

Landslide Inventory Mapping From Bitemporal Images Using Deep Convolutional Neural Networks

Tao Lei[✉], Member, IEEE, Yuxiao Zhang, Zhiyong Lv[✉], Shuying Li[✉],
Shigang Liu, and Asoke K. Nandi[✉], Fellow, IEEE

Abstract—Most of the approaches used for Landslide inventory mapping (LIM) rely on traditional feature extraction and unsupervised classification algorithms. However, it is difficult to use these approaches to detect landslide areas because of the complexity and spatial uncertainty of landslides. In this letter, we propose a novel approach based on a fully convolutional network within pyramid pooling (FCN-PP) for LIM. The proposed approach has three advantages. First, this approach is automatic and insensitive to noise because multivariate morphological reconstruction is used for image preprocessing. Second, it is able to take into account features from multiple convolutional layers and explore efficiently the context of images, which leads to a good tradeoff between wider receptive field and the use of context. Finally, the selected PP module addresses the drawback of global pooling employed by convolutional neural network, FCN, and U-Net, and, thus, provides better feature maps for landslide areas. Experimental results show that the proposed FCN-PP is effective for LIM, and it outperforms the state-of-the-art approaches in terms of five metrics, *Precision*, *Recall*, *Overall Error*, *F-score*, and *Accuracy*.

Index Terms—Change detection, deep convolutional network, landslide inventory mapping (LIM), multivariate morphological reconstruction (MMR).

I. INTRODUCTION

LANDSLIDES are important geomorphic agents that occur over a wide variety of spatial and temporal scales in many mountains and, thus, reshape the landscape and change

Manuscript received October 22, 2018; revised December 4, 2018; accepted December 17, 2018. Date of publication January 18, 2019; date of current version May 21, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61461025, Grant 61871259, Grant 61811530325 (NSFC-Royal Society), Grant 61701387, Grant 61873155, and Grant 61672333, and in part by the National Science Foundation of Shanghai under Grant 16JC1401300. (*Corresponding author:* Shuying Li.)

T. Lei is with the School of Electronical and Information Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China, and also with the School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: leitao@sust.edu.cn).

Y. Zhang is with the School of Electronical and Information Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China (e-mail: xyu@sust.edu.cn).

Z. Lv is with the School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China (e-mail: Lvzhiyong_fly@hotmail.com).

S. Li is with the School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710121, China (e-mail: angle_lsy@163.com).

S. Liu is with the School of Computer Science, Shaanxi Normal University, Xi'an 710119, China (e-mail: shgliu@gmail.com).

A. K. Nandi is with the Department of Electronic and Computer Engineering, Brunel University London, Uxbridge UB8 3PH, U.K., and also with the Key Laboratory of Embedded Systems and Service Computing, College of Electronic and Information Engineering, Tongji University, Shanghai 200092, China (e-mail: asoke.nandi@brunel.ac.uk).

Color versions of one or more of the figures in this letter are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2018.2889307

local topography. Landslides can be caused by earthquakes, rainfalls, water-level change, storm waves, human activities, and so on. A sudden and rapid landslide event is often associated with fatalities, environmental degradation, damages to businesses, buildings, roads, public utilities, and so on. For instance, landslides cause damages of nearly U.S. \$1 billion in China and more than U.S. \$3 billion in Japan annually [1].

Landslide inventory mapping (LIM) focuses on outlining slide boundaries and neglecting the wealth of information revealed by internal deformation features. It is able to provide some significant information, e.g., the sizes, spatial distributions, number, and types of landslides, for disaster relief strategy and hazard prevention. Thus, LIM is one of the most important features in landslide risk assessment.

So far, LIM relies on the visual interpretation of aerial photographs and intensive field surveys, which are highly labor-intensive and time-consuming for mapping of large areas. With the rapid progress of machine learning and remote sensing technologies, a large number of advanced approaches used for LIM have been proposed in recent years. Most of them depend on change detection that aims to detect the changed information of target areas by analyzing the multitemporal images acquired in different time of the same geographical area. The popular ones can be roughly divided into three categories: threshold-based approaches [1], [2], approaches based on feature extraction and feature classification [3]–[5], and deep learning approaches [6].

