Geo-Parcel-Based Change Detection Using Optical and SAR Images in Cloudy and Rainy Areas

Nan Zhou, Xiang Li, Zhanfeng Shen, Tianjun Wu, and Jiancheng Luo

***Abstract*—In this article, we deal with the problem of change detection in cloudy and rainy areas using multisource remote sensing images. While previous methods mostly focus on change detection on pixel or super-pixel levels, in this article, we introduce the concept of geo-parcel and use it as the basic processing unit for our change detection method. Concretely, we first extract geo-parcel from an optical high spatial resolution remote sensing image. Then, we divide each geo-parcel into fine-grained segments with refined boundaries using image segmentation methods. These fine-grained segments are used as the basic processing units for our change detection method. After that, an unsupervised learning-based method is adopted to obtain the difference map by comparing synthetic aperture radar images of two periods. Training samples with labels are automatically generated from the difference map. Finally, a deep neural network is trained using the generated samples and is further used to predict the refined change map. Experiments on the collected images from Gui'an, Guizhou Province, China demonstrate the effectiveness of the proposed method for change detection in a cloudy and rainy area with an overall accuracy surpasses 94%.**

***Index Terms*—Change detection, cloudy and rainy, geo-parcel, multisource images.**

I. Introduction

CHANGE detection is a long-standing problem in the field of remote sensing (RS) and has been widely used in many real-world applications [1], [2]. Existing change detection methods can be roughly divided into two categories; pixel-based method and object-based methods.

In pixel-based methods, each pixel in the image is treated as the basic analysis unit, without considering the relations among pixels. In general, these methods start with some algebraic operations to generate the difference images, followed by thresholding and clustering methods to produce the change detection results [3]–[5]. For example, Rignot *et al.* proposed a change detection method for ERS-1 SAR data based on differences in the magnitude of the signal intensity between two dates. Nevertheless, the pixel-based methods have limited capabilities when dealing with high-resolution images, where neighboring pixels have strong correlations and contextual information should be taken into consideration [6]. However, the capability of pixel-based change detection methods is limited when applied to high-resolution images because the pixels are not spatially independent. What is more, the increased variability present in high-resolution images and the difficulty of modeling the contextual information further weaken the performance of traditional pixel-based change detection approaches.

Unlike pixel-based methods, object-based methods conduct change detection on the object level. These methods usually include three main components: image segmentation, object-level feature extraction, and object classification. Extensive research attention has been paid to improve the quality of image segmentation, feature engineering, as well as classification modules in order to improve the change detection performance. Recently, with the prosperity of deep learning methods in the remote sensing field, researchers also developed various deep learning-based methods for change detection [6], [7]. One can refer to [8] for a brief summary of the object-based change detection methods.

In remote sensing, optical images and synthetic aperture radar (SAR) images are generally used to conduct change detection. Optical images are easy to be interpreted and cheap to be obtained. However, optical RS images are sensitive to atmospheric and sunlight conditions and cannot meet the need for change detection in cloudy and rainy conditions. On the contrary, SAR images are hard to interpret but can be collected in all weather and all time. It, therefore, has the advantage of being insensitive to atmospheric and sunlight conditions. To take the advantages of these two types of images, many researchers also develop change detection methods that combine these two kinds of data [9], [10]. For example, Zhao *et al.* [9] proposed a collaborative change detection method for optical RS images and SAR images based on deep learning. It overcomes the consistency problem of the reference space for RS image change detection between different sensors by projecting different types of RS images to a unified feature space. Although these methods use optical images in coordination with SAR images, and they all take advantage of

Manuscript received August 8, 2020; revised November 9, 2020; accepted November 10, 2020. Date of publication November 16, 2020; date of current version January 6, 2021. This work was supported in part by the National Key Research and Development Program of China under Grants 2017YFB0504204 and 2018YFB0505000, in part by the National Natural Science Foundation of China under Grant 41971375 and 41631179, and in part by the Xinjiang Uygur Autonomous Region Flexible Talent Award in 2018. *(Nan Zhou and Xiang Li contributed equally to this work.) (Corresponding author: Zhanfeng Shen.)*

Nan Zhou is with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100049, China, and also with the University of Chinese Academy of Sciences, Beijing 100864, China (e-mail: zhounan@aircas.ac.cn).

