Landslide Inventory Mapping With Bitemporal Aerial Remote Sensing Images Based on the Dual-Path Fully Convolutional Network

ZhiYong Lv, TongFei Liu, XiangBing Kong, Cheng Shi, and Jón Atli Benediktsson

***Abstract*—This article presents a novel dual-path full convolutional network (DP-FCN) model for constructing a landslide inventory map (LIM) with bitemporal very high-resolution (VHR) remote sensing images. Unlike traditional methods for drawing LIM, the proposed DP-FCN directly draws LIMs from the bitemporal aerial images with VHR through a trained deep neural network without generating the change magnitude map. Thus, the proposed approach can effectively reduce the effects of pseudo changes caused by phenological differences rather than landslide events. The proposed DP-FCN model contains two modules, namely, deep feature extraction, and joint feature learning networks. Deep feature extraction aims to reduce redundancy while extracting the high-level deep features from bitemporal images. Joint feature learning establishes the relationship between the deep features of bitemporal images and the ground reference map. Experiments on the real datasets of the landslide sites in Lantau Island of Hong Kong, China, demonstrate the feasibility and superiority of the proposed approach in drawing LIM with VHR remote sensing images. Moreover, compared with the results obtained by the state-of-the-art algorithms, the proposed DP-FCN method achieves the best performance in terms of accuracy for landslide inventory mapping.**

***Index Terms*—Change detection (CD), landslide inventory map (LIM), natural disaster, remote sensing images.**

I. Introduction

LANDSLIDES cause serious economic losses and many casualties annually. For example, at least two billion US dollars were expended in landslides in the United States each year [1], [2]. Furthermore, many people were killed from landslide occurrences. For example, 73 people died in the landslide event of Shenzhen City, China [3], and approximately 500 fatalities have been recorded in the past six decades in the Hong Kong area of China [4], [5]. To evaluate disaster assessments, carry out rescues, or analyze the influencing factors, obtaining landslide inventory maps (LIMs), which depict the sizes, shapes, locations, spatial distributions, and dates of occurrences of landslides, is essential [6]–[9]. To date, many LIM methods have been proposed through remote sensing techniques, especially those that use very high-resolution (VHR) remote sensing images.

Landslides change the ground surface of the Earth; thus, the progress of mapping landslides can be considered the land cover change detection (LCCD) problem [10]–[12]. Recently, landslides were captured and depicted in enhanced details with the convenience of VHR remote sensing images (satellite, airborne, and UAV platforms) [13], [14]. However, although VHR remote sensing images have high spatial resolution, they are insufficient in spectral domain that results in salt-and-pepper noises, which become a common phenomenon in the detection results when using VHR images [15]–[17]. To smoothen the noise in the detection maps, contextual information is usually adopted to reduce the pseudo change of LIMs. For example, level-set evolution theory has been developed for LCCD with bitemporal aerial photographs for LIM. Specifically, two strategies called edge-based level-set evolution (ELSE) and region-based level-set evolution (RLSE) were proposed on the basis of the theory in [18]. Zhang *et al.* [19] extended the level-set method with local uncertainty constraints (LSELUC) for landslide mapping. Li *et al.* [20] proposed the Markov random field-based LCCD method to draw LIMs. Although landslides can be detected through LCCD techniques in theory, these LIM drawing methods still have following limitations. 1) The differences in bitemporal images in terms of atmospheric conditions, sun angle, and phenological variation usually cause pseudo changes in the detection map. Thus, the process of maintaining real changes while removing the pseudo ones remains a challenge in LCCD [21]–[24]. 2) Binary threshold is required to divide the change magnitude image (CMI) into binary LCCD maps; however, determination of the optimal binary threshold for a given CMI is usually time consuming and subjective [25], [26]. Therefore, developing an automatic LIM method without constructing CMIs and binary thresholds is important and promising.

In recent years, the development of deep learning approaches allows for effective LCCD and rapid landslide mapping. The convolutional neural network (CNN) is a classical deep

Manuscript received January 15, 2020; revised February 27, 2020; accepted March 11, 2020. Date of publication March 16, 2020; date of current version August 24, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFC0504501, in part by the National Natural Science Foundation of China under Grants 61501200, 61701396, and 61902313, and in part by the Natural Science Foundation of Shaan Xi Province under Grant 2017JQ4006. *(Corresponding author: XiangBing Kong.)*

ZhiYong Lv, TongFei Liu, and Cheng Shi are with the School of Computer Science and Engineering, Xi'An University of Technology, Xi'An 710048, China (e-mail: lvzhiyong_fly@hotmail.com; liutongfei_home@hotmail.com; chengc_s@163.com).

XiangBing Kong is with the Yellow River Institute of Hydraulic Research, Zhengzhou, Henan 450003, China (e-mail: kongxb@foxmail.com).

Jón Atli Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, IS-107 Reykjavik, Iceland (e-mail: benedikt@hi.is).

Digital Object Identifier 10.1109/JSTARS.2020.2980895

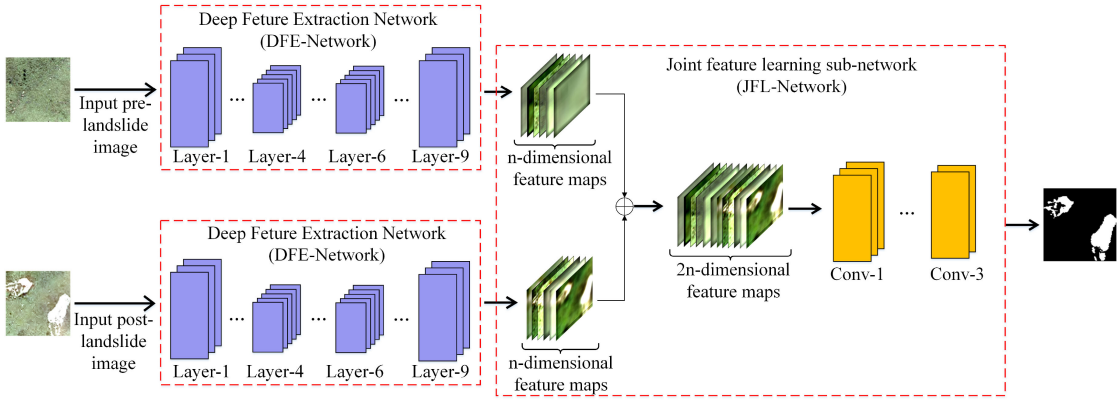


Fig. 1. Framework of the proposed DP-FCN approach.

