Landslide Mapping Using Multilevel-Feature-Enhancement Change Detection Network

Lukang Wang, Min Zhang, Xiaoqi Shen, and Wenzhong Shi

***Abstract*—Landslide mapping (LM) from bitemporal remote sensing images is essential for disaster prevention and mitigation. Although bitemporal change detection technology has been applied for LM, there remains room for improvement in its accuracy and automation. In this article, a multilevel feature enhancement network (MFENet) is proposed for LM based on modules built in convolutional neural networks (CNNs) like CNN-Attention. MFENet mainly consists of three modules: the postevent feature enhancement module (PFEM), the bifeature difference enhancement module (BFDEM), and the flow direction calibration module (FDCM). Specifically, the main role of PFEM is to selectively fuse postevent multilayer features to provide discriminative postevent features. BFDEM fuses the multilayer differences of both pre-event and postevent features to generate high-quality change detection features, which are sufficiently powerful to distinguish foreground from background. FDCM uses a digital elevation model to calibrate the flow direction of each pixel of the landslide detection results to complete the LM task. Experiments were conducted to test the effectiveness of MFENet on two real-world regions, Lantau Island and Sharp Peak, Hong Kong, where landslides occurred after rainstorms. Compared with other state-of-the-art general change detection methods and landslide-specific change detection methods, the proposed method outperforms all metrics, with its intersection over union reaching 87.23%. The availability of additional features and the generalization performance of MFENet are demonstrated experimentally. It is anticipated that the proposed network will further contribute to disaster prevention and mitigation.**

***Index Terms*—Change detection, convolutional neural network (CNN), flow direction, landslide mapping (LM), remote sensing images.**

I. Introduction

LANDSLIDES are one of the major natural disasters, causing huge economic losses and casualties worldwide [1], [2]. Under the background of frequent occurrences of severe weather around the world caused by environmental problems, landslide disasters appear to be occurring more frequently [3], [4]. Therefore, there is positive significance in quickly and accurately obtaining the location, boundary, and flow direction of landslides that follow disasters by the means of landslide mapping (LM) for postevent rescue, disaster investigation, and disaster prevention and mitigation [5].

During recent years, the development of remote sensing has provided new opportunities for LM [6], [7]. Many LM methods based on remote sensing images have been developed, and such methods can be divided into two types: object detection based on postevent images and change detection based on pre-event and postevent bitemporal images [8]. Deep learning technology shines in remote sensing image processing [9], [10], making it possible to achieve more accurate and automated LM [11].

LM using postevent images is usually based on textural, spectral, and morphological features. Of specific interest is that it can also be assisted by additional data such as that gained by the digital elevation model (DEM) [7], [11]. From Landsat-8 images, Yu et al. [12] proposed a method for LM, which introduced saliency enhancement to recognize landslides and used additional DEM data to remove the ground objects of plain areas in which landslides are less likely to occur. Cheng et al. [13] proposed a classification method based on the k-nearest neighbor, combining the bag-of-visual-words and the probabilistic latent semantic analysis model, which automatically divides a given image into landslide areas and nonlandslide areas. Yu et al. [14] used a deep convolutional neural network (CNN) to initially determine the landslide and then used an improved region-growing algorithm to extract the region and boundary of the landslide. Qi et al. [15] combined the characteristics of U-Net and residual network (ResNet) to design a deep learning model to complete automatic LM. "Different objects with the same spectral features" is the specific difficulty regarding this type of method [16], as it is necessary to, either manually or by using additional data, remove exposed rocks, roads, and other ground objects similar to the specific spectral landslide characteristics. This action leads to poorer LM results, whereas postprocessing reduces the associated automation.

The accumulation of historical data has created new possibilities for LM in the use of bitemporal remote sensing images [17]. Landslides mostly occur in vegetation-covered areas, in which the texture and spectral features of the ground before and after

Manuscript received 12 December 2022; revised 3 February 2023; accepted 10 February 2023. Date of publication 15 February 2023; date of current version 13 April 2023. This work was supported in part by the Otto Poon Charitable Foundation Smart Cities Research Institute, The Hong Kong Polytechnic University (Work Program: CD03), in part by the Ministry of Science and Technology of the People's Republic of China under Grant 2019YFB2103102, in part by the Beijing Key Laboratory of Urban Spatial Information Engineering under Grant 2020101, and in part by The Hong Kong Polytechnic University under Grant ZVU1. *(Corresponding author: Wenzhong Shi.)*

Lukang Wang and Xiaoqi Shen are with the School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China (e-mail: wanglukang@cumt.edu.cn; shenxiaoqi@cumt.edu.cn).

Min Zhang and Wenzhong Shi are with the Smart Cities Research Institute and the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong (e-mail: lsgi-min.zhang@polyu.edu.hk; john.wz.shi@polyu.edu.hk).

Digital Object Identifier 10.1109/JSTARS.2023.3245062

a landslide are generally different. However, disturbed objects like rocks still maintain their original features. Therefore, LM using bitemporal image change detection methods has become one of the mainstreams [6]. Such methods are generally divided into two steps: the generation of change detection images (CDIs) [18], [19], [20], and the segmentation of CDIs for LM. Mondini et al. [21] used the normalized differential vegetation index (NDVI), image spectral angle, principal component analysis (PCA), and independent component analysis (ICA) to generate CDIs and then employed multivariate classification techniques, such as logistic regression, linear discriminant analysis, and quadratic discriminant analysis, to detect landslides. Li et al. [22] applied change vector analysis to generate CDIs and employed a more robust threshold method to generate the landslide mask. Then, a change-detection-based Markov random field (CDMRF) method was proposed for LM using the spectral and spatial contextual features of landslides. Lu et al. [23] first generated CDIs through NDVI, PCA, and ICA and then combined the CDMRF method for LM. Lv et al. [8] proposed an approach based on adaptive region shape similarity, which used the neighborhood features of pixels to obtain richer context information to generate CDIs, and finally used the threshold segmentation method to segment the CDIs to complete LM. These methods rely on the quality of CDIs and also require empirical manual selection of optimal features and thresholds, resulting in a low degree of automation. In addition, bitemporal SAR images are also used for LM, and their effect is better than that of optical remote sensing images when there is cloud cover [24], [25]. But SAR-based LM relies on additional data such as DEM.

Deep learning methods can automatically extract more effective features through deep convolutional layers to complete the end-to-end LM [11], [26]. It is a potential LM method with a high degree of automation and accuracy equal to or better than traditional methods. Zhang et al. [27] used a deep CNN to learn landslide features from historical images, then proposed a change detection CNN to detect landslides from bitemporal images, and postprocessed landslide objects to obtain attribute information such as their trail, flow direction, and source points. Similarly, fully convolutional network (FCN)-based methods were proposed to extract features from differential images or directly superimpose depth features from bitemporal images for LM [26], [27]. Fang et al. [30] combined a generative adversarial network and a Siamese neural network to generate bitemporal feature images and used pixelwise Euclidean distance for LM. Based on U-Net and ResNet, Su et al. [31] proposed the LanD-CNN model to detect landslides, which superimposed bitemporal images and DEM as the input of the model. Amankwah et al. [32] tested the change detection methods proposed on nonlandslide datasets, such as the spatial–temporal attention neural network and Siamese nested U-Net (SNUNet) on the landslide dataset, and achieved good results. These methods can be summarized as using three feature extraction and fusion strategies

1) using differential images as the model input,
2) using the direct superimposition of bitemporal images as model inputs to extract features,
3) using the same encoder to extract features for bitemporal images and then superimposing these features into the decoding stage.

The pre-event and postevent features are equally weighted in the model.