Xiang Li is with the NYU Multimedia, and Visual Computing Lab, NYU Tandon, and Abu Dhabi, Abu Dhabi 129188, UAE (e-mail: xl1845@nyu.edu).

Zhanfeng Shen and Jiancheng Luo are with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100049, China (e-mail: shenzf@aircas.ac.cn; luojc@radi.ac.cn).

Tianjun Wu is with the School of Science, Chang'an University, Xi'an 710064, China (e-mail: tjwu@chd.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2020.3038169

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

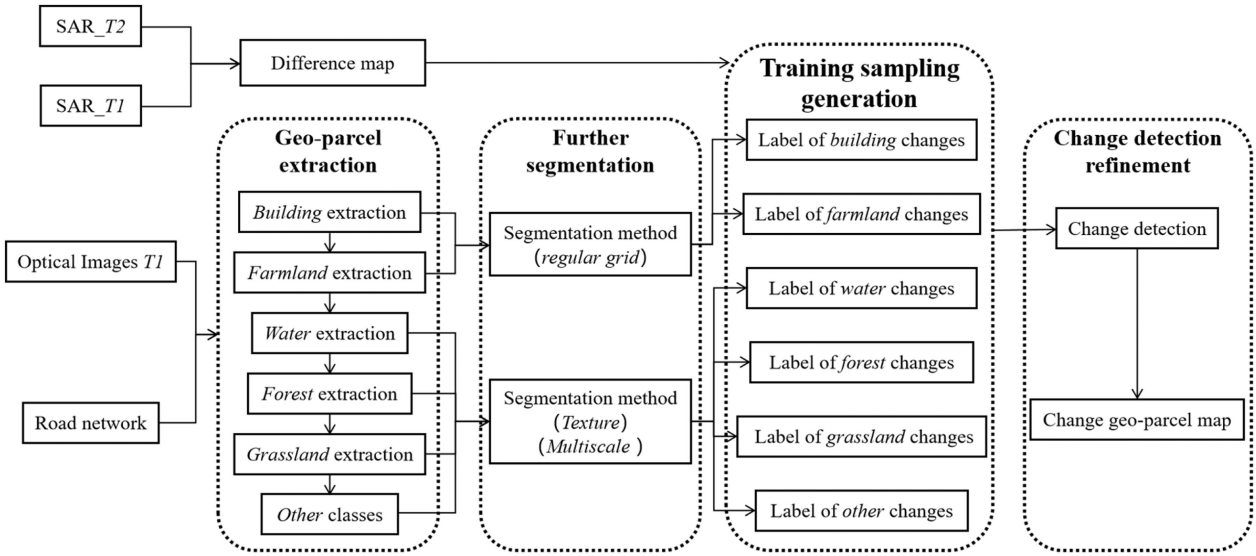


Fig. 1. Overview of the proposed geo-parcel based method for change detection. Our method takes as input early SAR and optical images acquired at the same time, the road network data, and the SAR images acquired at the detection time. First, we use a hierarchical classification method to extract geo-parcels for each land cover type from the optical image and produce the geo-parcel map, including buildings, farmland, water, forest, grassland, and other land types those cannot be extracted. The road network is used as guidance for geo-parcel partition. Second, we divide each geo-parcel into fine-grained segments using various kinds of segmentation algorithms. The segmentation method is chose based on the land cover type. These fine-grained segments will be used as the basic units in the process of change detection. Third, the difference map is extracted from the SAR images acquired in different periods and is further matched with fine-grained segments to obtain training samples. Finally, a deep neural network-based classification model is trained using the generated samples to refine the change detection results. The refined classification results are mapped back to geo-parcels and produce the final geo-parcels-based change map.

multiphase optical images, it is hard to obtain the two-phase optical images in a cloudy and rainy area. Therefore, it is still difficult to carry out the application of change detection in cloudy and rainy areas.

In this article, we propose a geo-parcel-based method for change detection in cloudy and rainy areas. Our method takes the advantages of optical images for accurate geo-parcel extraction and conducts change detection on SAR images with the geo-parcel as basic processing units. The main contributions of this article are summarized as follows.