The approaches in the first category can generate landslide areas by computing one or more thresholds for the difference image of a pair of bitemporal images. However, they are sensitive to noise and have a low robustness for different landslide images. Although Lv *et al.* [1] employed a multithreshold method and a voting strategy to improve the difference image, the approach requires a presegmentation and has a low computational efficiency.

The second category of approaches is often composed of two parts: feature extraction and feature classification [7]. Most of them use unsupervised learning algorithms, such as *k*-means, fuzzy *c*-means clustering (FCM), Gaussian mixture model, etc., [8] to achieve change detection from a difference image. Because the difference image of bitemporal images includes lots of noise caused by imaging devices or illumination, some image preprocessing and postprocessing operations are necessary to improve LIM results. Li *et al.* [3] employed edge-based level set evolution (ELSE) and region-based level set evolution (RLSE) to track the initial change detection profiles, leading to better landslide candidate areas and LIM

results. Furthermore, they also proposed change detection based on Markov random field (CDMRF) for LIM [4]. However, the performance of ELSE, RLSE, and CDMRF seriously depend on the quality of difference images and parameter selection. To reduce such dependencies, Lei *et al.* [5] employed morphological reconstruction and a fast clustering approach to distinguish changed and unchanged areas for LIM. The method efficiently utilizes the structuring information of bitemporal images, providing better LIM results than ELSE, RLSE, and CDMRF.

Based on deep learning technologies, Gong *et al.* [9] proposed a change detection approach using a deep neural network for synthetic aperture radar images. As the proposed network architecture included only a few hidden layers, and it adopted full connection without using convolutional operation, the context of images is utilized inefficiently. Moreover, it is difficult to train the network due to the full connection. To obtain better results, Liu *et al.* [10] proposed a new deep convolutional coupling network that is fully unsupervised without using any labels. To apply convolutional neural network (CNN) [11] to landslide recognition, Ding *et al.* [12] employ CNN and texture change detection to recognize landslides. Because CNN employs multiple pooling layers and a fully connectional layer to achieve classification tasks, the final result is coarse and has a low recognition accuracy.

To address the aforementioned issues, we propose a symmetric fully convolutional network within pyramid pooling (FCN-PP) that is able to learn better image features to improve LIM results. Before applying the FCN-PP, a multivariate morphological reconstruction (MMR) [13] is performed on training and testing images. Our main contributions are summarized as follows.

- 1) MMR is employed to filter noise and nonlandslide areas in difference images, which is helpful for improving training and testing images.
- 2) We design a powerful deep convolutional network that is able to tradeoff the use of context and the localization accuracy, and the network has an elegant architecture.

II. METHODOLOGY

The proposed approach is composed of two parts. The first part is image preprocessing using MMR. Because a pair of bitemporal high-resolution remote sensing images is captured at different time, different imagery environmental factors lead to a poor difference image that usually includes many falsely changed areas or misses many truly changed areas. Clearly, it is necessary to implement image filtering for preprocessing. The second part is to construct a deep CNN that is able to utilize the context efficiently and provides high localization accuracy for LIM.

A. MMR for Image Preprocessing

The difference image of bitemporal images includes lots of noise and falsely changed regions. There are two ways to address the problem. One is to implement image filtering on bitemporal images, and then compute the difference image; the other one is to compute the difference image and

then implement image filtering. Because the former requires performing image filtering on preevent and postevent images, respectively, it needs a long running time. To speed up the implementation of the proposed approach, we adopt the second way, i.e., performing image filtering on the difference image.

Since the frequency domain filtering is unsuitable for removing falsely changed areas and preserving the structuring information of truly changed areas, we use a morphological filter that is an efficient spatial filter, to achieve preprocessing. As traditional morphological filters are only effective for grayscale images, we use MMR operators here. MMR has two advantages for multichannel image filtering. First, MMR is able to filter noise while maintaining object details. Second, MMR has a low computational complexity, for fast lexicographic order denoted by \leq_{PCA} , is employed.