When doing LM, more attention is paid to the new landslide. New landslides exist in the postevent image and generally occur in vegetation-covered mountainous areas. Its spectral, textural, and morphological features are obvious in the postevent image and are well differentiated from the surrounding vegetation. Taking advantage of this property combined with change detection techniques, in this article, a network based on multilevel feature enhancement is proposed for LM. First, the postevent feature enhancement module (PFEM) is designed to enhance the postevent image features to make them more discriminative. Here, the additional DEM or partial landslide segmentation results of the postevent image can be input into the network through concatenate or side supervision to further enhance the postevent feature. Second, the pre-event image feature and the enhanced postevent image feature are input to the bifeature difference enhancement module (BFDEM), which highlights the differential features of bitemporal images to obtain reliable landslide change detection results. Finally, the flow direction calibration module (FDCM) uses an eight-direction pour point model (D8) [33] to calibrate the landslide flow direction. Our main contributions are as follows.

1) We propose an end-to-end change detection network named multilevel feature enhancement network (MFENet) for LM. MFENet enhances features from two levels and can use side supervision and concatenation to input additional features.
2) PFEM and BFDEM are proposed to enhance postevent image features and bitemporal image difference features. The modules make the results have clear boundaries and consistent internal logic, and they reduce false detections and missed detections caused by interference factors.
3) FDCM uses the D8 method to calibrate the flow direction of the LM with the DEM to meet the application requirements of subsequent landslide analysis.
4) Compared with existing general change detection methods and landslide-specific change detection methods, both of which are based on deep learning, the proposed method yielded state-of-the-art (SOTA) results on landslide detection.

II. Methodology

In this section, we introduce the overview of our proposed network and then introduce the details of the modules in the network.

A. Overview

MFENet is an end-to-end network, and its architecture is shown in Fig. 1. It consists of the following four parts:

1) a feature extractor based on a Siamese network,
2) a PFEM for enhanced postevent features,

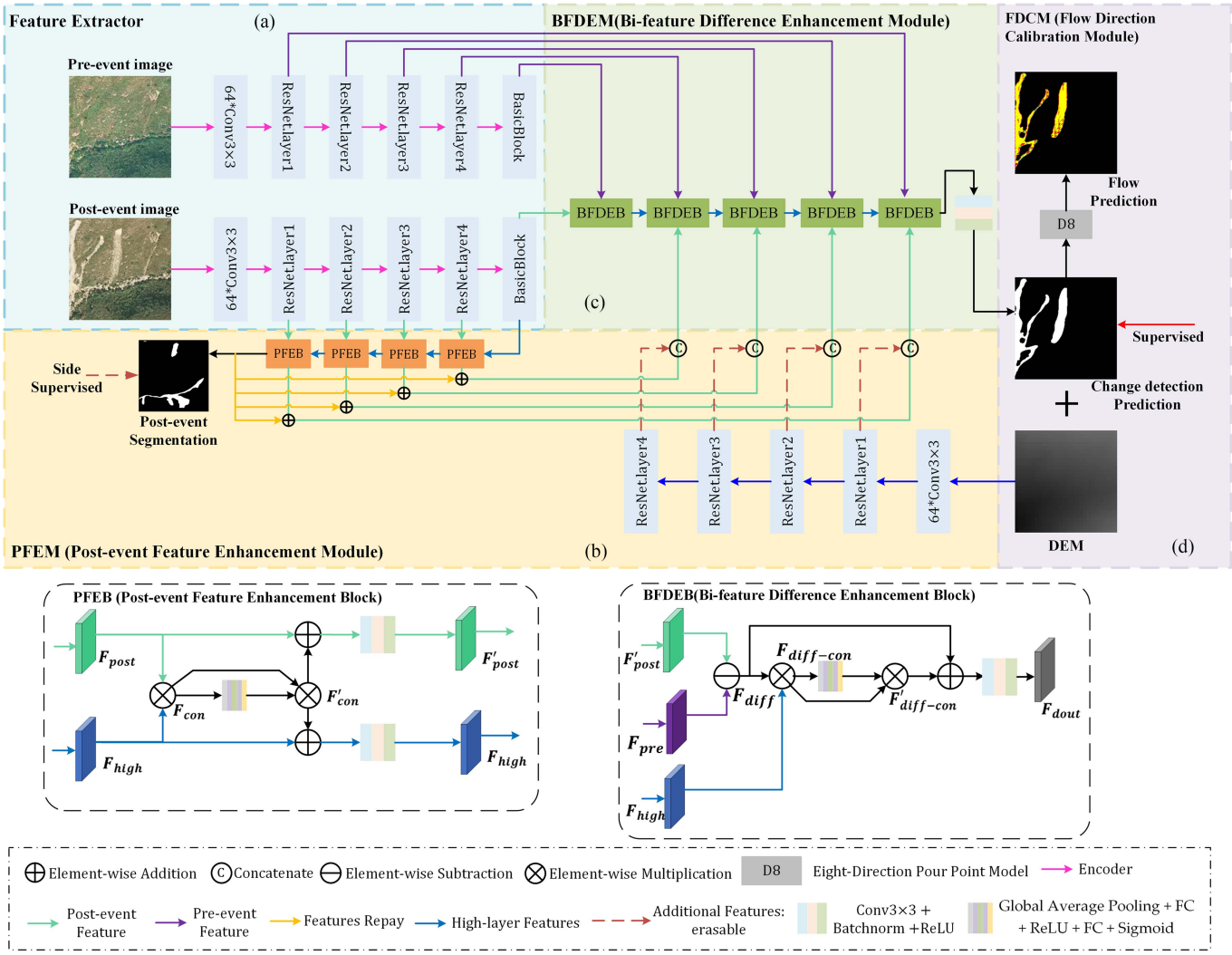

Fig. 1. Overall framework of the proposed MFENet. The framework mainly consists of four parts. (a) ResNet-34 is used as the backbone encoder for feature extractor. (b) PFEM fuses multilayer features to obtain discriminative postevent features. (c) BFDEM uses BFDEB to fuse the multilayer difference features to generate high-quality change detection features, and additional features can be input into the network in this part for end-to-end training. (d) FDCM uses DEM to calibrate the flow direction.

3) a BFDEM for enhanced bitemporal difference features,
4) an FDCM to calibrate the flow direction.

Let I_{pre} and I_{post} represent pre- and postevent images, respectively. The flow of MFENet is as follows.

Step 1: The Siamese network is a neural network framework with two branches. The "Siamese" of the neural network is realized by sharing weights. The I_{pre} and I_{post} are input into two weight-sharing branches of the Siamese network to obtain multilayer features of pre- and postevent, represented as $F_N = \{f_N^1, f_N^2, f_N^3, f_N^4, f_N^5\}$, N = {pre, post} , superscripts 1–5 represent layer1–layer5.

Step 2: The F_{post} is fed into a PFEM to get enhanced postevent features, i.e., $F'_{\text{post}} = \{f_{\text{post}}^1, f_{\text{post}}^2, f_{\text{post}}^3, f_{\text{post}}^4\}$. PFEM contains the following three stages of enhancement.

1) As the core block of PFEM, the postevent enhancement block (PFEB) fuses the different layers' features to enhance F_{post}.
2) F_{post} is corrected and refined using the feature repay (FR) mechanism before feeding it to *step 3*.
3) As optional inputs, DEM and partial landslide segmentation results are used to further enhance F_{post}.