1) Different from the traditional pixel/object-level change detection method, we propose a spatio-temporal collaborative method for changes extraction on the geo-parcel level. The change detection results are, therefore, more in line with the morphological characteristics of ground objects.
2) To alleviate the need for a large number of samples for deep neural network training, our method can automatically generate training samples using an unsupervised classification method. The proposed method is, therefore, much more efficient and labor-saving.

II. Methodology

Our proposed method for change detection in rainy and cloudy areas contains four main components (see Fig. 1). The first component is "Geo-parcel extraction." In this component, we leverage deep neural networks to extract the geo-parcels for each land cover type using a hierarchical classification method. The second component is "Further segmentation in geo-parcels." In this component, we divide each geo-parcel into fine-grained segments using various kinds of segmentation algorithms. These fine-grained segments will be used as the basic units in the process of change detection. The third component is "Training sampling generation." In this component, the difference map between two SAR images is obtained on the basis of the logarithmic ratio method. A threshold-based method is then adopted to determine whether a segment is changed or not. Note that the generated samples may include incorrect labels due to data noise and improper threshold. The fourth component is "Change detection refinement." In this component, a deep neural network model with more powerful generalization abilities is trained on generated samples in order to produce refined change detection results.

A. Geo-Parcels Extraction

The features in the optical RS images show obvious visual differences. Especially, roads and river systems divide the complex land surface by continuous spatial extension. Generally speaking, artificial land objects have regular boundaries, uniform texture, and consistent color. In contrast, the natural land objects have blurred boundaries, mottled texture, and inconsistent color. Therefore, traditional low-level feature-based classification methods tend to have inferior classification performance and limited generalization abilities. In this article, we use a hierarchical classification method to extract geo-parcels for each land cover type from the optical image and produce the geo-parcel map, including buildings, farmland, water, forest, grassland, and other land types that cannot be extracted. A

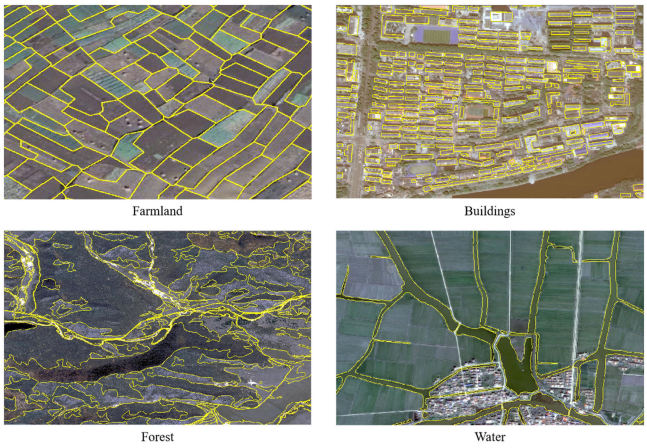


Fig. 2. Geo-parcels of farmland, buildings, forest, and water.

geo-parcel is generally defined as the smallest geographical entity for constructing geographical scenes that can be visually interpreted with determined land-use properties. The optical image is divided into different blocks under the district controlling of linearly extending features (roads, rivers, transition zones, etc.). For geo-parcel extraction, we train separate models for different land cover types. Specifically, we use TernausNetV2 [11] for building extraction, CRF [12] for farmland extraction, and D-LinkNet [13] model for the extraction of other classes.

We note that for geo-parcel extraction, there are sufficient optical datasets that can be used to train a model for geo-parcel extraction. Optical images are easy to be accessed and interpreted. One can use some public datasets, such as ISPRS 2D Semantic Labeling Dataset [14], 2017 IEEE GRSS Data Fusion Contest [15], Gaofen Image Dataset (GID) [16], to pretrain a model and then fine-tune the model in the experimental areas. For regions without a sufficiently dense road network, one can use public road network datasets, such as Massachusetts Roads Dataset [17] and SpaceNet Roads Dataset [18], to pretrain a model and then fine-tune the model in the experimental areas. Fig. 2. shows some examples of the extracted geo-parcels.