learning model [27], and has been widely used for image recognition [28]–[30], scene classification [31], and object detection [32]. Many methods have been proposed for LCCD by using synthetic aperture radar (SAR) images with deep neural networks (DNNs) [33], [34] and local restricted CNN (LRCNN) [35]–[37]. In these methods, abstract and deep features are learned for generating differing images (DIs) to measure the change magnitude among the bitemporal SARs. To avoid the effect of uncertain DI noise on the detection results, a joint classifier of the bitemporal images was introduced into a DNN based on the restricted Boltzmann machine to omit the progress of generating DIs [38]. In the field of LCCD with optical VHR images, certain networks have been investigated to concentrate on deep feature exploration and suitable feature space transformation. For instance, the authors in [39] exploited the spatiotemporal features of a series of monitoring satellite images for LCCD using deep CNN. The authors in [40] designed a general end-to-end two-dimensional (2-D) CNN framework used for learning discriminative features and detecting land cover change from hyperspectral images. The authors in [41] proposed a fully atrous CNN for LCCD through the learning scale representation feature with VHR remote sensing images. Recently, a full convolutional network within pyramid pooling (FCN-PP) was also proposed for LIM by using bitemporal aerial images [42]. An in-depth literature survey about LCCD with optical VHR remote sensing images can be found in [43] and [44]. The advantage of the machine learning-based LCCD methods lies in detecting the "change" and "unchanged" area directly without generation of CMI. Compared with traditional methods, avoiding the generation of CMI can circumvent the determination of binary threshold which is usually a time-consuming and experience-dependent progress [45].

As mentioned previously, although many methods based on deep learning techniques have been developed for detecting land cover changes with VHR remote sensing images, filters or other pre-operations in these deep learning techniques are usually required for processing raw bitemporal images for these commonly used networks. Furthermore, CMIs are first generated on the basis of the explored high-level features, and changes are also indirectly detected from the CMIs. Moreover, from the viewpoint of application, the majority of existing methods have focused on LCCD, and only limited methods with deep learning techniques have been designed specifically for LIM with VHR remote sensing images. Therefore, obtaining LIM with VHR remote sensing images remains a challenge due to the general great differences between the pre- and postevent images of landslides in spectra, moisture, and phenomenon.

In this article, we design a dual-path full convolutional network (DP-FCN) model for rapid LIM with VHR remote sensing images to address the abovementioned problems. The motivation of the proposed DP-FCN lies in three aspects. 1) Reducing the pseudo changes caused by the differences in the bitemporal images due to the different seasons, sun height, and soil moisture. 2) Investigating the feasibility and performance of deep learning techniques for the task of LIM with VHR remote sensing images, which have a remarkable difference in spectral reflectance. 3) Learning robust and abstract features from the bitemporal images for LIM without needing to generate CMIs and other preprocessing operations.

The rest of this article is organized as follows. Section II presents the details of the proposed approach. Section III discusses the experiments on certain landslide site images to verify the feasibility and performance of the proposed approach. Finally, Section IV concludes the article.

II. Proposed Dual-Path Full Convolutional Network

In this section, we present the proposed DP-FCN model for drawing LIM with VHR remote sensing images. The main motivation of the proposed DP-FCN model is to reduce pseudo changes and improve the detection accuracy of LIM by omitting the process of CMI generation. The proposed DP-FCN contains two cooperative modules, namely, the deep feature extraction (DFE-network) and joint feature learning (JFL-network) networks. DFE-network is designed to exploit the robust features of bitemporal images without any preprocessing, whereas JFL-network is used for learning the relationship between the bitemporal deep features and the ground reference map. The framework of the proposed model is shown in Fig. 1. The details of the proposed DP-FCN model will be presented in the following sections.

TABLE I
DETAILED ARCHITECTURE OF THE PROPOSED DP-FCN

	Layer	NO. of convs	NO. of channels	Conv. kernel size	Conv. stride	Padding	Deconv/ Pool size	Deconv/ Pool stride	Output size
DFE-network	Layer-1	2	64	3×3	1	SAME	2×2	2	64×64
	Layer-2	2	128	3×3	1	SAME	2×2	2	32×32
	Layer-3	2	256	3×3	1	SAME	2×2	2	16×16
	Layer-4	2	512	3×3	1	SAME	2×2	2	8×8
	Layer-5	2	1024	3×3	1	SAME	2×2	2	16×16
	Layer-6	2	512	3×3	1	SAME	2×2	2	32×32
	Layer-7	2	256	3×3	1	SAME	2×2	2	64×64
	Layer-8	2	128	3×3	1	SAME	2×2	2	128×128
	Layer-9	2	64	3×3	1	SAME	2×2	2	128×128
JFL-network	Conv-1	1	64	3×3	1	SAME	–	–	128×128
	Conv-2	1	64	3×3	1	SAME	–	–	128×128
	Conv-3	1	2	1×1	1	SAME	–	–	128×128

A. Deep Feature Extraction Network (DFE-Network)

The DFE-network is designed to extract the high-level features of bitemporal images and constructed on the basis of the U-net, which is an efficient deep feature extraction model [46] that has been proven to be effective in many tasks [30], [47], [48]. To the best of our knowledge, the U-net is first introduced for LIM with VHR remote sensing images in this article.

Fig. 1 shows the structure of the DFE-network with nine layers. This network contains the encoder and decoder parts. The encoder process reduces the spatial dimension with four convolution layers and four downsampling layers. This process also restores the detailed information and the spatial dimension with four convolution and four upsampling layers. In addition, the corresponding feature maps of encoders and decoders are concatenated to further preserve the details of landslides. The output feature map of each remote sensing image has *n* channels (*n* is a per-set parameter). Thus, it can be considered as a 3-D feature cube. The output features of the DFE-network have the same spatial size as the inputs.

In the proposed DFE-network, the activation function is required to learn the nonlinear relationship between every network layer. Rectified linear unit (ReLU), which has been successfully used in [49], is adopted as the activation function for each convolutional layer. ReLU can be defined as follows:

$$f(x) = \max(0, x) \tag{1}$$

where x is the output feature from each convolution layer. In addition, to overcome the overfitting issue and achieve the optimal local minimum, the dropout is exploited after each layer.

Notably, the proposed DFE-network is an extension application of the U-net. Thus, it inherits the characteristics of the U-net. For the task of LIM, the DFE-network learns the high-level features from the bitemporal VHR remote sensing image. Concurrently, the low-level features of the bitemporal images can be captured. The preservation of high- and low-level features is conducive to accurately mine the relationship between the landslide and nonlandslide areas.

B. Joint Feature Learning Network (JFL-Network)

Many studies have achieved the LCCD map by generating the CMIs. However, the bitemporal images of LCCD are usually different in terms of season, sun height, and even soil moisture. Moreover, the bitemporal images used for landslide mapping are typically located in mountain areas covered with various types of vegetation. These issues will result in pseudo change metrics in the CMIs and cause pseudo changes in the final detection map. To address these problems, the JFL-network in the proposed DP-FCN approach is suggested to directly detect the landslide areas from the bitemporal deep features without the generation of CMIs. Details of the proposed JFL-network are given in the following paragraphs.