Step 3: The F_{pre} and F'_{post} are input into BFDEM to obtain the output change detection map. The core block of BFDEM, the bifeature difference enhancement block (BFDEB), fuses F_{pre} and F'_{post} to obtain the change detection feature $F_{dout} = \{f_{dout}^1, f_{dout}^2, f_{dout}^3, f_{dout}^4, f_{dout}^5\}$. Convolutions are employed to reduce the feature dimension of f_{dout}^1 to 1, that is, the output change detection map. A combined loss function that integrates binary cross entropy [34], structure similarity [35], and intersection over union (IoU) [36] loss is used to calculate the distance between the change detection map and the ground truth to complete the training. It can be defined as follows:

$$l = l_{\text{BCE}} + l_{\text{SSIM}} + l_{\text{IoU}} \tag{1}$$

The BCE loss function calculates the loss value for each pixel for the image as follows:

$$l_{\text{bce}} = -\frac{1}{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left[gt_{ij}\log p_{ij} + (1-gt_{ij})\log(1-p_{ij})\right] \tag{2}$$

where gt_{ij} and p_{ij} represent the true value at the pixel (i, j) and the predicted value, respectively. H and W represent the height and width of the image, respectively. The SSIM loss function is integrated to focus on the integrity of the local area. Its calculation formula is given as follows:

$$l_{\text{ssim}} = 1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3}$$

where x and y represent the reference image and the predicted image, respectively. μ_x, μ_y and σ_x, σ_y represent the mean and standard deviation of x and y, and σ_{xy} represent the covariance. C_1 and C_2 are two constants to avoid the denominator being 0, in this study, $C_1 = 0.01^2$ and $C_2 = 0.03^2$. The IoU loss function focuses on the global structural information. Its calculation formula is given as follows:

$$l_{\text{IoU}} = 1 - \frac{\sum_{i=1}^{H}\sum_{j=1}^{W}(g_{ij} * p_{ij}) + 1}{\sum_{i=1}^{H}\sum_{j=1}^{W}(g_{ij} + p_{ij} - g_{ij} * p_{ij}) + 1} \tag{4}$$

where g_{ij} and p_{ij} represent the true value and the predicted value, respectively.

Step 4: The proposed FCDM uses DEM and D8 to calibrate the flow direction of landslides.

B. Feature Extractor

The feature extractor is a Siamese network with two weight-sharing branches, as shown in Fig. 1(a). First, in each branch, convolutions with a stride width of 1 and kernel size of 3 (*Conv3×3_str1*) are used to increase the feature dimension to 64. Second, the first four layers of ResNet34 [37], i.e., *layer1–layer4,* each of which contains a varying number of *BasicBlock* of ResNet, are used as the backbone for feature extraction. *BasicBlock* for each layer contains the same number of *Conv3×3_str1* as the number of channels and uses a skip connection known as "shortcut." Skip connection solves the degradation problems of vanishing gradients and exploding gradients when training deep neural networks. Feature extraction by using skip connections enables the network to learn features with multiscale information and various receptive fields. A *BasicBlock* as *layer5* is used to obtain deeper features for feature enhancement. Finally, *layer1 –layer5* of features can be obtained, and their channels are 64, 128, 256, 512, and 512, respectively.

C. Postevent Feature Enhancement Module

The PFEM contains three stages of enhancement as shown in Fig. 1(b), i.e., PFEB, FR, and additional features.

Stage 1: The PFEB is the core block of the PFEM, which selectively fuses postevent multilayer features to provide discriminative postevent features. Due to the difference in the receptive field, low-layer features retain more complete details, such as localization information and clearer boundaries, but they suffer from background noise, whereas high-layer features have a clear background and semantic information. PFEB reduces feature variance by fusing high-level and low-level features that contain different information. As shown in Fig. 1, PFEB is a bi-input and bioutput structure. The postevent feature of the ith layer and the high-layer feature of the $i+1$ layer are used as input. The enhanced postevent feature and the high-layer feature of the ith layer are the output. An overview of PFEB can be expressed as follows:

$$(f_{\text{post}}^{i})',\ f_{\text{high}}^{i} = \text{PFEB}\left(f_{\text{post}}^{i},\ f_{\text{high}}^{i+1}\right) \tag{5}$$

where $i = \{4, 3, 2, 1\}$ represents the layer of network.

The feature f'_{post} output by each layer of PFEB contains more comprehensive information, but it also retains the differences in the features of each layer. The specific process of PFEB is as follows.

Step 1: The consistent part of the multilayer features of the bi-input is obtained by elementwise multiplication, denoted as F_{con}. This process can be expressed as follows:

$$F_{\text{con}} = F_{\text{post}} \otimes F_{\text{high}} \tag{6}$$

where $\otimes$ represents elementwise multiplication.

Step 2: The Squeeze-and-Excitation network (SE-Net) [38] is used for the "feature recalibration" of F_{con}. As shown in Fig. 2, first, the Squeeze uses the global average pooling (GAP) to compress the spatial dimension in each channel into a global feature constant. Second, the Excitation is used to capture the dependencies between feature channels to generate weights for each feature channel, which is realized through layers

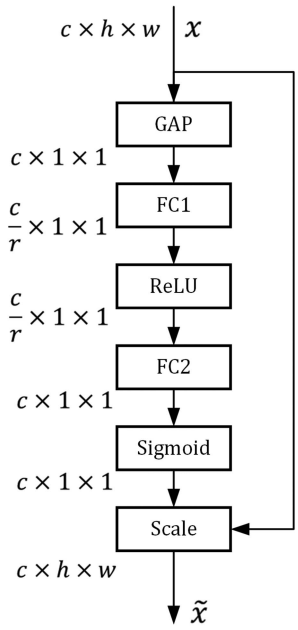

Fig. 2. Structure of SE-Net.

FC + Relu + FC + Sigmoid, that is

$$\text{Weigths} = \sigma\left(\text{FC}_2\left(\text{Relu}\left(\text{FC}_1\left(F_{\text{con}}\right)\right)\right)\right) \tag{7}$$

where $\sigma(\cdot)$ represents a Sigmoid activation function. FC_1 and FC_2 represent fully connected layers with dimensions c/r and c, respectively. c represents the feature dimension of F_{con}, and r is the dimension reduction coefficient, generally $r = 16$.

Finally, the Scale weights feature channel-by-channel by multiplication to complete the "feature recalibration" in the channel dimension. Scale makes the model more discriminative to the characteristics of each channel, which is similar to the attention mechanism. This process can be expressed as follows:

$$F'_{\text{con}} = \text{Weigths} \circledast F_{\text{con}} \tag{8}$$

where $\circledast$ represents dot product by channel.

Step 3: The recalibrated consistent features are applied to enhance saliency cues in F_{post} and F_{high}, thereby yielding fused features

$$F'_{\text{post}} = \text{Relu}\left(\text{BN}\left(\text{Conv}\left(F'_{\text{con}} \oplus F_{\text{post}}\right)\right)\right) \tag{9}$$

$$F_{\text{high}} = \text{Relu}\left(\text{BN}\left(\text{Conv}\left(F'_{\text{con}} \oplus F_{\text{high}}\right)\right)\right) \tag{10}$$

where $\oplus$ represents elementwise addition, and Conv, BN, and Relu represent *Conv3×3_str1*, batch normalization, and rectified linear unit in the common deep learning network, respectively.

Compared with methods using direct addition or concatenation to fuse the features of different layers, the proposed PFEB can remove the information generated during the fusion process that may muddy the original features. Through upward propagation, F_{post} continuously learns useful information from F_{high} so that the enhanced F'_{post} contains rich information, such as clear boundaries, accurate localization, and rich semantics.

Stage 2: The f^1_{high} output of the *layer1* PFEB is relatively complete, and the FR is a mechanism that downsamples the f^1_{high} to the same dimension as each layer in F'_{post} to further correct and refine F'_{post}. This process can be expressed as follows:

$$F'_{\text{post}} = F'_{\text{post}} \oplus f^1_{\text{high}}. \tag{11}$$

Stage 3: Additional features can be used to further enhance F'_{post}. The first kind of additional information is DEM; because of the independence between DEM features and optical image features, concatenation is directly used to combine them. This process can be expressed as follows:

$$F'_{\text{post}} = F'_{\text{post}} \copyright F_{\text{dem}} \tag{12}$$

where $\copyright$ represents the concatenation of two feature matrices.

The second one is the labeling of some landslide segmentation results on the postevent image. This network can upsample the output f^1_{high} to the original image size as a rough landslide segmentation result. Labeled landslides are used to side-supervise the corresponding image regions and update model parameters by using back-propagation.