B. Fine-Grained Segmentation in Geo-Parcels

After generating the semantic geo-parcel map, we further divide each geo-parcel into fine-grained segments, which will be used as the basic units for change detection. Note that the land objects with different land use types have quite different spectral and texture characteristics. We, therefore, develop a series of image segmentation algorithms based on the land use type of geo-parcels. For the geo-parcels with artificial land use types, such as building, farmland, and road, their boundaries are more regular. We use simple linear iterative clustering [19] algorithm to generate an orderly division segments unit in the range of the corresponding geo-parcel. For the geo-parcels with nonartificial land use types such as forest, grassland, and water, their boundaries are irregular. We adopt the Quickshift image segmentation algorithm [20], which is sensitive to the change of color information and is more suitable for the subdivision of these irregular geo-parcels.

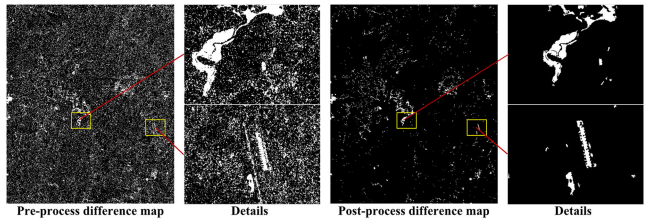


Fig. 3. Difference maps before (left) and after (right) post-processing.

C. Training Sample Generation

After getting the fine-grained segmentation results, a direct way to get the change detection results is to compare the difference map of two SAR images at different times. However, SAR images tend to include various types of data noise. While directly applying change detection algorithms on the noised inputs will lead to incorrect change detection results. Therefore, we propose to conduct an image processing on the difference map and fuse the processed difference map with geo-parcels to generate the change samples. In this way, our method not only retains the morphology outlines of the changing parts but also reduces the time of training samples production.

In this article, we use the logarithmic ratio method proposed in [3] to generate training samples. The different map is calculated as: $D_t = |\log(\mathcal{I}1) - \log(\mathcal{I}2)|$. where $\mathcal{I}1$ and $\mathcal{I}2$ denote the signals (values) of SAR images before and after change, respectively. After that, a Gaussian filter is then used to eliminate data noise and generate a smooth difference map. Finally, a pixel-level change map can be generated by using a simple technique of threshold binarization (see Fig. 3). In this article, we use the Otsu algorithm [21] that automatically determines the optimal threshold to get the pixel-level change map.

To determine whether a segment is changed or not, we further transfer the pixel-level change map to each segment. To achieve this, we calculate the proportion of changed pixels in each segment. A suitable threshold value is set to determine the changing pattern for each segment. Here, when the proportion of changed pixels of a segment is less than 0.5, we determine this segment as no change. When the proportion value gets greater than 0.8, we determine this segment as changed. When the proportion value of a segment falls between 0.5 and 0.8, we further divide this segment into smaller segments until each small piece can be determined as changed or not. The generated samples may include incorrect labels due to data noise and improper threshold, so we manually select some training samples according to the difference map.

D. Change Detection Refinement

After generating training samples, we develop a deep neural network-based classification model to further refine the change detection results. We note that deep neural networks have more powerful feature extraction abilities and are more robust to input variations (e.g., color variations, textural variations). In this article, we use the D-LinkNet [13] to build the classification model. The D-LinkNet model is extended from Link-Net [22]

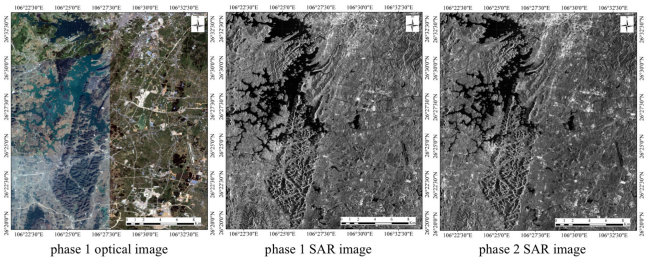


Fig. 4. Two phases of optical and SAR images in Gui'an district.

by introducing dilated convolutional layers and pretrained technique to the encoder part. The introduced dilated convolutional layers enlarge the receptive field and strengthen classification capability. The detailed network architecture of D-LinkNet can be found at [13]. Finally, the refined classification results are mapped back to geo-parcels and produce the final geo-parcels-based change map.

Note that, in this article, we do not intend to develop a new method for image classification. We focus on change detection in cloudy and rainy areas where real-time optical images are inaccessible, but the SAR images are easy to access. We introduce a brand-new framework for change detection in cloudy and rainy areas. It would be possible to use other state-of-the-art methods, such as DeepLabv3+ [23] to get more accurate classification, but this is not the focus of our paper.