Let $\mathbf{v}(R, G, B)$ represent a color pixel, and $\mathbf{v}'(P_f, P_s, P_t)$ denote the transformed color pixel using the principal component analysis (PCA), the operation \leq_{PCA} is defined as follows [13]:

$$\mathbf{v}'_1 \leq_{\text{PCA}} \mathbf{v}'_2 \Leftrightarrow \begin{cases} P_f^1 < P_f^2, & \text{or} \\ P_f^1 = P_f^2, & \text{and } P_s^1 < P_s^2 \text{ or} \\ P_f^1 = P_f^2, & \text{and } P_s^1 = P_s^2 \text{ and } P_t^1 \leq P_t^2. \end{cases} \quad (1)$$

Let $\vec{\epsilon}$ and $\vec{\delta}$ denote multivariate morphological erosion and dilation, respectively. According to (1), $\vec{\epsilon}$ and $\vec{\delta}$ are defined as follows:

$$\begin{cases} \vec{\epsilon}_B(\mathbf{g}) = \wedge_{b \in B} \mathbf{g}_{-b} \\ \vec{\delta}_B(\mathbf{g}) = \vee_{b \in B} \mathbf{g}_{-b} \end{cases} \quad (2)$$

where \vee_{PCA} and \wedge_{PCA} denote the supremum and infimum based on lexicographical ordering \leq_{PCA} , respectively. \mathbf{g} represents a color image, and B is a structuring element. If we let \mathbf{f}, \mathbf{g} are a marker and a mask image, respectively, where $\mathbf{f} \leq_{\text{PCA}} \mathbf{g}$. We propose a morphological closing reconstruction operation, denoted by R^C , and it is defined as follows:

$$R_f^C(\mathbf{g}) = R_f^{\vec{\epsilon}}(R_f^{\vec{\delta}}(\mathbf{g})) \quad (3)$$

where $R_f^{\vec{\epsilon}}(\mathbf{g}) = \vec{\epsilon}_f^{(i)}(\mathbf{g})$, $R_f^{\vec{\delta}}(\mathbf{g}) = \vec{\delta}_f^{(i)}(\mathbf{g})$. Like conventional morphological reconstruction operation, $\vec{\epsilon}_f^{(1)}(\mathbf{g}) = \vec{\epsilon}(\mathbf{g}) \vee_{\text{PCA}} \mathbf{f}$, $\vec{\epsilon}_f^{(i)}(\mathbf{g}) = \vec{\epsilon}(\vec{\epsilon}_f^{(i-1)}(\mathbf{g})) \vee_{\text{PCA}} \mathbf{f}$, duality, $\vec{\delta}_f^{(1)}(\mathbf{g}) = \vec{\delta}(\mathbf{g}) \wedge_{\text{PCA}} \mathbf{f}$, $\vec{\delta}_f^{(i)}(\mathbf{g}) = \vec{\delta}(\vec{\delta}_f^{(i-1)}(\mathbf{g})) \wedge_{\text{PCA}} \mathbf{f}$, and $2 \leq i \leq n$, $i, n \in N^+$, where n satisfies $\vec{\epsilon}_f^{(n)}(\mathbf{g}) = \vec{\epsilon}_f^{(n-1)}(\mathbf{g})$ and $\vec{\delta}_f^{(n)}(\mathbf{g}) = \vec{\delta}_f^{(n-1)}(\mathbf{g})$.

We applied MMR to a difference image to remove small image structures while preserving large object structures, which is helpful for subsequent feature learning.

B. Network Structure of FCN-PP

In practical applications, our purpose is to detect landslide areas in a postevent image. It is clear that we can use image segmentation algorithms to obtain landslide areas from the difference image of bitemporal images. Because the final output only includes changed and unchanged areas, the problem is viewed as a pixel-level binary classification, i.e., a binary segmentation task. The typical use of deep convolutional

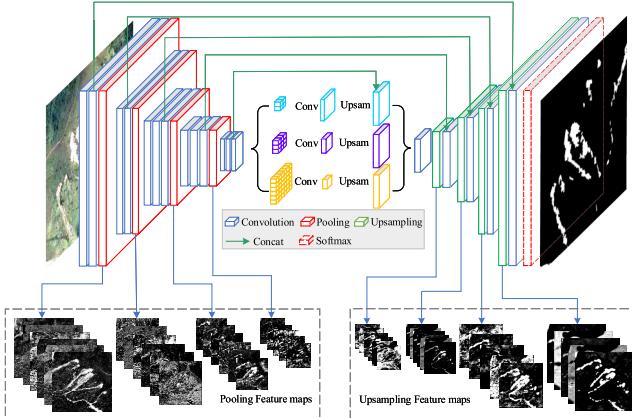


Fig. 1. Structure of FCN-PP. It is a symmetric fully convolutional network that includes four pooling layers and four deconvolutional layers.

networks is on classification tasks. However, a pixel-level classification task (semantic segmentation) is more complex due to the requirement of localization.