D. Bifeature Difference Enhancement Module

The architecture of the BFDEM is shown in Fig. 1(c). The BFDEM uses its core block BFDEB as the backbone of the decoder to obtain change detection results by continuously upsampling and upward propagation. The structure of BFDEB is shown in Fig. 1; it generates a differentiated feature map by enhancing the difference between pre- and postevent features. The process of BFDEB is as follows.

First, the difference feature matrix obtained by elementwise subtraction is denoted as F_{diff}. This process can be represented as follows:

$$F_{\text{diff}} = F'_{\text{post}} \ominus F_{\text{pre}} \tag{13}$$

where $\ominus$ represents elementwise subtraction.

Second, similar to the PFEB, elementwise multiplication is used to get the consistent part of F_{diff} and F_{high} as follows:

$$F_{\text{diff-con}} = F_{\text{diff}} \otimes F_{\text{high}}. \tag{14}$$

Third, similar to (7) and (8), SE-Net is used to achieve "feature recalibration" to get $F'_{\text{diff-con}}$. The calculation process will not be repeated.

Finally, $F'_{\text{diff-con}}$ is combined with the difference feature F_{diff} by an elementwise addition to obtain the feature-enhanced change detection features. This process can be expressed as follows:

$$F_{dout} = \text{Relu}\left(\text{BN}\left(\text{Conv}\left(F'_{\text{diff-con}} \oplus F_{\text{diff}}\right)\right)\right). \tag{15}$$

Compared with methods using direct subtraction or concatenation to fuse pre- and postevent features, the designed BFDEB can highlight spectral, textural, and morphological differences between features to obtain more accurate results.

After BFDEB of *layer1*, *Conv3×3_str1* is used in f^1_{dout} to reduce its dimension to get the final change detection result. In addition, a hybrid loss function is applied to the supervised training process, as suggested in [39].

E. Flow Direction Calibration Module

FDCM uses D8 to calibrate the landslide flow direction as shown in Fig. 1(d). D8 is a single-flow-direction algorithm using DEM for flow direction analysis. The idea of the single-flow-direction algorithm is that the central grid has only one outflow grid, and all the "water" in the central grid is transferred to the outflow grid after the flow direction is determined. D8 assumes that the "water flow" in a single grid can only flow into the eight adjacent grids and use the steepest slope method to determine the direction of the flow.

Specifically, D8 calculates the elevation weight drop between the central grid and each adjacent grid on the 3×3 DEM grid and takes the grid with the largest elevation weight drop as the outflow grid of the central grid. On the final flow direction prediction map, we use eight different values to represent different flow directions, as shown in Fig. 3. D8 is a simple and effective method for the detection task of landslide flow direction.

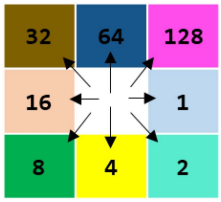

Fig. 3. D8 direction code. Different numbers and different colors represent different flow directions.

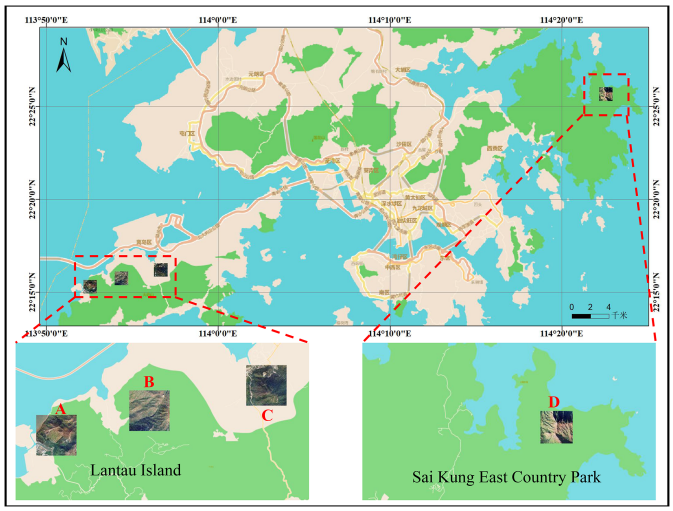

Fig. 4. Study area. Areas A–C are located in Lantau Island, Hong Kong, and Area D is located in the sharp park of Sai Kung East Country Park, Hong Kong.

III. Experiments

A. Study Area and Dataset

Under the influence of a torrential rainstorm on June 7, 2008, thousands of landslides with different sizes, shapes, and spatial distributions were caused in Hong Kong, causing huge losses to human life and property. The Hong Kong government attaches great importance to the prevention and postprocessing of such disasters [40]. Historical landslide data are recorded through the Enhanced Natural Terrain Landslide Inventory, which provides a wealth of prior knowledge for landslide identification. Our study area consists of four areas A–D, as shown in Fig. 4. Areas A–C are located in Lantau Island, the largest island in Hong Kong, and Area D is located in the sharp park of Sai Kung East Country Park. The pre-event and postevent aerial photos of areas A–D were collected by an aerial survey system equipped with a Zeiss RMK TOP 15 aerial survey camera at a flight altitude of 2400 m. The spatial resolution of the photos is 5 m, and the size is 2698 $\times$ 2698 pixels. At the same time, the postevent DEM data were also obtained from the relevant departments in Hong Kong. In Fig. 5, we illustrate study area A. A more detailed introduction to the study area can be found in [27].

Due to the limitations of the GPU and the need for continuous downsampling of the model, we cropped four large-format photos into several 256 $\times$ 256 pixel images with a 20% overlap rate. Flip, rotate, blur, and GridMask [41] are randomly adopted to enhance the data, and finally, 1296 sets of data were obtained. We use the following two data division methods.

1) Randomly divide the training set, validation set, and test set according to the ratio of 3:1:1. This division method is used for comparative experiments to verify the accuracy of the model.
2) Divide the data according to the A–C areas as the training set and the D area as the test set to verify the stability and generalization performance of the model.

B. Comparative Methods

To demonstrate the superiority of MFENet, six SOTA deep-learning-based change detection methods were selected for comparison, which included the three SOTA general change detection methods proposed on nonlandslide datasets and the three landslide-specific methods proposed on landslide datasets. These six methods are briefly introduced. Methods 1)–3) are proposed on nonlandslide data, and methods 4)–6) are proposed on landslide data.

1) *FC-Siam-diff [42]:* This network is proposed based on the U-Net architecture, which uses an encoder composed of a Siamese network to extract bitemporal features in parallel as the input to the decoder.
2) *DSIFN [43]:* This network uses pretrained vgg16 as an encoder, spatial and channel attention mechanisms in the feature fusion stage, and deep supervised learning to supervise feature fusion at different scales.
3) *SNUNet [44]:* This network is a densely connected Siamese network that maintains high-resolution and fine-grained features through dense skip connections. Meanwhile, the deep supervision module of the ensemble channel attention module is used to refine the most representative features of different semantic levels.
4) *FCN-PP [28]:* This network uses FCN as the encoder of the network in order to construct a U-shaped network. The network is improved by adding a pyramid pooling (PP) module at the bottom of the U-shaped network. PP consists of three convolutions of different sizes and different stride widths to obtain features in the different receptive fields.
5) *DP-FCN [29]:* This network uses Siamese FCNs to encode and decode pre- and postevent images, respectively, and fuses pre- and postevent features by concatenation in the intermediate stage of decoding.
6) *LanDCNN [31]:* This network also uses U-Net as the main structure, replacing the encoder with ResNet50. The DEM and the image are concatenated as the input of the network, which proves the effectiveness of the DEM as an additional feature.

C. Implementation Details and Evaluation Metrics

The proposed method is implemented based on Pytorch using python3.8 + cuda10.2. The Adam optimizer is used with an initial learning rate of 0.0001, and the training batch data size is set to 8. The optimizer of the comparison method is only mentioned by SNUNet in its paper and also uses the Adam optimizer; our optimizer is unified as Adam during training. The epoch is set to 100 in order to make the results obtained by all networks converge to their optimal values. All of our experiments are conducted on the NVIDIA GeForce RTX 3090 24 GB.