The structure of the JFL-network is shown in Fig. 1. We first concentrate on the two groups of 3-D cube feature maps obtained by using the DEF-network into a new 3-D feature cube. The concentrated 3-D feature cube is then considered the input of the JFL-network. The JFL-network contains two full convolution layers and a dense convolution layer to further learn the relationship between the features of the pre- and postevent images. At the end of the JFL-network, a softmax classifier is applied to each pixel to obtain the final detection results. The possibility presents a pixel that can either be "landslide" or "nonlandslide" and is defined as follows:

$$P_k^n = \frac{e^{w_k^n}}{\sum_{j=0}^{j=1} e^{w_j^n}} \, (k = k|0, 1) \tag{2}$$

where P_k^n indicates the probability that a pixel belongs to the *k*th class, and w_k^n denotes the weight matrix of the F^k feature map. In this article, the proposed DP-FCN aims to obtain LIM by dividing an image scene into landslide and nonlandslide areas, where "$k = 1$" and "$k = 0$" represent landslide and nonlandslide areas, respectively. As mentioned above, the landslide (w_1^n) and nonlandslide (w_0^n) weight matrixes are generated by using the dense convolution layer in the JFL-network. The label (L) of each pixel x^n can be predicted on the basis of the following maximal probability P_k^n:

$$L(x^n) = \text{argmax} P_k^n, (k = k|0, 1) \tag{3}$$

From the discussion above, the landslide areas are detected by determining the relationship between the ground reference and the bitemporal 3-D feature maps in the proposed JFL-network. Unlike the traditional methods, the pseudo change caused by the generation of CMI can be reduced because the calculation of CMI and the binary threshold is avoided in the proposed DP-FCN.

TABLE II
DESCRIPTION OF THE QUANTITATIVE CRITERIA FOR THE EXPERIMENTS

Criterion	Formula	Description
Completeness (CP)	$\frac{TP}{TP+FN}$	Obtains the ratio of the achieved real changed pixels compared with the real changed pixel in the ground reference map.
Correctness (CR)	$\frac{TP}{TP+FP}$	Demonstrates the correlation between the detected real changed pixels and the detected changed pixels in the detected map.
Quality (QA)	$\frac{TP}{TP+FP+FN}$	Presents the percentage of the real detected changed pixel compared with the changed pixels in the reference map and the acquired LIM.
Overall accuracy (OA)	$\frac{TP+TN}{TP+TN+FP+FN}$	Determines which proportions of the ground reference map were detected correctly.
Average accuracy (AA)	$\frac{1}{2} \times (\frac{TP}{TP+FP} + \frac{TN}{TN+FN})$	Presents the percentage of the real detected changed pixel compared with the changed pixels in the reference map and the acquired LIM.

C. Network Training for the Proposed DP-FCN

Two co-registered images, namely, I_1 and I_2, are assigned as the pre- and postevent images for LIM. Notably, the proposed DP-FCN only accepts inputting image blocks with a size of 128×128 pixels with the three spectral channels. Thus, the raw image scene should be separated into regular blocks with the size of 128×128 pixels. In the proposed DP-FCN approach, the main goal of two parallel DFE-networks lies in extracting two *n*-dimension feature map vectors from the inputs I_1 and I_2 which can be described as $F_1 = \{f_1^1, f_1^2, \ldots, f_1^n\}$ and $F_2 = \{f_2^1, f_2^2, \ldots, f_2^n\}$, respectively, where *n* is set to 64 in the proposed DP-FCN. Subsequently, F_1 and F_2 are concatenated as $F = \{f_1^1, f_1^2 \ldots f_1^n, f_2^1, f_2^2, \ldots, f_2^n\}$ and *F* is placed into the convolutional layer of the JFL-network for joint learning. Then, the probability of each pixel is calculated by using the softmax layer for determining whether each pixel belongs to the landslide or nonlandslide class. Finally, LIM is predicted by using the proposed DP-FCN approach. The entire algorithm is depicted in Algorithm I.

In the progress of training the DP-FCN model, the weight matrix of the convolutional layers are initialized by using the Xavier method [50]. Furthermore, the detailed configuration of DP-FCN is shown in Table I. To fine-tune the proposed DP-FCN, we adopt the classical cross-entropy method as the loss function for training the model. The cross-entropy function can be defined as follows:

$$\text{Loss}_i = -y_i \log(y_i^{:}) \tag{4}$$

where y_i represents the label of the ground reference image, and $y_i^{:}$ denotes the predicted label. Hence, loss_i is the loss value of the *i*th training block. We minimize the loss simultaneously by using the AdamOptimizer with backpropagation, and the learning rate for training the DP-FCN model is set to 10^{-4}. To avoid the problem of overfitting, we set the parameter keep_prob of the dropout to 0.6. On the basis of the above parameter setting, the proposed DP-FCN can be trained and fine-tuned. Finally, when the bitemporal landslide images are fed into the proposed DP-FCN, the corresponding LIM can be automatically acquired.

Algorithm 1: The Framework of the Proposed DP-FCN Approach.

Input: The bitemporal images (I_1 and I_2) are divided into *m* image blocks with the size of 128×128 pixels. I_1^i and I_2^i are a pair of blocks for input, where $i \in [1, m]$, and *m* is the number of training blocks.

1: **for** $i = 1 : m$ **do**
2: Extract the feature map vectors $f_1^i = \{f_1^1, f_1^2, \ldots f_1^n\}$ and $f_2^i = \{f_2^1, f_2^2, \ldots f_2^n\}$ from the pairwise image blocks I_1^i and I_2^i by using parallel DFE-networks, respectively.
3: Concatenate the feature map vectors f_1^i and f_2^i into $f^i = \{f_1^1, f_1^2 \ldots f_1^n, f_2^1, f_2^2, \ldots, f_2^n\}$ in the JFL-network.
4: Learn the contacted relationship of the feature vector f^i through the convolution layer in JFL-network.
5: Calculate the probability of each pixel by using formula (2).
6: The label of each pixel is determined by using equation (3).
7: AdamOptimizer with backpropagation is employed for fine-tuning the proposed DP-FCN.
8: $i \leftarrow i + 1$;
9: **end for**

Output: the trained DP-FCN model.

We implement the proposed DP-FCN model on a workstation with Intel CPU W-2123, 3.6 GHz, four cores, 64 GB RAM, and NVIDIA GTX 2080Ti GPU. We have deployed the entire project in PyCharm-2019, which configures Python 3.6 and TensorFlow 1.12.0.

III. Experiments

In this section, the experiments based on real landslide sites in Hong Kong are investigated to validate the effectiveness and performance of the proposed network for LIM with VHR remote sensing images. First, the general information of the experiments, including the quantitative evaluating criteria and parameter setting of the comparison methods, are detailed. Second, the bitemporal images for investigating landslide sites are described. Finally, the visual and quantitative performance of the LIMs are shown for comparison. Each part will be detailed as follows.