Although CNN is able to achieve effective image classifications, it provides a poor result on image segmentation since it employs global pooling and misses the spatial information of images. FCN [14] overcomes the problem by taking into account the combined features from multiple convolutional layers, which results in a better localization and the use of context. Consequently, FCN provides better image segmentation results than CNN. Inspired by the idea of FCN, we presented an FCN-PP to obtain better LIM results. The proposed FCN-PP is able to capture wider receptive field, it eventually overcomes the drawback of global pooling. Fig. 1 shows the proposed network structure.

In Fig. 1, the FCN-PP similarly yields a U-shape architecture that includes four pooling layers and corresponding four upsampling layers. It also has an elegant architecture structuring because pooling layers on the left and upsampling layers on the right are symmetric. Furthermore, the PP module is integrated into FCN-PP to overcome the global pooling problem.

From fine to coarse, we choose a three-level PP module that includes three different scales (convolutional kernels: 5×5 , 10×10 , and 15×15 ; strides: 5, 10, and 15), where the first scale (5×5) is marked by cyan color, the second and the third scales (10×10 , 15×15) are marked by purple and yellow, respectively. Then, we use 1×1 convolution to reduce the dimension of the three different size feature maps to achieve upsampling. Here, bilinear interpolation is used for upsampling to obtain feature maps of the same size as the original feature map. The final output of pyramid module is a fusion result of multilevel feature maps. The average pooling is chosen in the pyramid module as it provides better global information than the max pooling.

C. Pyramid Pooling

To verify the validity of the PP module for landslide feature learning, we extract the outputs from convolutional layer at different scales. Fig. 2 shows that a large convolutional kernel

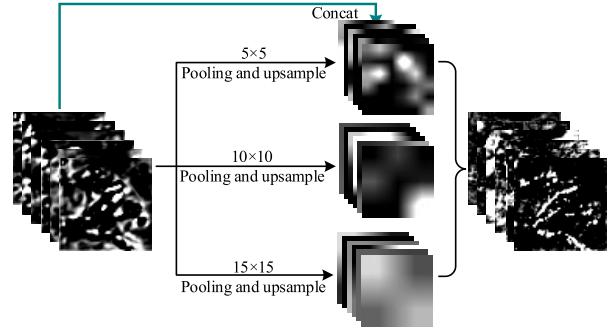


Fig. 2. PP module that includes three layers, where a small convolutional kernel of size 5×5 is used for the first layer, a middle convolutional kernel of size 10×10 is used for the second layer, and a larger convolutional kernel of size 15×15 is used for the last layer.

means a wider received field that is helpful for global feature representation, while, a small convolutional kernel means a narrower received field that is helpful for local detail feature representation. We combined multiscale scaled features to achieve a stronger feature representation than single-scale feature learning. As landslide areas have a serious spatial uncertainty, it is difficult to learn effective landslide features. The PP module is able to overcome the difficulty and it is suitable for feature representation of landslide areas.

In Fig. 2, the input is the same as feature maps of FCN, which are pooled by using different-size convolutional kernels. Then, these results are upsampled and fused with the feature maps from previous convolutional layers, leading to a final output with more accurate localization and better semantic information than the feature maps of FCN. Consequently, the proposed FCN-PP achieves a tradeoff between the use of context and the localization accuracy, making a significant improvement over the CNN approach in LIM. Moreover, the elegant architecture leads to a requirement of a small number of training samples.

III. EXPERIMENTS

The proposed FCN-PP is compared with two groups of popular approaches, i.e., unsupervised (ELSE [3], RLSE [3], CDMRF [4], and change detection-based fast FCM (CDFFCM) [5]) and supervised (CNN [12], FCN [14], and U-Net [15]) change detection for LIM. The first group of comparative approaches is implemented with MATLAB 2017b and the second group is implemented with PyTorch on a workstation with Intel Xeon CPU E5-1620v4, 3.5 GHz, four cores, 64-GB RAM, and double NVIDIA GTX 1080 GPU.