III. Experiments

A. Study Area and Materials

We conduct change detection experiments in the Gui'an district, one of the rainy and cloudy areas in southwest China. According to historical weather statistics, from 2011 to 2019, the Gui'an district has only 90 sunny days. As of January 31, 2019, there are only 24 scenes of Chinese Gaofen2 (GF2) RS images with less than 5% cloud coverage in the Gui'an district, which are not enough to cover the whole Gui'an district. So, the optical images in our experiments were collected from Google Earth. The acquired images contain three-band of RGB channels with a spatial resolution of 0.5 m. The image was captured in June 2017. The SAR images in our experiments were collected from the Cosmo-SkyMed satellite. The SAR image of the first phase was captured on June 29, 2017. The SAR image of the second phase was captured on January 30, 2018. The preprocessed SAR images have a spatial resolution of 2.5 m. Both SAR images have been geometrically registered with the optical image under WGS-84 coordinates (see Fig. 4). Also, we perform preprocessing for SAR images to reduce speckle using ENVI software.

B. Results and Analysis

In our experiments, we collect 20 image tiles from GoogleEarth in the Guizhou area, China. Each tile has a size of 2000x2000 pixels. We annotate the images with the six land cover classes, including building, farmland, water, forest, grassland, and others. We build this large-scale benchmark dataset to enable geo-parcel extraction in cloudy and rainy areas.

TABLE I
Detailed Information of the Geo-Parcel Map

Land types	IoU	Area	Number of geo-parcels
Water	0.88	35.84	155
Building	0.85	0.95	2324
Farmland	0.91	35.51	23367
Forest	0.82	42.1	244
Grassland	0.79	0.72	197

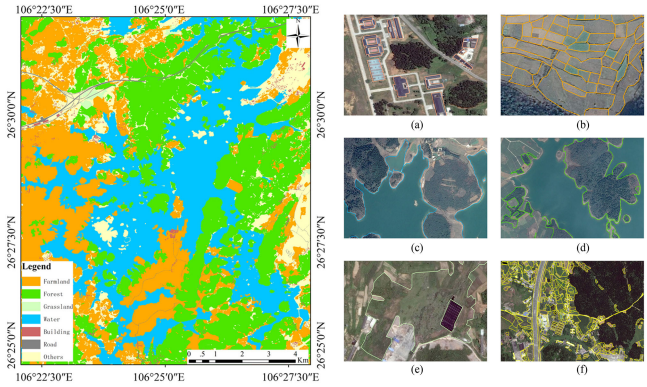


Fig. 5. Left figure shows the overall geo-parcel map of the Gui'an area. The right figure shows examples of the geo-parcel results for different land use types, (a) building, (b) farmland, (c) water, (d) forest, (e) grassland, (f) merged geo-parcels.

For model evaluation, part of the extracted geo-parcels were selected as the test set. The testing area covers 129.4 km^2, where water area is 35.8 km^2 (27.67% in testing are) and a total of 155 geo-parcels are extracted; building area is 0.95 km^2 (0.73%), and a total of 2324 geo-parcels are extracted; farmland area is 35.5 km^2 (27.43%) and a total of 23 367 geo-parcels are extracted; forest area is 42.1 km^2 (32.53%) and a total of 244 geo-parcels are extracted; grassland area is 0.72 km^2 (0.56%) and a total of 197 geo-parcels are extracted.

In our experiment, four metrics are used to evaluate the performance of the proposed method for change detection, including precision, recall, F1-score, and intersection-over-union (IoU). IoU indicates geo-parcel extraction performance. Detailed information are shown in Table I.

In order to obtain a more reliable geo-parcel map for change detection, we add a small amount of manual editing to the automatically extracted results. Fig. 5 shows the geo-parcel extraction results.

After getting the geo-parcel map, we divide each geo-parcel into fine-grained segments using image segmentation techniques. For the geo-parcels with artificial and nonartificial land use types, we adopt different segmentation methods. Then, we integrate the fine-grained segmentation results with the SAR difference map to get the change samples. Fig. 6 shows some training samples.