A. General Information

First, to quantitatively evaluate the proposed DP-FCN and demonstrate its superiority for acquiring LIM with VHR remote sensing images, five widely used criteria are adopted for the experimental comparison. The details of these adopted criteria are listed in Table II.

Table II presents that the true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP) that were calculated in all cases. Specifically, TP and TN are the total number of correctly detected pixels for "changed" and "unchanged" pixels, respectively; FN is the number of changes that have not been detected by the method; and FP is the number of false alarms.

Second, to demonstrate the superiority of the proposed DP-FCN, it was compared with three traditional but relatively new landslide detection approaches (e.g., ELSE, RLSE, and LSELUC [19]) and three classical deep learning methods (e.g., FCN, U-net, and FCN-PP [42]). The parameters of the three traditional methods used in the following experiments are given as follows: 1) $\alpha = 1.5, C_0 = 1.0, \sigma_1 = \sigma_2 = 1.0$ are set for ELSE, the parameters of RLSE are $\alpha = 1.0$ the template size of the Gaussian filter is fixed at 9×9, and time step $\Delta t = 8$. 2) The parameters of the three classical deep learning methods are presented as follows: the learning rate and keep_prob for the FCN are set to 10^{-4} and 0.5, respectively; 10^{-4} and 0.6 are set for the learning rate and keep_prob of the U-net; and the filter size in the FCN-PP is 1×1; concurrently, the learning rate and keep_prob are fixed at 10^{-4} and 0.6, respectively, in the FCN-PP network.

B. Data Description

In this article, 13 landslide sites in Lantau Island of Hong Kong, China are investigated to test the effectiveness of the proposed network. The pre- and postevent images were acquired by using a Zeiss RMK TOP 15 Aerial Survey Camera at a flying height of approximately 2400 m on December 2007 and November 2014, respectively. The spatial resolution of the bitemporal aerial photos is 0.5 m/pixel. Then, the preprocessing, such as coregistration and radiometric corrections, was conducted by the Civil Engineering and Development Department of Hong Kong. In this article, the image pairs from "A" to "E" are assigned as the dataset for training the deep learning model, and the image pairs from "F" to "M" are utilized for testing the effectiveness and performance of each selected approach.

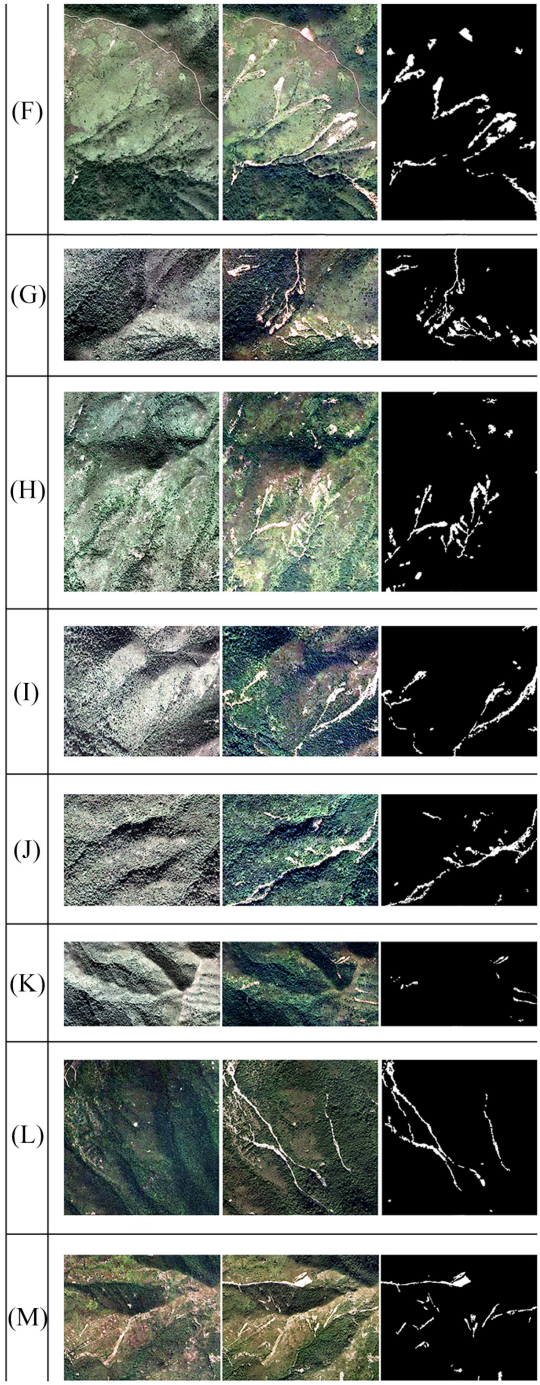


Fig. 2. Test dataset for landslide sites and their reference maps. The column from left to right are the dataset-ID, pre-event images, post-event images, and ground reference maps.

The details of each bitemporal image are shown in Fig. 2, and the visual performance of the testing data is presented in Fig. 2. Notably, the original blocks for training the learning model are extended through "rotation" to obtain a robust model. In addition, the following aspects can be found by observing the visual performance of these bitemporal images of the landslide sites.

1) The spectral differences of the bitemporal images caused by the difference of seasonal phenology are serious, as