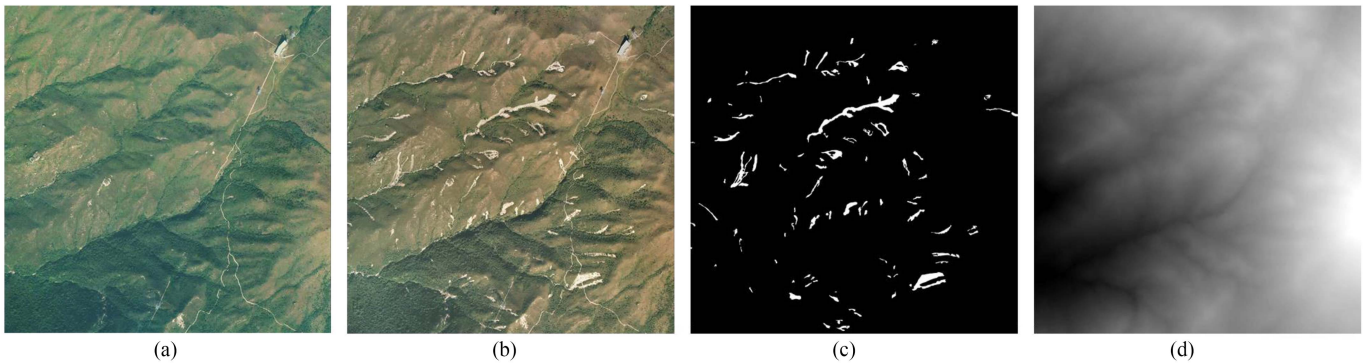

Fig. 5. Data example from study area A. (a) Pre-event image. (b) Postevent image. (c) Ground truth of landslide. (d) DEM.

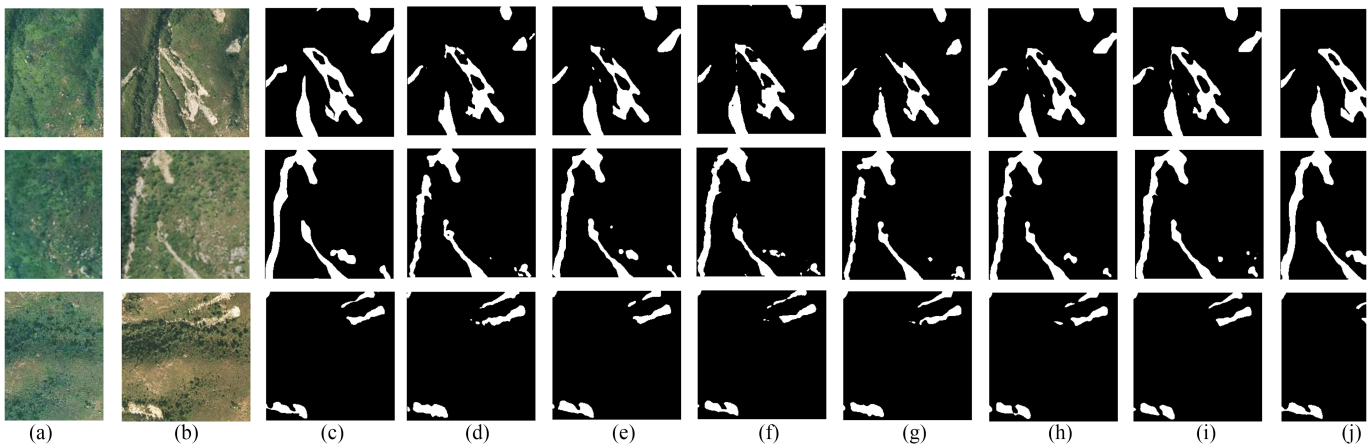

Fig. 6. Visualization results of comparisons. (a) Pre-event image. (b) Postevent image. (c) Ground truth. (d) FC-Siam-diff. (e) DSIFN. (f) SNUNet. (g) FCN-PP. (h) DP-FCN. (i) LanDCNN. (j) MFENet (ours). The detected changes are white pixels, and the unchanged parts are black pixels. It is worth noting that MFENet does not use additional features in this section.

To evaluate the performance of the proposed method, we used four evaluation metrics: Precision (*P*), Recall (*R*), F1 score (F1), and IoU. *P* represents the proportion of correctly detected changed pixels in the model-predicted changed pixels. The higher *P* indicates that fewer false pixels are detected. *R* represents the proportion of correctly detected changed pixels in the true value of changed pixels. The higher the *R*, the fewer missing pixels are detected. The F1 score can be regarded as the harmonic means of the model's *P* and *R*. IoU represents the ratio of the intersection and union between the predicted result of the change pixels and the real change pixels. They are calculated as follows:

$$P = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \times 100\% \tag{16}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100\% \tag{17}$$

$$F1 = \frac{2 \times P \times R}{P+R} \times 100\% \tag{18}$$

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}+\mathrm{FP}} \times 100\% \tag{19}$$

TABLE I
COMPARISON RESULTS

Methods	P (%)	R (%)	IoU (%)	F1 (%)	Params(M)
FC-Siam-diff	86.41	79.74	70.86	82.94	5
DSIFN	85.22	**94.65**	81.30	89.69	193
SNUNet	92.08	88.67	82.39	90.34	46
FCN-PP	89.27	84.86	77.01	87.01	475
DP-FCN	90.48	87.67	80.26	89.05	113
LanDCNN	91.21	87.90	81.04	89.53	143
MFENet	**92.87**	93.51	**87.23**	**93.19**	175

The bold numbers indicate the best results.

where TP represents the correct detection as changed pixels, FN and FP represent the missed and falsely detected changed pixels, and TN represents the correct detection as unchanged pixels.

D. Comparisons and Analysis

It is worth noting that MFENet does not use additional features in this section. According to the first data division

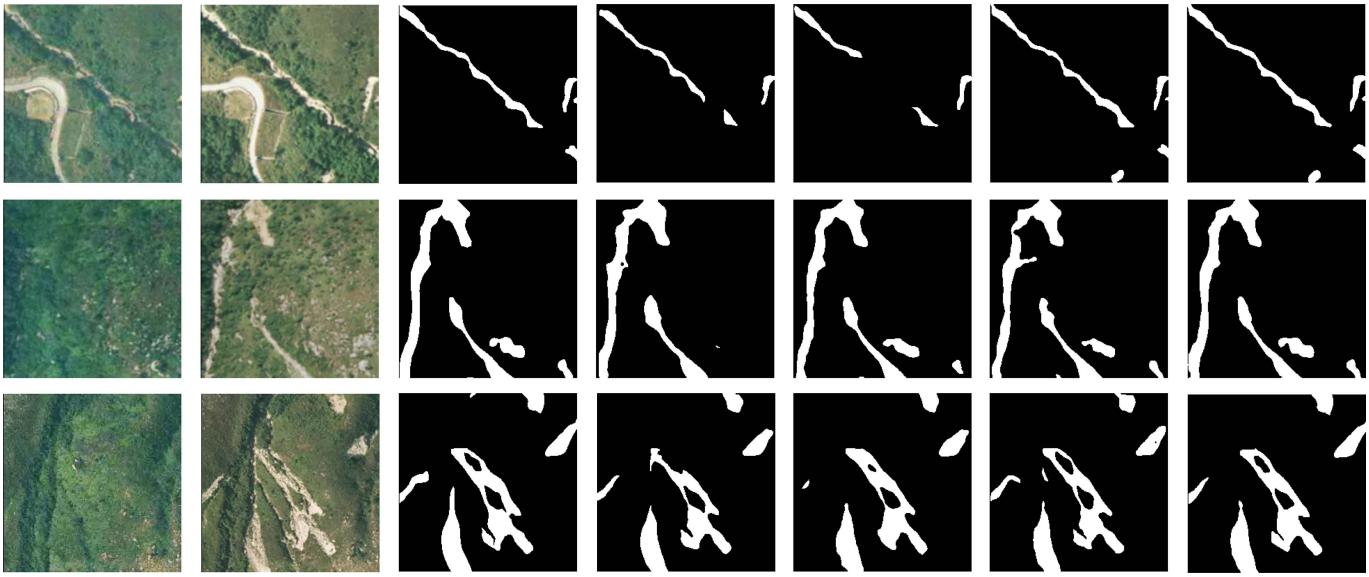

Fig. 7. Visualization results of ablation study. (a) Pre-event image. (b) Postevent image. (c) Ground truth. (d) BASE. (e) BASE+BFDEM. (f) BASE+PFEM. (g) MFENet. The detected changes are white pixels, and the unchanged parts are black pixels.