TABLE II
SAMPLES EXPERIMENTAL RESULTS

Method	Manual samples	Automatic samples	Accuracy	Precision	Recall	F1-score	Time
Ratio	0	0	0.8117	0.4201	0.4223	0.4212	0.5h
D-linkNet	15	0	0.8836	0.5794	0.8552	0.6908	6.2h
FCN	0	30	0.9174	0.6559	0.7694	0.7081	1.5h
Ours	0	30	0.9403	0.7405	0.8236	0.7798	1.6h

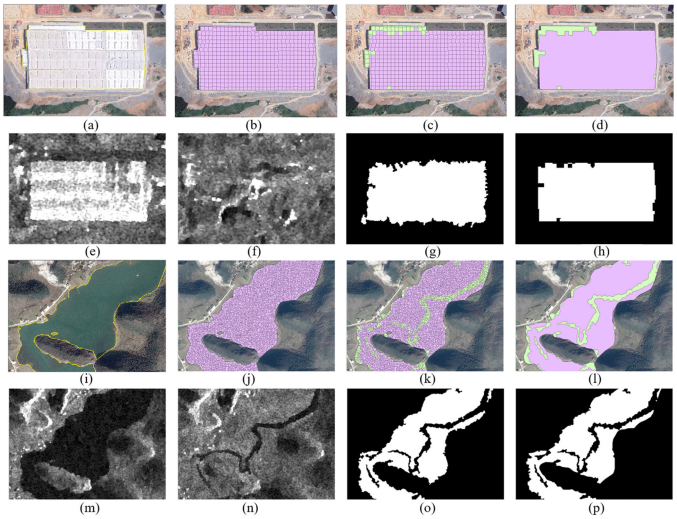


Fig. 6. Illustration of geo-parcels fine-grained segmentation and training sample generation. The geo-parcels with regular boundaries, such as buildings, are segmented in the form of the grid, (a)–(h). (a) building geo-parcels, (b) fine-grained segments, (c) segments with difference values, (d) merged parcels, (e) before SAR image, (f) after SAR image, (h) difference map, (i) change samples. The geo-parcels with irregular boundaries such as water are segmented in the form of texture, (i)–(p). (i) water geo-parcels, (j) fine-grained segments, (k) segments with difference values, (l) merged parcels, (m) before SAR image, (n) after SAR image, (o) difference map, (p) change samples.

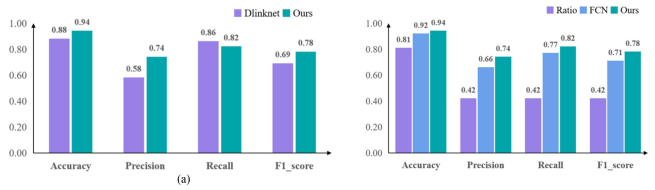


Fig. 7. Quantitative change detection performance of different methods.

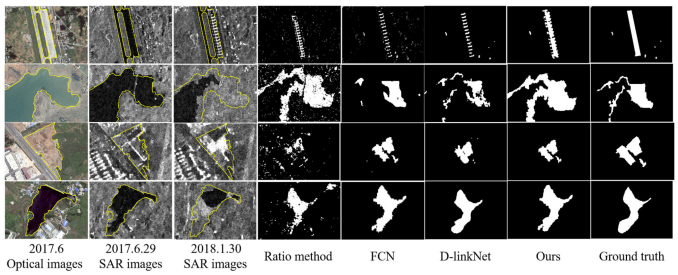


Fig. 8. Qualitative change detection performance of different methods.

Finally, we select 30 samples as the training samples, which were generated by the proposed method. These training samples are used to train our D-linkNet model and the comparing FCN [24] model. To compare the efficiency for sample generation, we manually draw 15 samples in the test area and train another D-linkNet with the same experimental settings. All experiments are done on a windows server with a GTX Titan X GPU.

Table II shows the change detection performance and overall processing time of four comparing methods: 1) ratio method, using the ratio method to obtain the difference map, and postprocessing techniques to get the changed geo-parcels; 2) D-linkNet method, using 15 manual samples to train the model and predict the changed geo-parcels; 3) FCN method, using 30 automatic samples to train the FCN model and predict the changed geo-parcels; 4) Our proposed method, using 30 automatic samples to train the D-linkNet model and predict the changed geo-parcels.