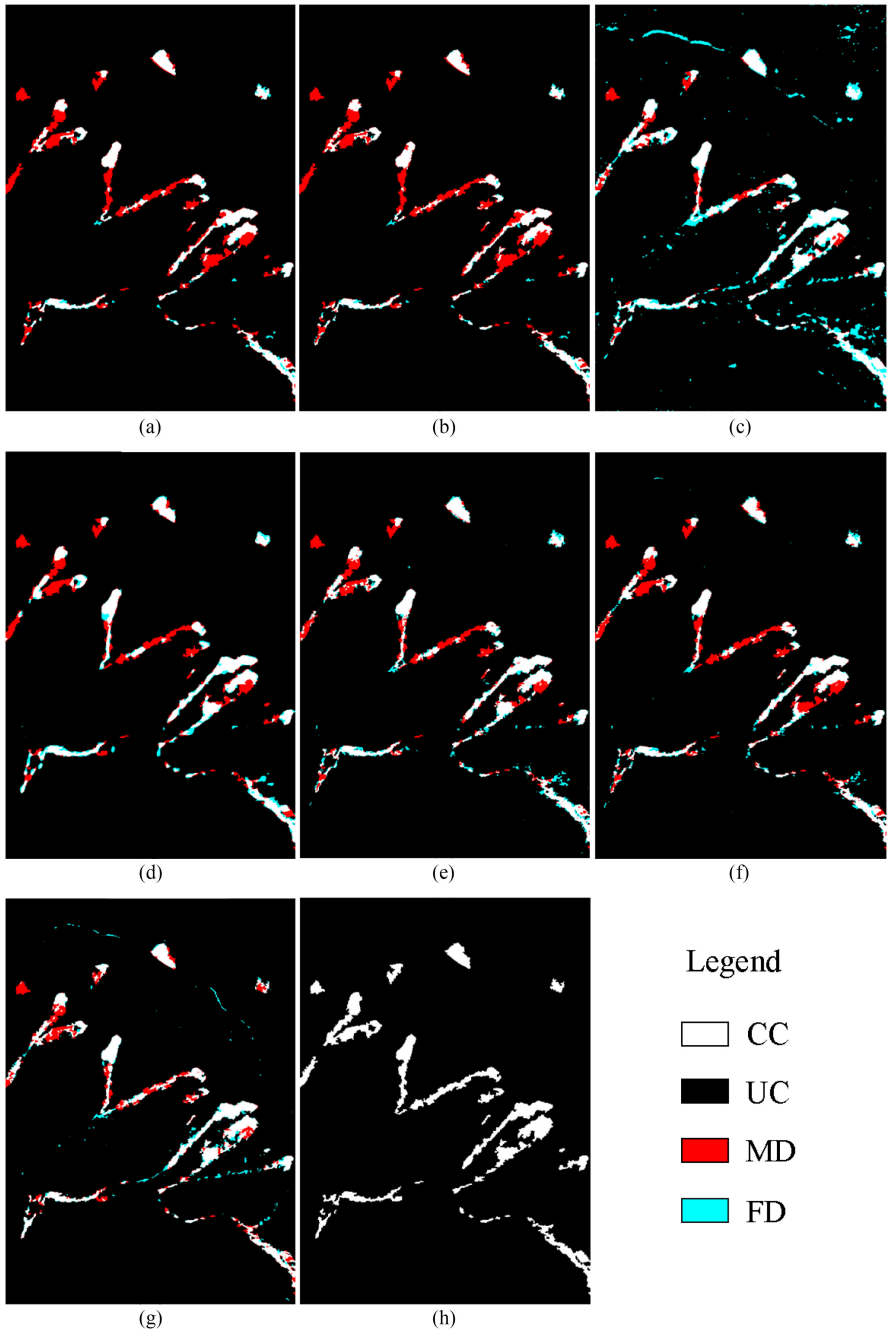


Fig. 3. Landslide inventory mapping obtained for data-F. (a) ELSE. (b) RLSE. (c) LSELUC. (d) FCN. (e) U-net. (f) FCN-PP. (g) Proposed DP-FCN. (h) Ground reference map.

shown in the pre- and postevent images in Fig. 2(I) and (J).

2) Given that certain areas are covered by tall trees, the landslides in these areas are heterogenous, as shown in Fig. 2(F).
3) Bare rocks are easily confused with landslides because the incident locations of landslides are usually located in mountainous areas. From these observations, acquiring accurate LIM by using VHR remote sensing images is still challenging. Details of the testing datasets are presented in Table III.

TABLE III
DESCRIPTION OF THE DATASETS IN THE EXPERIMENTS

Data		Size (pixels)	Block (128 × 128 pixels/block)	Ground Reference Map (change pixels/unchanged pixels)
Training data	A	754 × 694	910	26,572/496,704
	B	1252 × 2199	4451	113,234/26,399,147
	C	923 × 593	1040	26,593/520,746
	D	1107 × 743	648	17,261/805,240
	E	826 × 725	271	7,105/591,745
Testing data	F	640 × 896	35	30,607/542,833
	G	896 × 640	35	29,774/543,666
	H	896 × 1152	63	37,696/994,496
	I	768 × 640	30	20,537/470,983
	J	896 × 640	35	27,006/546,434
	K	1664 × 896	91	138,21/147,712,3
	L	1024 × 1024	64	31,898/1,016,678
	M	1280 × 1024	80	38,348/1,272,372

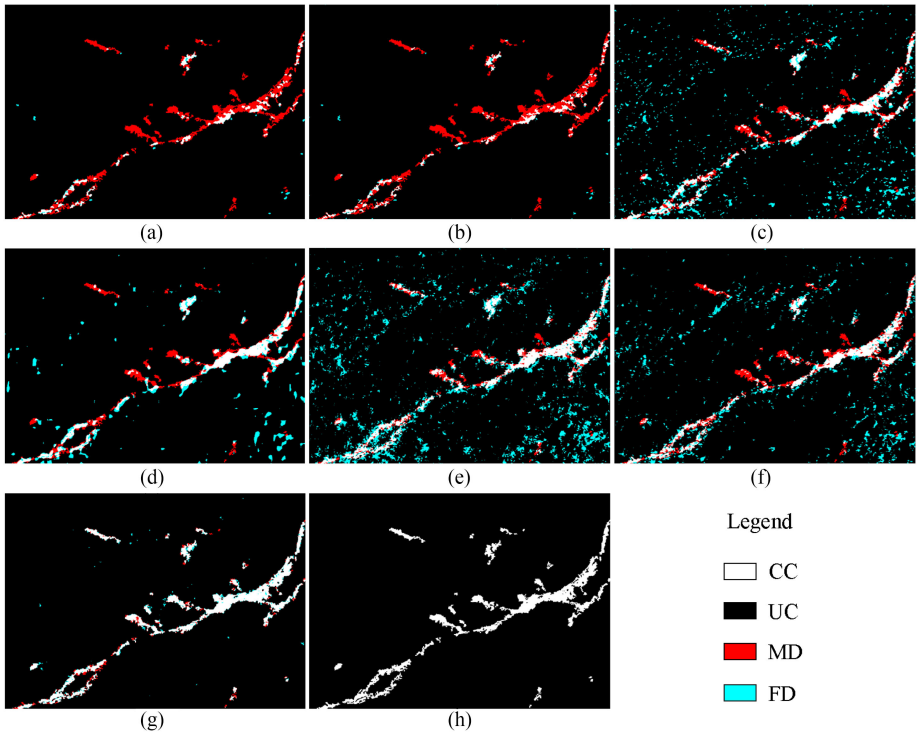


Fig. 4. Landslide inventory mapping obtained for data-J. (a) ELSE. (b) RLSE. (c) LSELUC. (d) FCN. (e) U-net. (f) FCN-PP. (g) Proposed DP-FCN. (h) Ground reference map.