In our experiments, the parameter values of the first group of comparative approaches follow the original papers. In the proposed FCN-PP, because the convolutional layers before the PP module are used for feature extraction, the corresponding parameters are initialized using the parameters of the first four convolutional layers in the VGG-16. We set a small learning rate that is 1×10^{-7} for pretrained convolutional layers to avoid overfitting. The stochastic gradient descent with a constant learning rate of 1×10^{-4} , weight decay of 0.0005, momentum of 0.99, minibatch size of 4, and epochs of 30 were used to

train the proposed network. In addition, the structure element employed by MMR is a disk of size 1×1 .

A. Data Description

Five pairs of bitemporal images on A–E areas in Hong Kong, were captured by the Zeiss RMK TOP 15 Aerial Survey Camera System at a flying height of approximately 2.4 km on December 2007 and on November 2014, respectively [1]. Due to the geometrical resolution of bitemporal images is 0.5 m, the captured images have a large size. We cropped A–E areas to obtain five interesting areas including different types of landslide areas. The size of A–E areas are 750×950 , 1252×2199 , 923×593 , 1107×743 , and 826×725 , respectively. Because it is impossible to build a large data set of bitemporal landslide images, we built a small data set that is considered as a training set in this letter. To distinguish the training data and testing data, we first cropped three typical areas from A–C areas, where different kinds of landslide occur. The rest of A–C areas and D–E areas are then used for training data. The training images are cropped overlappingly. The testing images and training images have the same size of 473×473 , and they have no overlapped areas. Finally, we obtain 139 training images that are overlapped. To increase the training data, each image is rotated by $\pm 30^\circ$, horizontally and vertically flipped, sheared by $\pm 30^\circ$, and scaled to 80% and 125% of its original size, but the final image has the same size as the original image. Finally, we obtain 1112 training image pairs of 473×473 and 3 testing image pairs.

B. Evaluation Metrics

To compare quantitatively the existing approaches with the proposed FCN-PP, five quantitative evaluation indices [5], [16] are presented: Precision = P_{lm}/P_l , Recall = P_{lm}/P_r , Overall error = $(P_{over} + P_{rum})/P_t$, F-score = $(2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$, and Accuracy = $P_{lm}/(P_l + P_{rum})$, where P_{lm} is the total number of pixels of the detected landslides that are matching with the relevant ground truth, P_r is the total number of pixels of the ground truth, P_l is the total number of pixels of the detected landslides, and P_{rum} is the total number of pixels of the relevant ground truth that is unmatched with the detected landslides. P_{over} is the total number of pixels of detected false landslides. P_t is the total number of pixels of the test image.

A large value of *Precision* means a small number of false alarms and a large value of *Recall* means a small number of missed detections. A small value of *Overall error* (*OE*) means a small sum of false alarms and missed detections. Interestingly, *OE*, *F-score*, and *Accuracy* reveal the overall detection performance, and a good approach for LIM corresponds to large values of *F-score*, and *Accuracy* but a small value of *OE*.

C. Results and Quantitative Evaluation

Due to the limitation of the length of this letter, we only presented the comparative results of A-area and B-area. Figs. 3(a) and (b) and 4(a) and (b) show two pairs of bitemporal images. We can see that the landslide areas are simple and

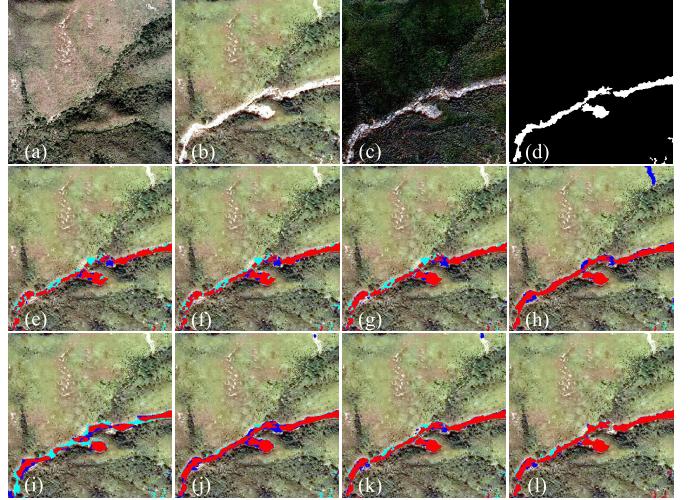


Fig. 3. Comparison of results on the A-area using different approaches, where the red areas are true-detected landslides, the cyan areas are false-detected landslides, and the blue areas are true-missed landslides. (a) Preevent image. (b) Postevent image. (c) Difference image. (d) Ground truth. (e) ELSE. (f) RLSE. (g) CDMRF. (h) CDFFCM. (i) CNN. (j) FCN. (k) U-Net. (l) FCN-PP.