TABLE II
ABLATION STUDY

Methods	P (%)	R (%)	IoU (%)	F1 (%)
BASE	87.95	86.84	77.61	87.39
BASE+BFDEM	**94.43**	87.15	82.89	90.94
BASE+PFEM	92.73	91.66	85.51	92.19
MFENet	92.87	**93.51**	**87.23**	**93.19**

The bold numbers indicate the best results.

method, a comparative test is carried out to verify the accuracy of the model. The quantitative results for the *P*, *R*, F1, and IoU of all methods are summarized in Table I. The method proposed in this article achieves the best in *P*, IoU, and F1 and improves by 0.79%, 4.84%, and 2.85% compared with the suboptimal method. In addition, MFENet achieves the highest accuracy with a modest number of parameters (Params). This shows that the MFENet has a good balance between computational complexity and accuracy. To evaluate the performance more intuitively, details of the experimental results are presented in Fig. 6. It can be found that the proposed method improves the results mainly in the following aspects.

1) *Clearer boundaries:* The postevent features enhanced by PFEM have a more complete boundary, which is preserved when the landslide change detection results are obtained.
2) *Consistent internal logic:* The detected landslides have fewer internal voids or partial truncations because the model preserves more detailed features of ground objects through feature enhancement.
3) *Less "salt and pepper:"* This is a spurious change consisting of a small number of pixels on the predicted map. The "salt and pepper" is the pollution of low-layer features by background noise and the loss of some semantic information between pixels during decoding. The proposed method fuses high-layer and low-layer features through PFEM and BFDEM, which preserves rich semantic information and suppresses background noise on the final feature map.
4) *Small targets can be caught:* Through the feature-enhancement modules PFEM and BFDEM, the major features are enhanced and the useless features are suppressed so that many small tributaries of landslides are not ignored by the network.

E. Ablation Study

To evaluate the effectiveness of PFEM and BFDEM, the following four ablation experiments are set up to further examine the modules and network structures.

1) Both PFEB and BFDEB of MFENet are replaced by the concatenated features and the *Conv3×3_str1* to achieve two feature fusions in the network (BASE).
2) PFEB of MFENet is replaced by the concatenated features and the *Conv3×3_str1* (BASE+BFDEM).
3) The BFDEB of MFENet is replaced by the concatenated features and the *Conv3×3_str1* (BASE+PFEM).
4) MFENet.

The experimental results are shown in Table II and Fig. 7. BFDEM and PFEM have significant effects regarding the respective improvement of the *P* and *R* of the model. Moreover, the overall performance of PFEM is higher than that of BFDEM, which shows that PFEM specially designed for this type of postevent change of landslide is very effective. The MFENet model, which combines PFEM and BFDEM, achieves the best performance, indicating that there is good compatibility between the two modules. Thus, and in summary, MFENet is able to be effectively applied for LM.

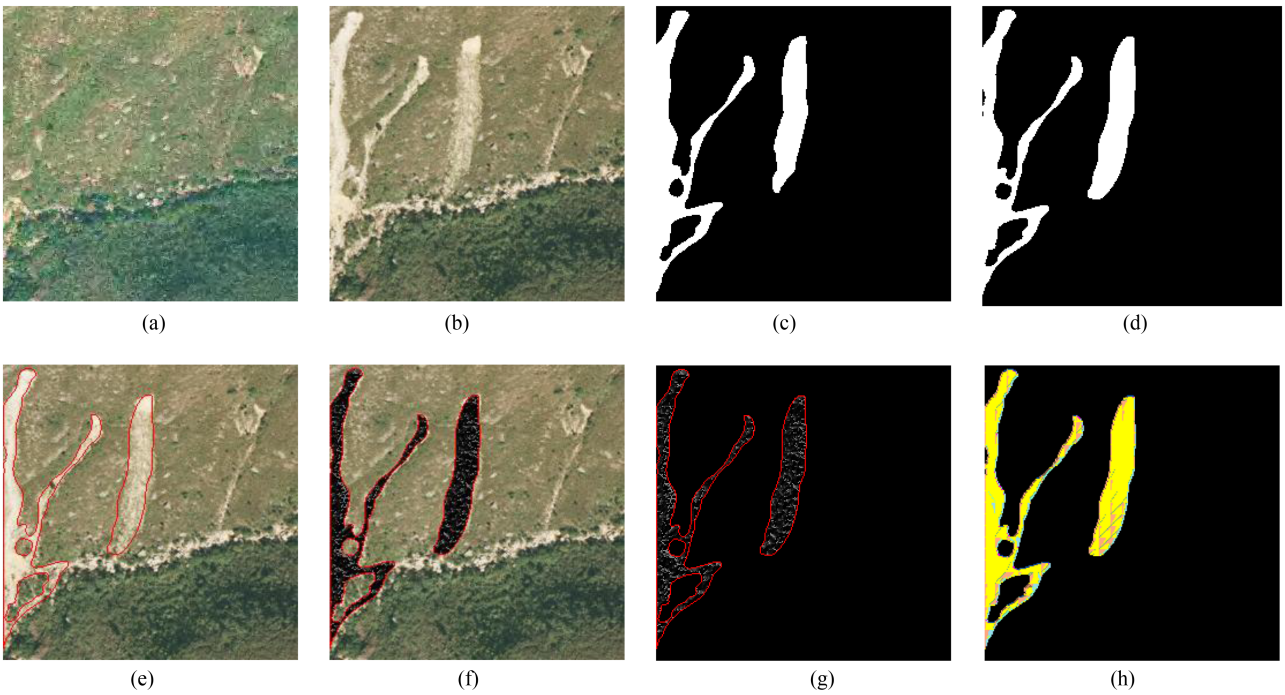

Fig. 8. Visualization results of LM. (a) Pre-event image. (b) Postevent image. (c) Ground truth. (d) Landslide predicted mask. (e) Prediction boundary on the postevent image. (f) Predicted boundary and predicted flow direction on the postevent image. (g) Predicted boundary and predicted flow direction on the predicted mask. (h) RGB display of predicted flow direction on the predicted mask.

F. Flow Direction

First, the landslide boundary is calibrated on the postevent image using the landslide detection results, which provide information for calculating the perimeter and area of the landslide. Similarly, the landslide boundaries are demarcated on the DEM. Then, the flow direction of each pixel of the landslide is calibrated by using the D8 method combined with the DEM and the landslide boundary, and the flow direction map of the landslide is obtained. Finally, the flow-to-grayscale image is converted to an RGB image for easy visualization. Fig. 8 visualizes the final LM results.

IV. Discussion

In this section, the availability of additional features and the generalization performance of the model are experimentally verified and discussed.

A. Availability of Additional Features

DEM and some postevent landslide segmentation results are input into the network as additional features. Using MFENet as the base, three sets of comparative experiments were set up: only additional input of DEM (MFENet_DEM), only additional input of segmentation results (MFENet_Seg), and additional input of DEM and segmentation results simultaneously (MFENet_DEM&Seg). The experimental results are shown in Table III. Inputting additional features can significantly improve the accuracy of the model. When two additional features are input simultaneously, *R*, IoU, and F1 are all optimal, but the *P* value is slightly lower than the input alone, which may be caused by a mutual disturbance between features.