In Fig. 7(a), we use D-linkNet to train manual samples and automatically generated samples and use ground truth for comparative analysis. It is difficult to draw the change detection label data, especially when the boundary of the changed geo-parcels is difficult to define. Deviations are likely to occur in the drawing process. For this reason, we have drawn a total of 15 test area changed label data. The average time consumption for drawing a sample is 20 min. Under the same training parameters, the overall accuracy of the model obtained by hand-drawn sample training is about 88%, which takes 6.2 h in total. The overall accuracy obtained by our proposed method reaches 94%, and the whole process takes 1.6 h. The efficiency has been greatly improved.

In Fig. 7(b), we use the automatically generated samples to train different models and compare the evaluation metrics with those of the ratio method. There is usually a misalignment between the difference map obtained by the traditional ratio method and the real changed geo-parcels, which leads to low accuracy of the ratio method. After obtaining training samples using our method, FCN and D-linkNet models are trained on the 30 automatically generated samples. The FCN model gets an overall accuracy of about 92%. Our D-linkNet model achieves an overall accuracy of about 94%. Compared to the FCN model, the proposed method achieves a higher accuracy for change detection in the study area.

Fig. 8 shows selected examples of the results generated by all comparing methods. From the figure, we can see that there are obvious speckle effects in the changed geo-parcels obtained by the ratio method, which further leads to a higher error rate. Using hand-drawn samples to train D-linkNet, there are still obvious mismatches between the predicted results and the

ground truth. The main reason is that it is hard to control the quality of hand-drawn samples and the number of hand-drawn samples is too small to meet the requirements of training a deep neural network model. More importantly, by using the proposed automatic sample generation technique, our method can greatly improve the efficiency of supervised deep learning models.

We also show the change detection results using automatically generated samples. As shown in Fig. 8, with automatically generated samples, both FCN and our model produce better results than the model trained using hand-drawn samples and unsupervised ratio method. Moreover, by using dilated convolutional layers with shortcuts in D-linkNet model, the boundaries of extracted changed geo-parcels are more in line with the real ones.

IV. Conclusion

This article proposed an accurate and efficient method for change detection in cloudy and rainy areas using multisource remote sensing images. Our method starts with extracting geo-parcels for each land cover type using a hierarchical classification method from optical images. The difference map was calculated by using two-phase SAR images, and the sample data required by change detection was obtained by fine-grained segmentation in the geo-parcels. Finally, the results of change detection were obtained through a deep learning model. In the experiment of the Gui'an test area with cloudy and rainy weather, the accuracy of our change detection method reached 94%. It was shown to be able to quickly discriminate the changed geo-parcels. In conclusion, this method has the possibility of application promotion.

References

[1] L. Li, C. Wang, H. Zhang, B. Zhang, and F. Wu, "Urban building change detection in SAR images using combined differential image and residual u-net network," *Remote Sens.*, vol. 11, no. 9, 2019, Art. no. 1091.

[2] U. Renu, R. Nema, M. Awasthi, and Y. Tiwari, "Change detection analysis of cropland using geospatial technique-a case study of narsinghpur district," *Int. J. Environ., Agriculture Biotechnol.*, vol. 2, no. 4, pp. 1726–1731, 2017.

[3] E. J. Rignot and J. J. Van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896–906, Jul. 1993.

[4] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.

[5] 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.

[6] M. Gong, T. Zhan, P. Zhang, and Q. Miao, "Superpixel-based difference representation learning for change detection in multispectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2658–2673, May 2017.

[7] Q. Wang, X. Zhang, G. Chen, F. Dai, Y. Gong, and K. Zhu, "Change detection based on faster R-CNN for high-resolution remote sensing images," *Remote Sens. Lett.*, vol. 9, no. 10, pp. 923–932, 2018.

[8] G. Chen, G. J. Hay, L. M. Carvalho, and M. A. Wulder, "Object-based change detection," *Int. J. Remote Sens.*, vol. 33, no. 14, pp. 4434–4457, 2012.

[9] W. Zhao, Z. Wang, M. Gong, and J. Liu, "Discriminative feature learning for unsupervised change detection in heterogeneous images based on a coupled neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7066–7080, Dec. 2017.