continuous in Fig. 3(d) but they are complex and discontinuous in Fig. 4(d). Therefore, it is more difficult to extract the landslide areas in Fig. 4(b) than in Fig. 3(b) using traditional approaches. Figs. 3(c) and 4(c) show the difference images of preevent and postevent images. It is clear that each difference image includes lots of noises that influence the detection of the true landslides areas. Figs. 3(e)–(h) and 4(e)–(h) show landslide areas detected by four conventional approaches, ELSE, RLSE, CDMRF, and CDFFCM. Because ELSE, RLSE, and CDMRF employ general image segmentation models to achieve landslide areas detection, they are sensitive to noise. The detected landslide areas include lots of discontinuous areas that are continuous in ground truths. Although CDFFCM addresses the problem by using image filtering and improved FCM algorithm that incorporates spatial information of images into its objective function, some false landslide areas are detected as shown in Figs. 3(h) and 4(h). Compared with unsupervised learning approaches, CNN is able to capture the semantic information of landslide areas, but the detected areas are coarse as shown in Figs. 3(i) and 4(i). FCN provides a better result than CNN and four unsupervised learning approaches. However, the detail of landslide areas is removed in Figs. 3(i) and 4(i) since the global pooling is adopted by FCN. Note that although the experimental results of the C-area are not shown here, the practical results also show that the proposed FCN-PP is superior to the comparative approaches.

For quantitative evaluation of the proposed FCN-PP, we compare experimental results with ground truths according to five performance indices. The experiments are shown in Tables I and II. It can be seen that *Precision* and *Recall* are inconsistent for the evaluation of results. The ELSE, RLSE, and CDMRF obtain high *Precision* but low *Recall* values, while the CDFFCM obtains low *Precision* but high

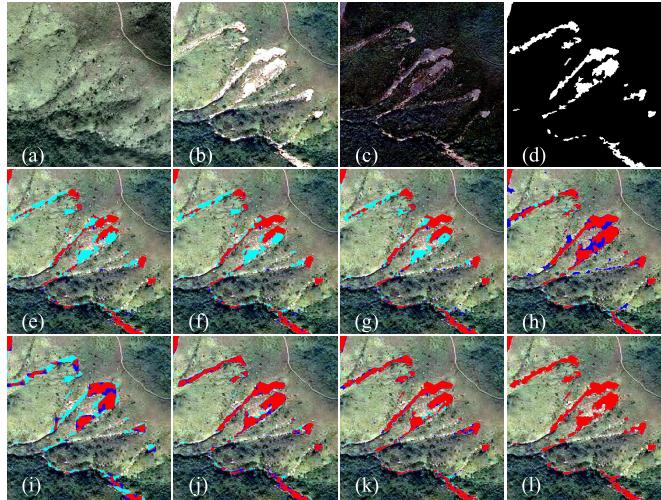


Fig. 4. Comparison of results on the B-area using different approaches. (a) Preevent image. (b) Postevent image. (c) Difference image. (d) Ground truth. (e) ELSE. (f) RLSE. (g) CDMRF. (h) CDFFCM. (i) CNN. (j) FCN. (k) U-Net. (l) FCN-PP.

TABLE I

QUANTITATIVE RESULTS OF A-AREA. LARGER VALUES ARE BETTER FOR *Precision*, *Recall*, *F-Score*, *Accuracy* WHILE SMALLER VALUES ARE BETTER FOR *OE*. THE BEST VALUES ARE IN **BOLD**

Methods	<i>Precision</i>	<i>Recall</i>	<i>OE</i>	<i>F-score</i>	<i>Accuracy</i>
ELSE [3]	86.79	74.02	1.75	79.90	66.53
RLSE [3]	86.61	75.87	1.68	80.89	67.91
CDMRF [4]	84.89	79.15	1.64	81.92	69.38
CDFFCM [5]	79.30	91.01	1.54	84.75	73.54
CNN [12]	75.47	62.16	2.73	68.17	51.71
FCN [14]	79.10	88.45	1.64	83.51	71.69
U-Net [15]	85.65	83.92	1.42	84.78	73.58
FCN-PP	89.84	90.32	0.93	90.08	81.95