TABLE III
Availability of Additional Features

Methods	P (%)	R (%)	IoU (%)	F1 (%)
MFENet	92.87	93.51	87.23	93.19
MFENet_DEM	**94.73**	94.67	89.93	94.70
MFENet_Seg	94.43	94.50	89.51	94.46
MFENet_DEM&Seg	94.29	**95.87**	**90.61**	**95.07**

The bold numbers indicate the best results.

The visualization of the results is shown in Fig. 9. As a commonly used data source for landslide detection or postprocessing, the effect of adding DEM to the network is significant. It can be seen from Fig. 9 that the DEM effectively eliminates the internal cavity of the landslide caused by disturbance factors and the false changes that may be identified as landslides on the spectrum. Segmentation results increase the supervision information of the network, which is equivalent to expanding the training data so that the network can learn more effective landslide spectral features to improve detection accuracy.

B. Generalization Performance

According to the second data division method, the generalization performance of the model is verified by comparative experiments. The experimental metric results are shown in Table IV. MFENet still outperforms other networks on *R*, IoU, and F1. The visualization of the results is shown in Fig. 10. Combining

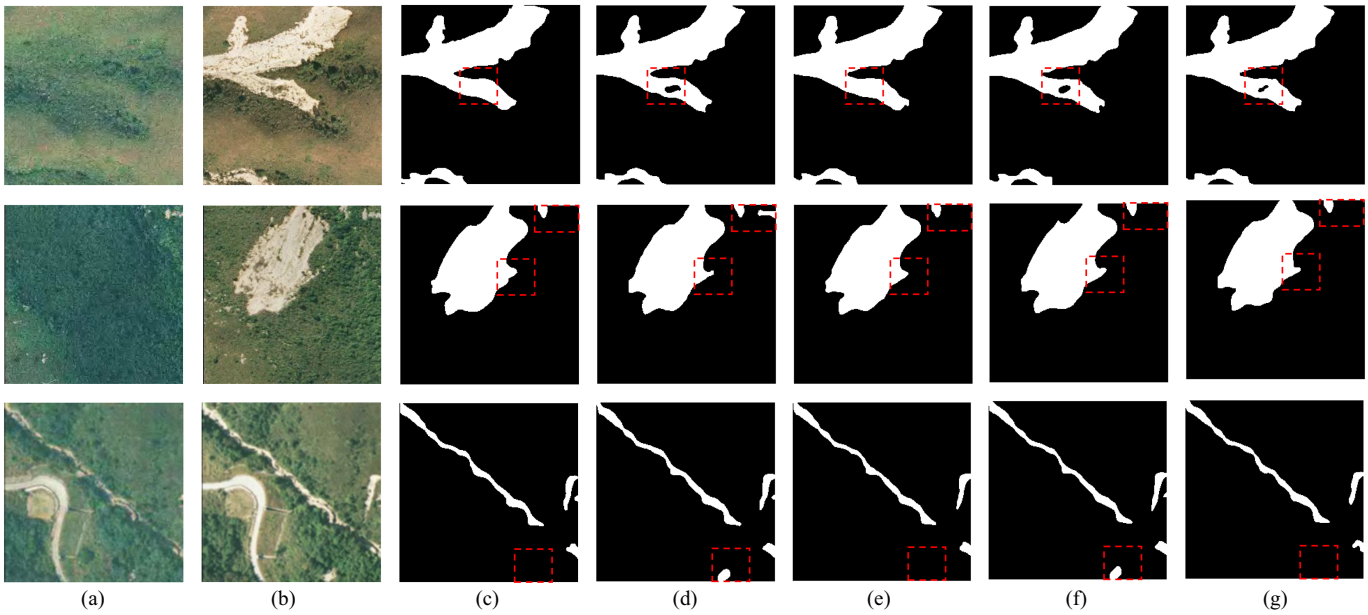

Fig. 9. Visualization results of availability of additional features. (a) Pre-event image. (b) Postevent image. (c) Ground truth. (d) MFENet. (e) MFENet_DEM. (f) MFENet_Seg. (g) MFENet_DEM&Seg. The detected changes are white pixels, and the unchanged parts are black pixels.

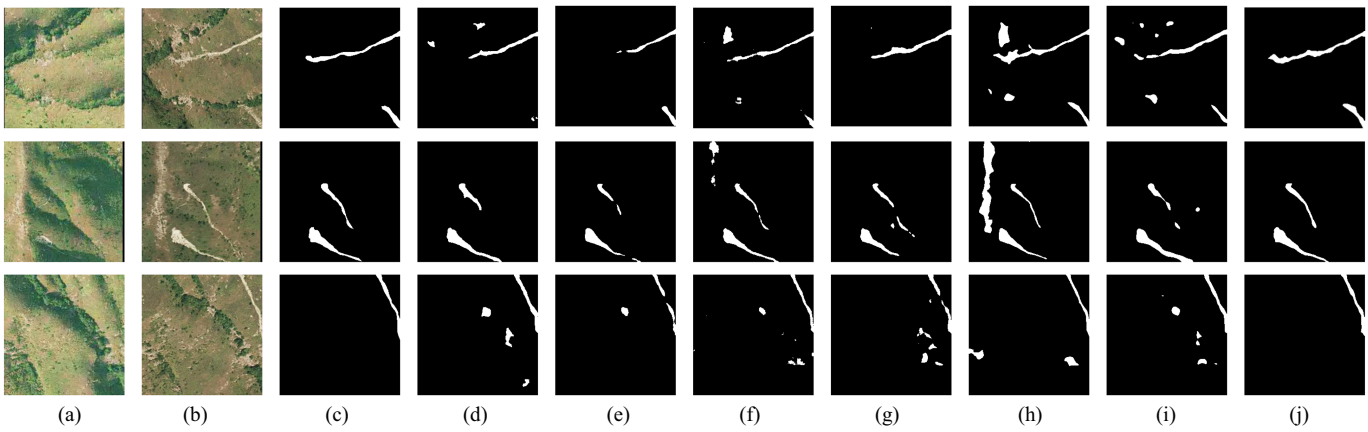

Fig. 10. Visualization results of generalization performance. (a) Pre-event image. (b) Postevent image. (c) Ground truth. (d) FC-Siam-diff. (e) DSIFN. (f) SNUNet. (g) FCN-PP. (h) DP-FCN. (i) LanDCNN. (j) MFENet (ours). The detected changes are white pixels, and the unchanged parts are black pixels.

metrics analysis and visualization results, it is found that the main reason for poor generalization performance is reflected in the following two aspects: conservative predictions lead to a large number of missed pixels, and aggressive predictions lead to a large number of false pixels. MFENet is neither too conservative to ensure that landslides can be completely detected, nor it is too aggressive to detect many false landslides. MFENet maintains high precision while achieving high recall. However, it can be seen that the generalization performance of MFENet is better than that of other models. It is also worth mentioning that DSIFN obtained the largest P or R in the two data division methods, respectively, showing that conservative and aggressive may not be fixed characteristics of the model and may be related to the dataset.

However, when compared with the first data partitioning method, all models show a significant decrease in accuracy. This appears to be due to the similarity of the geographical environments of the A–C areas, which leads to similar types of landslides afterward. The models only learn the features of landslides in this specific geographical environment, and the decrease in model accuracy is predictable. If a fully supervised model is to have better generalization performance, model training in different scenarios is essential.