C. Experimental Results and Analysis

For the comparison of visual performance of LIMs, the correctly changed (CC), unchanged (UC), missing detected (MD), and false detected (FD) pixels are assigned as white, black, red, and aqua colors in the figures, respectively. To demonstrate the visual performance of different methods, Figs. 3 and 4 were selected to show the LIM visual performance of seven different methods on the landslide sites F and J in Lantau Island, Hong Kong, China. These figures demonstrate that FD and MD pixels emerged in all the achieved LIMs because despite these aerial images with VHR perform the better visual performance, they are insufficient in spectral bands. This phenomenon results in large intraclass variance and brings considerable noises in a group of pixels that consist a ground target. Compared with the LIMs from the traditional methods and the classical deep learning approaches, the proposed DP-FCN achieves the best performance with the fewest FD pixels. Furthermore, although the spectral reflection of the land cover types in the bitemporal images is very different [Fig. 2(I) and (K)], the traditional methods ELSE and RLSE missed detecting most of the parts of landslide areas, and the LIMs acquired by other methods emerged with many FD pixels. Compared with these visual observations and quantitative comparisons in Table IV, the proposed DP-FCN achieves the best LIMs for most sites of the landslide sites. The best LIMs for the landslide sites I and K can be obtained by the proposed DP-FCN, the reason is that CMI is omitted in the proposed network, and the error in the corresponding pixels of the bitemporal images can be avoided, and the error between the corresponding pixels of the bitemporal images can be avoided. Thus, the proposed DP-FCN can directly learn the deep feature and relationship between the bitemporal images and the ground reference map. Therefore, the proposed DP-FCN can avoid the antipseudo change brought about by differences in seasonal phenology spectra.

The quantitative comparative bars for each dataset are shown in Table IV. From the viewpoint of quantitative evaluation, the proposed DP-FCN achieves the best accuracies in terms of five evaluation measurements (CP, CR, QA, OA, and AA) which are defined in Table II. Hence, the quantitative comparisons have clearly demonstrated and supported the superiority of the proposed DP-FCN method.

D. Discussion

To promote the potential application of the proposed DP-FCN approach, we discussed the advantages and the limitation of the proposed approach as follows.

1) Advanced landslide area detection maps are obtained with the proposed DP-FCN. As shown in the Table IV, it clearly demonstrated that the proposed DP-FCN outperforms the widely used traditional LCCD methods and the basic deep learning approaches, in terms of visual performance and five quantitative evaluation criteria.
2) The simple and no-parameter approach will be more acceptable and easier to promote in a piratical application. Despite the proposed DP-FCN requires training samples for learning parameters and optimizing the model, a landslide inventory map can be achieved by the proposed approach without any parameters.

TABLE IV
QUANTITATIVE COMPARISONS FOR DIFFERENT METHODS ON THE DIFFERENT DATASETS

Datasets	Methods	CP	CR	QA	OA	AA
F, G	ELSE	54.15, 30.95	89.82, 89.81	51.02, 29.92	97.23, 96.23	76.90, 65.38
	RLSE	51.14, 28.91	**90.25**, **91.97**	48.46, 28.20	97.10, 96.18	75.41, 64.39
	LSELUC	82.13, 65.95	65.31, 48.12	57.19, 38.55	96.72, 94.54	89.84, 81.03
	FCN	61.98, 45.73	79.55, 69.09	53.46, 37.96	97.12, 96.12	80.54, 72.30
	U-net	66.13, 62.50	81.47, 49.85	57.49, 38.38	97.39, 94.79	82.64, 79.53
	FCN-PP	60.25, 48.35	84.36, 62.08	54.20, 37.33	97.28, 95.78	79.81, 73.37
	Proposed DP-FCN	**73.54**, **84.94**	85.62, 90.94	**65.46**, **78.31**	**97.93**, **98.78**	**86.42**, **92.24**
H, I	ELSE	19.40, 13.17	70.36, 67.31	17.96, 12.38	96.76, 96.10	59.55, 56.44
	RLSE	17.89, 13.70	72.83, 64.88	96.76, 12.76	96.76, 96.08	58.82, 56.69
	LSELUC	57.06, 56.35	31.03, 12.27	25.15, 11.20	93.80, 81.34	76.13, 69.39
	FCN	29.08, 58.62	52.69, 10.30	23.06, 9.59	96.46, 76.93	64.05, 68.17
	U-net	43.21, 70.68	39.88, 9.61	26.17, 9.24	95.55, 70.99	70.37, 70.84
	FCN-PP	32.73, 55.87	53.59, 12.96	25.50, 11.75	96.51, 82.47	65.83, 69.75
	Proposed DP-FCN	**89.20**, **94.88**	**73.56**, **78.42**	**67.54**, **75.24**	**98.43**, **98.70**	**93.99**, **96.87**
J, K	ELSE	23.99, 1.34	87.15, 7.19	23.17, 1.14	96.25, 98.93	61.91, 50.59
	RLSE	23.71, 1.72	87.34, 7.29	22.92, 1.41	96.25, 98.89	61.77, 50.76
	LSELUC	61.64, 85.65	47.63, 1.06	36.74, 1.06	95.00, 25.59	79.15, 55.33
	FCN	56.41, 8.23	57.79, 0.417	39.95, 0.399	96.01, 80.95	77.19, 44.93
	U-net	71.92, 16.56	37.29, 0.459	32.55, 0.449	92.98, 65.96	82.97, 41.49
	FCN-PP	58.56, 8.35	46.85, 0.355	35.19, 0.342	94.92, 77.43	77.64, 43.21
	Proposed DP-FCN	**88.20, 70.57**	**89.63, 93.12**	**80.03, 67.07**	**98.96, 99.68**	**93.85, 85.26**
L, M	ELSE	23.99, 75.28	87.15, 83.58	23.17, 65.58	96.25, 98.84	61.91, 87.42
	RLSE	23.71, 72.93	87.34, 84.52	22.92, 64.34	96.25, 98.82	61.77, 86.27
	LSELUC	61.64, **90.23**	47.63, 46.03	36.74, 43.84	95.00, 96.62	79.15, 93.52
	FCN	56.41, 82.08	57.79, 59.36	39.95, 52.55	96.01, 97.83	77.19, 90.19
	U-net	71.92, 84.65	37.29, 60.01	32.55, 54.12	92.98, 97.90	82.97, 91.47
	FCN-PP	58.56, 82.37	46.85, 69.50	35.19, 60.50	94.92, 98.43	77.64, 90.64
	Proposed DP-FCN	**88.20**, 83.67	**89.63**, **89.26**	**80.03**, **76.01**	**98.96**, **99.23**	**93.85**, **91.68**

The bold entity indicate the best accuracy in terms of the corresponding measurement and datasets.

3) One limitation of the proposed DP-FCN lies in that it requires training samples for training the network. However, labeling training samples is time-consuming and labor-intensive.