TABLE II

QUANTITATIVE RESULTS OF B-AREA. LARGER VALUES ARE BETTER FOR *Precision*, *Recall*, *F-Score*, *Accuracy* WHILE SMALLER VALUES ARE BETTER FOR *OE*. THE BEST VALUES ARE IN **BOLD**

Methods	<i>Precision</i>	<i>Recall</i>	<i>OE</i>	<i>F-score</i>	<i>Accuracy</i>
ELSE [3]	91.24	51.29	3.93	65.67	48.89
RLSE [3]	91.14	56.00	3.63	69.37	53.10
CDMRF [4]	87.42	60.01	3.56	71.17	55.24
CDFFCM [5]	73.52	88.42	3.18	80.28	67.06
CNN [12]	71.44	51.44	5.07	59.82	42.67
FCN [14]	83.94	84.39	2.33	84.17	72.66
U-Net [15]	87.88	87.93	1.77	87.91	78.42
FCN-PP	96.03	95.62	0.61	95.82	91.99

Recall values. Considering that *OE*, *F-score*, and *Accuracy* evaluate the overall performance of an approach, CDFFCM is superior to ELSE, RLSE, and CDMRF according to Tables I and II. Amongst the existing supervised approaches, U-Net is superior to CNN and FCN. What is abundantly clear that the proposed FCN-PP obtains the best performance index as shown in Tables I and II.

IV. CONCLUSION

In this letter, we proposed an FCN-PP to achieve end-to-end LIM. The FCN-PP adopts multiple layer connection to incorporate the low- and high-dimensional features into the final feature map. The PP module is integrated into the FCN-PP and is able to efficiently exploit the spatial multiscale features of landslide areas, which addresses the drawback of global pooling and, thus, outputs a better feature map with stronger feature representation capability than CNN, FCN, and U-Net. Experimental results show that the proposed FCN-PP generates satisfactory LIM results without hard-tuning parameters, the FCN-PP clearly outperforms the state-of-the-art approaches for LIM, and it performs the best in five metrics, i.e., *Precision*, *Recall*, *OE*, *F-score*, and *Accuracy*.

REFERENCES

- [1] Z. 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, Mar. 2018.
- [2] Y.-Q. Jin, "Change detection of enhanced, no-changed and reduced scattering in multi-temporal ERS-2 SARimages using the two-thresholds EM and MRF algorithms," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 6, Jul. 2005, pp. 3994–3997.
- [3] Z. Li, W. Shi, S. 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, Mar. 2016.
- [4] 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, Dec. 2016.
- [5] T. Lei, D. Xue, Z. Lv, S. Li, Y. Zhang, and A. K. Nandi, "Unsupervised change detection using fast fuzzy clustering for landslide mapping from very high-resolution images," *Remote Sens.*, vol. 10, no. 9, p. 1381, Aug. 2018.
- [6] 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. 2018, doi: [10.1109/TGRS.2018.2849692](https://doi.org/10.1109/TGRS.2018.2849692).
- [7] Q. Wang, X. He, and X. Li, "Locality and structure regularized low rank representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: [10.1109/TGRS.2018.2862899](https://doi.org/10.1109/TGRS.2018.2862899).
- [8] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, and A. K. Nandi, "Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 3027–3041, Oct. 2018.
- [9] 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.
- [10] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 545–559, Mar. 2018.
- [11] Q. Wang, S. Liu, J. Chanussot, and X. Li, "Scene classification with recurrent attention of VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: [10.1109/TGRS.2018.2864987](https://doi.org/10.1109/TGRS.2018.2864987).
- [12] A. Ding, Q. Zhang, X. Zhou, and B. Dai, "Automatic recognition of landslide based on CNN and texture change detection," in *Proc. 31st Youth Academic Annu. Conf. Chin. Assoc. Automat.*, Nov. 2016, pp. 444–448.
- [13] T. Lei, X. Jia, Y. Zhang, S. Liu, H. Meng, and A. K. Nandi, "Superpixel-based Fast fuzzy C-means clustering for color image segmentation," *IEEE Trans. Fuzzy Syst.*, to be published, doi: [10.1109/TFUZZ.2018.2889018](https://doi.org/10.1109/TFUZZ.2018.2889018).
- [14] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
- [15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Nov. 2015, pp. 234–241.