[10] L. Wan, Y. Xiang, and H. You, "A post-classification comparison method for SAR and optical images change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 7, pp. 1026–1030, Jul. 2019.

[11] V. Iglovikov, S. S. Seferbekov, A. Buslaev, and A. Shvets, "Ternausnetv2: Fully convolutional network for instance segmentation," in *Proc. Conf. Comput. Vis. Pattern Recognit. Workshops*, vol. 233, 2018, Paper 237.

[12] Y. Liu, M.-M. Cheng, X. Hu, K. Wang, and X. Bai, "Richer convolutional features for edge detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 3000–3009.

[13] L. Zhou, C. Zhang, and M. Wu, "D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction." in *Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 182–186.

[14] J. Niemeyer, F. Rottensteiner, and U. Soergel, "Contextual classification of lidar data and building object detection in urban areas," *ISPRS J. Photogrammetry Remote Sens.*, vol. 87, pp. 152–165, 2014.

[15] N. Yokoya *et al.*, "Open data for global multimodal land use classification: Outcome of the 2017 IEEE GRSS data fusion contest," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1363–1377, May. 2018.

[16] X.-Y. Tong *et al.*, "Learning transferable deep models for land-use classification with high-resolution remote sensing images," 2018.

[17] V. Mnih, *Machine Learning for Aerial Image Labeling*. Princeton, NJ, USA: Citeseer, 2013.

[18] A. Van Etten, D. Lindenbaum, and T. M. Bacastow, "Spacenet: A remote sensing dataset and challenge series," 2018, *arXiv:1807.01232*.

[19] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[20] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in *Proc. Eur. Conf. Comput. Vis*. Springer, 2008, pp. 705–718.

[21] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[22] A. Chaurasia and E. Culurciello, "Linknet: Exploiting encoder representations for efficient semantic segmentation," in *Proc. IEEE Visual Commun. Image Process.*, 2017, pp. 1–4.

[23] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 801–818.

[24] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3431–3440.

Nan Zhou received the B.S. degree in remote sensing science and technology from Wuhan University, Wuhan, China, in 2014. He is currently working toward the Ph.D. degree in aerospace information research Institute, Chinese Academy of Sciences, Beijing, China.

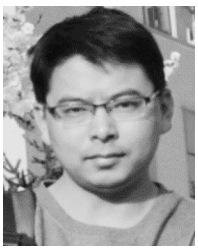
His research interests include deep learning, computer vision, and remote sensing image recognition.

Xiang Li received the B.S. degree in remote sensing science and technology from Wuhan University, Wuhan, China, in 2014. He received the Ph.D. degree in cartography and GIS from the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2019.

He is currently a Postdoctoral Associate with the Department of Electrical and Computer Engineering, New York University, New York, NY, USA and the Department of Electrical and Computer Engineering, New York University Abu Dhabi, Abu Dhabi, UAE. His research interests include deep learning, computer vision, and remote sensing image recognition.

Zhanfeng Shen received the B.S. and M.S. degrees in exploration from Daqing Petroleum Institute, Daqing, China, in 1999 and 2002, respectively, and the Ph.D. degree in cartography and geography information system from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2005

From 2005 to 2007, he was a Postdoctoral Fellow with the Institute of Remote Sensing Applications, Chinese Academy of Sciences. He is currently a Professor with the Aerospace Information Research Institute, CAS, Beijing, China. His research interests include information extraction from medium and high resolution remotely sensed images.



Tianjun Wu received the B.S. degree in information science and the M.S. degree in applied mathematics from Chang'an University, China, in 2009 and 2012, respectively, and the Ph.D. degree in cartography and geographical information system from Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2015.

He is currently an Associate Professor with the School of Science, Chang'an University. His research interests include intelligent information extraction from remote sensing images, spatial-temporal data mining.

Jiancheng Luo received the B.S. degree in remote sensing from Zhejiang University, Hangzhou, China, in 1991, the M.S. degree in GIS from the Chinese Academy of Sciences (CAS), Beijing, China, in 1996, and the Ph.D. degree in GIS from Aerospace Information Research Institute, CAS, in 1999.

He is currently a Professor with the Aerospace Information Research Institute, CAS. His research interest is artificial intelligence techniques in remote sensing.