[11] J. R. Hu and Y. Z. Zhang, ''Seasonal change of land-use/land-cover (LULC) detection using MODIS data in rapid urbanization regions: A case study of the pearl river delta region (China),'' *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 4, pp. 1913–1920, Aug. 2013.

[12] Z. X. Zhang and Y. H. Zhang, ''Remote sensing research issues of the national land use change program of China,'' *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 6, pp. 461–472, Dec. 2007.

[13] J. S. Deng, K. Wang, Y. H. Deng, and G. Qi, ''PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data,'' *Int. J. Remote Sens.*, vol. 29, no. 16, pp. 4823–4838, 2008.

[14] F. Z. Dou, H. C. Sun, X. Sun, W. H. Diao, and K. Fu, ''Remote sensing image change detection method based on DBN and object fusion,'' *Comput. Eng.*, vol. 44, no. 4, pp. 294–298 and 304, 2018.

[15] M. Zhao and Y. D. Zhao, ''Object-oriented and multi-feature hierarchical change detection based on CVA for high-resolution remote sensing imagery,'' *J. Remote Sens.*, vol. 22, no. 1, pp. 119–131, Jan. 2018.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ''ImageNet classification with deep convolutional neural networks,'' in *Proc. Adv. Neural Inf. Process. Syst.*, Stateline, NV, USA, Dec. 2012, pp. 1097–1105.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Apr. 2015, *arXiv:1409.1556v6*. [Online]. Available: https://arxiv.org/abs/1409.1556v6
[18] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. 28th IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Mar. 2015, pp. 3434–3440.
[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," May 2015, *arXiv:1505.04597*. [Online]. Available: https://arxiv.org/abs/1505.04597
[20] F. Gao, J. Dong, B. Li, and Q. Xu, "Automatic change detection in synthetic aperture radar images based on PCANet," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1792–1796, Dec. 2016.
[21] Y. Y. Dian, S. H. Fang, and C. H. Yao, "The geographic object-based method for change detection with remote sensing imagery," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 39, no. 8, pp. 906–912, 2014.
[22] M. G. Gong, Y. Cao, and Q. Wu, "A neighborhood-based ratio approach for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 307–311, Mar. 2012.
[23] K. M. Sun and Y. Chen, "The application of objects change vector analysis in object-level change detection," in *Proc. Int. Conf. Comput. Intell. Ind. Appl.*, Hong Kong, 2010, pp. 383–389.
[24] W. P. Quan, W. H. Li, and X. C. Li, "Detecting the change of high-resolution remote sensing images by adaptive multi-feature fusion," *Electron. Opt. Control*, vol. 22, no. 3, pp. 45–49, 2015.
[25] Y. Wang, N. Shu, and Y. Gong, "A study of land use change detection based on high resolution remote sensing images," *Remote Sens. Land Resour.*, vol. 24, no. 1, pp. 43–47, Mar. 2012.
[26] G. Cao, Y. P. Li, Y. Liu, and Y. Shang, "Automatic change detection in high-resolution remote-sensing images by means of level set evolution and support vector machine classification," *Int. J. Remote Sens.*, vol. 35, no. 16, pp. 6255–6270, Aug. 2014.
[27] K. Tang, B. Zou, Z. H. Tang, K. Lin, and B. Liu, "A detection method of remote sensing images change by fusedpixel-level and object-level," *Sci. Surv. Mapping*, vol. 42, no. 5, pp. 106–112, May 2016.
[28] M. H. Xiao, W. Q. Feng, and H. G. Sui, "Remote sensing image change detection algorithm based on super-pixel segmentation and multiple difference maps," *Bull. Surv. Mapping*, no. 10, pp. 93–97 and 121, Mar. 2018.
[29] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 580–587.
[30] Z. Shu, X. Hu, and J. Sun, "Center-point-guided proposal generation for detection of small and dense buildings in aerial imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 7, pp. 1100–1104, Jul. 2018.
[31] M. Kampffmeyer, A. B. Salberg, and R. Jenssen, "Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Las Vegas, NV, USA, Jun./Jul. 2016, pp. 680–688.
[32] Q. Q. Hou, F. Wang, and L. Yan, "Extraction of color image texture feature based on gray-level co-occurrence matrix," *Remote Sens. Land Resour.*, vol. 25, no. 4, pp. 26–32, 2013.
[33] P. Xiao, X. Zhang, D. Wang, M. Yuan, X. Feng, and M. Kelly, "Change detection of built-up land: A framework of combining pixel-based detection and object-based recognition," *ISPRS J. Photogramm. Remote Sens.*, vol. 119, pp. 402–414, Sep. 2016.
[34] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 67, pp. 93–104, Jan. 2012.

LU XU received the B.E. degree in computer science and technology from Northeast Forestry University, Harbin, China, in 2017, where she is currently pursuing the M.E. degree. Her current research interests include computer vision and machine learning.

WEIPENG JING received the Ph.D. degree from the Harbin Institute of Technology, China. He is currently an Associate Professor with Northeast Forestry University, China. He has published over 50 research articles in refereed journals and conference proceedings, such as CPC, PUC, and FGCS. His research interests include modeling and scheduling for distributed computing systems, fault tolerant computing and system reliability, cloud computing, and spatial data mining.

HOUBING SONG (M'12–SM'14) received the M.S. degree in civil engineering from the University of Texas, El Paso, TX, USA, in 2006, and the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, USA, in 2012.

In 2017, he joined the Department of Electrical, Computer, Software, and Systems Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA, where he is currently an Assistant Professor and the Director of the Security and Optimization for Networked Globe Laboratory (SONG Lab). He has served on the Faculty of West Virginia University, from 2012 to 2017. In 2007, he was an Engineering Research Associate with the Texas A&M Transportation Institute. He is the editor of six books, including *Big Data Analytics for Cyber-Physical Systems: Machine Learning for the Internet of Things* (Elsevier, 2019), *Smart Cities: Foundations, Principles, and Applications* (Hoboken, NJ: Wiley, 2017), *Security and Privacy in Cyber-Physical Systems: Foundations, Principles, and Applications* (Chichester, U.K.: Wiley-IEEE Press, 2017), *Cyber-Physical Systems: Foundations, Principles and Applications* (Boston, MA: Academic Press, 2016), and *Industrial Internet of Things: Cybermanufacturing Systems* (Cham, Switzerland: Springer, 2016). He is the author of more than 100 articles. His research interests include cyber–physical systems, cybersecurity and privacy, the Internet of Things, edge computing, big data analytics, unmanned aircraft systems, connected vehicle, smart and connected health, and wireless communications and networking.

Dr. Song is a Senior Member of the ACM. He was a recipient of the Navigation and Surveillance Technologies (ICNS) Conference, the very first recipient of the Golden Bear Scholar Award, the highest campus-wide recognition for research excellence at the West Virginia University Institute of Technology (WVU Tech), in 2016, the prestigious Air Force Research Laboratory's Information Directorate (AFRL/RI) Visiting Faculty Research Fellowship, in 2018, and the Best Paper Award from 2019 Integrated Communication. He has served as an Associate Technical Editor for the *IEEE Communications Magazine* and a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (J-SAC), IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and IEEE NETWORK.

GUANGSHENG CHEN is currently a Doctoral Supervisor and a Professor with Northeast Forestry University, China. He is also a member of the National Innovation Methods Research Institute and the Executive Director of the Education Information Technology Council of Education Ministry. He has published over 30 academic papers and one monograph. His research interests include biomass material prediction, intelligent detection of new composite materials, and big data on forestry.