IV. CONCLUSION

This article presents a novel deep learning approach called DP-FCN for landslide inventory mapping through VHR remote sensing images. This method consists of the following main modules: 1) deep feature extraction network with the aims of learning abstract and robust features from the pre- and postevent images, respectively, and 2) joint feature learning network used for concatenating the bitemporal deep features and learning the relationship between these deep features and the landslide reference map. Then, the two modules are constructed as the DP-FCN, and the entire network is trained and fine-tuned by training the datasets with the classical cross-entropy loss function and a specific learning ratio. The experimental results on 13 landslide sites of VHR aerial photos and their comparison with RLSE, ELSE, LSELUC, FCN, U-net, and FCN_PP shows that the proposed DP-FCN is effective and superior for drawing LIM with VHR aerial images. To the best of our knowledge, this article is the first attempt at developing a deep learning network for LIM through VHR remote sensing images without measuring the change magnitude. The major advantages and contributions of the proposed DP-FCN are briefly summarized as follows:

1) Better performance and higher accuracies of LIMs are obtained by the proposed DP-FCN. For eight of thirteen landslide sites with VHR aerial photos, the LIMs clearly demonstrate that the proposed DP-FCN outperforms six widely used LCCD methods, namely, RLSE, ELSE, LSELUC, FCN, U-net, and FCN_PP, in terms of visual performance and five quantitative evaluation criteria.

2) No parameters are needed for tuning in the proposed DP-FCN. Moreover, the parameters of the learning ratio and keep_prob in the proposed DP-FCN must only be initialized in the training stage. The parameters need not be tuned when they are implemented on real applications. The experimental results also show that the proposed DP-FCN can obtain satisfactory LIM results without hard turning the parameters.

3) The removal of the calculation of CMIs in the proposed DP-FCN successfully reduces the pseudo change and improves the performance of the detected result. The spectral differences in the bitemporal images are general and inevitable, especially for the VHR remote sensing images of landslides that usually occur in mountainous areas. Furthermore, the spectral differences of a ground target will cause the pseudomagnitude in the CMIs, and these pseudomagnitude errors may be transformed into the subsequent steps. Thus, the proposed DP-FCN obtains the LIM by directly learning the relationship between deep features of the bitemporal images and the ground reference map without calculating the CMI, thereby reducing the pseudo change in the results.

Although the advantages of the proposed DP-FCN have been investigated well, the widespread application of this proposed model should be further verified. In our future research, when other sourcing sensors' bitemporal images for a landslide event are available, the applicability of the proposed DP-FCN will be further tested. Even the bitemporal images with different spatial resolutions for a landslide event will also be considered in the further investigation of the proposed DP-FCN method.

ACKNOWLEDGMENT

The authors would like to thank the Editor-in-Chief, Associate Editor, and Reviewers for their insightful comments and suggestions.

REFERENCES

[1] M. K. Iwamoto and P. S. Ooi, "Calibration of a transient hydrological model useful for developing a landslide warning system," in *Proc. Transp. Res. Board 97th Annu. Meeting*, 2018, pp. 3918–3919.

[2] H. Chapella, W. Haneberg, M. Crawford, and A. Shakoor, "Landslide inventory and susceptibility models, Prestonsburg 7.5-min quadrangle, Kentucky, USA," in *Proc. IAEG/AEG Annu. Meeting*, vol. 1, 2019, pp. 217–226.

[3] C. Wang *et al.*, "Formation of the 2015 Shenzhen landslide as observed by SAR shape-from-shading," *Scientific Rep.*, vol. 7, 2017, Art. no. 43351.

[4] K. Choi and R. W. Cheung, "Landslide disaster prevention and mitigation through works in Hong Kong," *J. Rock Mech. Geotechnical Eng.*, vol. 5, no. 5, pp. 354–365, 2013.

[5] C. Wills and T. McCrink, "Comparing landslide inventories: The map depends on the method," *Environmental Eng. Geosci.*, vol. 8, no. 4, pp. 279–293, 2002.

[6] A. Stumpf, J.-P. Malet, and C. Delacourt, "Correlation of satellite image time-series for the detection and monitoring of slow-moving landslides," *Remote Sens. Environ.*, vol. 189, pp. 40–55, 2017.

[7] A. Mondini, F. Guzzetti, P. Reichenbach, M. Rossi, M. Cardinali, and F. Ardizzone, "Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1743–1757, 2011.

[8] H. Shahabi and M. Hashim, "Landslide susceptibility mapping using gis-based statistical models and remote sensing data in tropical environment," *Scientific Rep.*, vol. 5, 2015, Art. no. 9899.

[9] C. Zhao and Z. Lu, "Remote sensing of landslides—A review," *Remote Sens.*, vol. 10, 2018, Art. no. 279.

[10] A. Si *et al.*, "Regional landslide identification based on susceptibility analysis and change detection," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 10, 2018, Art. no. 394.

[11] Z. Y. Lv, W. Shi, X. Zhang, and J. A. Benediktsson, "Landslide inventory mapping from bitemporal high-resolution remote sensing images using change detection and multiscale segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1520–1532, May 2018.

[12] W. Zhao, A. Li, X. Nan, Z. Zhang, and G. Lei, "Postearthquake landslides mapping from Landsat-8 data for the 2015 nepal earthquake using a pixel-based change detection method," *IEEE J. Sel. Topics. Appl. Earth. Observations. Remote. Sens.*, vol. 10, no. 5, pp. 1758–1768, May 2017.

[13] U. Niethammer, S. Rothmund, U. Schwaderer, J. Zeman, and M. Joswig, "Open source image-processing tools for low-cost UAV-based landslide investigations," *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 38, no. 1, pp. 161–166, 2011.

[14] H.-R. Ma, X. Cheng, L. Chen, H. Zhang, and H. Xiong, "Automatic identification of shallow landslides based on Worldview2 remote sensing images," *J. Appl. Remote Sens.*, vol. 10, no. 1, 2016, Art. no. 016008.

[15] M. Herold, M. E. Gardner, and D. A. Roberts, "Spectral resolution requirements for mapping urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1907–1919, Sep. 2003.

[16] L. Zhang, X. Huang, B. Huang, and P. Li, "A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2950–2961, Oct. 2006.

[17] X. Huang, Q. Lu, and L. Zhang, "A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas," *ISPRS J. Photogrammetry Remote Sens.*, vol. 90, pp. 36–48, 2014.

[18] Z. Li, W. Shi, S. W. Myint, P. Lu, and Q. Wang, "Semi-automated landslide inventory mapping from bitemporal aerial photographs using change detection and level set method," *Remote Sens. Environ.*, vol. 175, pp. 215–230, 2016.

[19] X. Zhang, W. Shi, P. Liang, and M. Hao, "Level set evolution with local uncertainty constraints for unsupervised change detection," *Remote Sens. Lett.*, vol. 8, no. 8, pp. 811–820, 2017.

[20] Z. Li, W. Shi, P. Lu, L. Yan, Q. Wang, and Z. Miao, "Landslide mapping from aerial photographs using change detection-based Markov random field," *Remote Sens. Environ.*, vol. 187, pp. 76–90, 2016.

[21] J. Chen, M. Lu, X. Chen, J. Chen, and L. Chen, "A spectral gradient difference based approach for land cover change detection," *ISPRS J. Photogrammetry Remote Sens.*, vol. 85, pp. 1–12, 2013.

[22] Y. Du, P. M. Teillet, and J. Cihlar, "Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection," *Remote Sens. Environ.*, vol. 82, no. 1, pp. 123–134, 2002.

[23] M. Lu, J. Chen, H. Tang, Y. Rao, P. Yang, and W. Wu, "Land cover change detection by integrating object-based data blending model of Landsat and MODIS," *Remote Sens. Environ.*, vol. 184, pp. 374–386, 2016.

[24] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[25] D. Lu, P. Mausel, M. Batistella, and E. Moran, "Land-cover binary change detection methods for use in the moist tropical region of the amazon: a comparative study," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 101–114, 2005.

[26] J. Im, J. Rhee, J. R. Jensen, and M. E. Hodgson, "An automated binary change detection model using a calibration approach," *Remote Sens. Environ.*, vol. 106, no. 1, pp. 89–105, 2007.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Advances Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[28] 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.

[29] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object detection via region-based fully convolutional networks," in *Proc. Advances Neural Inf. Process. Syst.*, 2016, pp. 379–387.

[30] M. Livne *et al.*, "A U-Net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease," *Frontiers Neurosci.*, vol. 13, 2019, Art. no. 97.

[31] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2811–2821, May 2018.

[32] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.

[33] 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.

[34] M. Gong, H. Yang, and P. Zhang, "Feature learning and change feature classification based on deep learning for ternary change detection in sar images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 129, pp. 212–225, 2017.

[35] F. Liu, L. Jiao, X. Tang, S. Yang, W. Ma, and B. Hou, "Local restricted convolutional neural network for change detection in polarimetric SAR images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 818–833, Mar. 2019.

[36] R. Wang, J. Zhang, J. Chen, L. Jiao, and M. Wang, "Imbalanced learning-based automatic SAR images change detection by morphologically supervised PCA-Net," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 554–558, Apr. 2019.

[37] F. Gao, X. Liu, J. Dong, G. Zhong, and M. Jian, "Change detection in SAR images based on deep semi-NMF and SVD networks," *Remote Sens.*, vol. 9, no. 5, 2017, Art. no. 435.

[38] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, "Change detection in synthetic aperture radar images based on deep neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 125–138, Jan. 2016.

[39] S. H. Khan, X. He, F. Porikli, and M. Bennamoun, "Forest change detection in incomplete satellite images with deep neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5407–5423, Sep. 2017.

[40] Q. Wang, Z. Yuan, Q. Du, and X. Li, "GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, Jan. 2019.

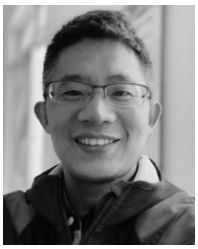
[41] C. Zhang, S. Wei, S. Ji, and M. Lu, "Detecting large-scale urban land cover changes from very high resolution remote sensing images using cnn-based classification," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 4, 2019, Art. no. 189.

[42] T. Lei, Y. Zhang, Z. Lv, S. Li, S. Liu, and A. K. Nandi, "Landslide inventory mapping from bitemporal images using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 982–986, Jun. 2019.

[43] 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.

[44] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 116, pp. 24–41, 2016.

[45] Z. Lv, W. Shi, X. Zhou, and J. A. Benediktsson, "Semi-automatic system for land cover change detection using bi-temporal remote sensing images," *Remote Sens.*, vol. 9, no. 11, 2017, Art. no. 1112.
[46] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241.
[47] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks," in *Proc. Annu. Conf. Med. Image Understanding Anal.*, 2017, pp. 506–517.
[48] V. Iglovikov and A. Shvets, "TernausNet: U-Net with vgg11 encoder pre-trained on imagenet for image segmentation," 2018, *arXiv:1801.05746*.
[49] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1026–1034.
[50] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.



ZhiYong Lv received the M.S. degree and Ph.D. degree from the School of Remote Sensing and Information Engineering at Wuhan University, Wuhan, China, in 2008 and 2014, respectively.

He was an Engineer of surveying and worked with The First Institute of Photogrammetry and Remote Sensing, Xi'An City, China, from 2008 to 2011. He is currently working with the School of Computer Science and Engineering, Xi'an University of Technology. His research interests include multihyperspectral and high-resolution remotely sensed image processing, spatial feature extraction, neural networks, pattern recognition, deep learning, and remote sensing applications.

TongFei Liu is currently working toward the master's degree in computer science with the Xi'an University of Technology, Shaanxi, China.

He research interests include spatial–spectral feature extraction, pattern recognition, ground target detection, and land cover/land use change detection, through remote sensing image with VHR, such as satellite imagery and aerial images.

XiangBing Kong received the Ph.D. degree from School of Remote sensing and information Engineering at Wuhan University, Wuhan, China, in 2012.

He is currently working with the Yellow River Institute of Hydraulic Research, Zhengzhou, China. His research interests include multisource remote sensing image processing and intelligent information extraction, ecological environment monitoring and evaluation.

Cheng Shi received the B.S. degree from the Xi'an University of Architecture and Technology, Shanxi, China, in 2009, the M.S. and ph.D. degrees from Xidian University, Shanxi, China, in 2012 and 2016, respectively.

She worked as the Postdoctoral Research Fellow with the University of Macau, China, in 2018. She is currently working with School of Computer Science and Engineering, Xi'an University of Technology. Her research interests include deep learning, image processing, and pattern recognition.

Jón Atli Benediktsson received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984, and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, USA, in 1987 and 1990, respectively.

He is currently a Pro Rector for academic affairs and Professor of electrical and computer engineering with the University of Iceland. He is a co-founder of the biomedical start up company Oxymap. He has authored or coauthored extensively in several fields. His research interests include remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing.

Prof. Benediktsson is a fellow of SPIE. He was the recepient of the Stevan J. Kristof Award from Purdue University, in 1991, as the Outstanding Graduate Student in remote sensing. In 1997, he was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award, the IEEE Third Millennium Medal, in 2000. In 2004, he was a co-recipient of the University of Iceland's Technology Innovation Award, and the recepient of the yearly research award from the Engineering Research Institute of the University of Iceland, in 2006, and the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society, in 2007. He was a co-recipient of the 2012 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Paper Award, and the IEEE GRSS Highest Impact Paper Award, in 2013. He was the recepient of the IEEE/VFI Electrical Engineer of the Year Award, in 2013. He was the President of the IEEE Geoscience and and Remote Sensing Society (GRSS) from 2011 to 2012 and has been on the GRSS AdCom, since 2000. He was Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), from 2003 to 2008, and has served as an Associate Editor for TGRS, since 1999; the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, since 2003; and IEEE ACCESS, since 2013. He is on the International Editorial Board of the *International Journal of Image and Data Fusion* and was the Chairman of the Steering Committee of IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (J-STARS) from 2007 to 2010.