V. Conclusion

In this article, an end-to-end change detection network based on multilevel feature enhancement is proposed for landslide

TABLE IV
GENERALIZATION PERFORMANCE

Methods	P (%)	R (%)	IoU (%)	F1 (%)
FC-Siam-diff	85.64	55.28	50.59	67.19
DSIFN	**87.48**	68.46	62.35	76.81
SNUNet	81.62	63.82	55.80	71.63
FCN-PP	83.01	42.24	38.88	55.99
DP-FCN	61.27	77.20	51.88	68.32
LanDCNN	76.23	60.99	51.24	67.76
MFENet (ours)	83.19	**81.95**	**70.30**	**82.56**

The bold numbers indicate the best results.

detection. The postevent feature recalibration and bitemporal difference feature recalibration are completed by means of two feature-enhancement modules, PFEM and BFDEM. Experiments show that MFENet outperforms both the SOTA general change detection methods and landslide-specific change detection methods for LM. The landslides detected by MFENet maintain clear boundaries and internal logical consistency and also reduce false detections and missed detections caused by disturbance factors. Finally, the landslide flow direction is calibrated by using the D8 method combined with the landslide detection results and DEM to complete the LM. On the basis of MFENet, additional features are input through side supervision and concatenation, which further improves the accuracy of the network to detect landslides. In addition, it is experimentally demonstrated that the generalization performance of MFENet outperforms that of other methods. The LM results of this study are shown to have high accuracy, less manual intervention, and richer landslide information, capable of being used for landslide sensitivity analysis, and have great significance as regards the success of the following: postdisaster rescue, disaster prevention, and mitigation. In the future, it will be worthwhile for there to be further focus on the development of LM networks with stronger generalization performance, enabled by combining weakly supervised learning or transfer learning with the current limited landslide remote sensing data, worldwide.

ACKNOWLEDGMENT

The authors would like to thank editors and anonymous reviewers for helpful comments on this article and those who provided the base model. They would also like to thank the Geotechnical Engineering Office (GEO) of the Hong Kong Civil Engineering and Development Department for providing aerial orthophotographs.

REFERENCES

[1] M. J. Froude and D. N. Petley, "Global fatal landslide occurrence from 2004 to 2016," *Natural Hazards Earth Syst. Sci.*, vol. 18, no. 8, pp. 2161–2181, 2018.

[2] G. Metternicht, L. Hurni, and R. Gogu, "Remote sensing of landslides: An analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments," *Remote Sens. Environ.*, vol. 98, no. 2/3, pp. 284–303, 2005.

[3] U. Haque et al., "The human cost of global warming: Deadly landslides and their triggers (1995–2014)," *Sci. Total Environ.*, vol. 682, pp. 673–684, 2019.

[4] S. L. Gariano and F. Guzzetti, "Landslides in a changing climate," *Earth-Sci. Rev.*, vol. 162, pp. 227–252, 2016.

[5] A. Perrone, V. Lapenna, and S. Piscitelli, "Electrical resistivity tomography technique for landslide investigation: A review," *Earth-Sci. Rev.*, vol. 135, pp. 65–82, 2014.

[6] C. Zhong et al., "Landslide mapping with remote sensing: Challenges and opportunities," *Int. J. Remote Sens.*, vol. 41, no. 4, pp. 1555–1581, 2020.

[7] C. Zhao and Z. Lu, "Remote sensing of landslides–A review," *Remote Sens.*, vol. 10, no. 2, 2018, Art. no. 279.

[8] Z. Lv, F. Wang, W. Sun, Z. You, N. Falco, and J. A. Benediktsson, "Landslide inventory mapping on VHR images via adaptive region shape similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Sep. 2022, Art. no. 5630211.

[9] W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan, "Change detection based on artificial intelligence: State-of-the-art and challenges," *Remote Sens.*, vol. 12, no. 10, 2020, Art. no. 1688.

[10] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[11] A. Mohan, A. K. Singh, B. Kumar, and R. Dwivedi, "Review on remote sensing methods for landslide detection using machine and deep learning," *Trans. Emerg. Telecommun. Technol.*, vol. 32, 2021, Art. no. 7, doi: 10.1002/ett.3998.

[12] B. Yu and F. Chen, "A new technique for landslide mapping from a large-scale remote sensed image: A case study of central Nepal," *Comput. Geosci.*, vol. 100, pp. 115–124, 2017.

[13] G. Cheng, L. Guo, T. Zhao, J. Han, H. Li, and J. Fang, "Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA," *Int. J. Remote Sens.*, vol. 34, no. 1, pp. 45–59, 2013.

[14] H. Yu, Y. Ma, L. Wang, Y. Zhai, and X. Wang, "A landslide intelligent detection method based on CNN and RSG_R," in *Proc. IEEE Int. Conf. Mechatron. Automat.*, 2017, pp. 40–44.

[15] W. Qi, M. Wei, W. Yang, C. Xu, and C. Ma, "Automatic mapping of landslides by the ResU-Net," *Remote Sens.*, vol. 12, no. 15, 2020, Art. no. 2487.

[16] T. R. Martha, N. Kerle, V. Jetten, C. J. van Westen, and K. V. Kumar, "Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods," *Geomorphology*, vol. 116, no. 1/2, pp. 24–36, 2010.

[17] F. Guzzetti, A. C. Mondini, M. Cardinali, F. Fiorucci, M. Santangelo, and K.-T. Chang, "Landslide inventory maps: New tools for an old problem," *Earth-Sci. Rev.*, vol. 112, no. 1/2, pp. 42–66, 2012.

[18] D. J. Hayes and S. A. Sader, "Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series," *Photogrammetric Eng. Remote Sens.*, vol. 67, no. 9, pp. 1067–1075, 2001.

[19] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and *k*-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[20] R. D. Johnson and E. S. Kasischke, "Change vector analysis: A technique for the multispectral monitoring of land cover and condition," *Int. J. Remote Sens.*, vol. 19, no. 3, pp. 411–426, 1998.

[21] A. C. 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.

[22] 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.

[23] P. Lu, Y. Qin, Z. Li, A. C. Mondini, and N. Casagli, "Landslide mapping from multi-sensor data through improved change detection-based Markov random field," *Remote Sens. Environ.*, vol. 231, 2019, Art. no. 111235.

[24] B. Adriano, N. Yokoya, H. Miura, M. Matsuoka, and S. Koshimura, "A semiautomatic pixel-object method for detecting landslides using multitemporal ALOS-2 intensity images," *Remote Sens.*, vol. 12, no. 3, 2020, Art. no. 561.

[25] Y. Aimaiti, W. Liu, F. Yamazaki, and Y. Maruyama, "Earthquake-induced landslide mapping for the 2018 Hokkaido Eastern Iburi earthquake using PALSAR-2 data," *Remote Sens.*, vol. 11, no. 20, 2019, Art. no. 2351.

[26] F. S. Tehrani, M. Calvello, Z. Liu, L. Zhang, and S. Lacasse, "Machine learning and landslide studies: Recent advances and applications," *Natural Hazards*, vol. 9, pp. 1–49, 2022.

[27] W. Shi, M. Zhang, H. Ke, X. Fang, Z. Zhan, and S. Chen, "Landslide recognition by deep convolutional neural network and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 4654–4672, Jun. 2020.
[28] 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.
[29] Z. Lv, T. Liu, X. Kong, C. Shi, and J. A. Benediktsson, "Landslide inventory mapping with bitemporal aerial remote sensing images based on the dual-path fully convolutional network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4575–4584, Mar. 2020.
[30] B. Fang, G. Chen, L. Pan, R. Kou, and L. Wang, "GAN-based Siamese framework for landslide inventory mapping using bi-temporal optical remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 391–395, Mar. 2020.
[31] Z. Su, J. K. Chow, P. S. Tan, J. Wu, Y. K. Ho, and Y.-H. Wang, "Deep convolutional neural network-based pixel-wise landslide inventory mapping," *Landslides*, vol. 18, no. 4, pp. 1421–1443, 2021.
[32] S. O. Y. Amankwah et al., "Landslide detection from bitemporal satellite imagery using attention-based deep neural networks," *Landslides*, vol. 19, no. 10, pp. 2459–2471, 2022.
[33] J. Fairfield and P. Leymarie, "Drainage networks from grid digital elevation models," *Water Resour. Res.*, vol. 27, no. 5, pp. 709–717, 1991.
[34] P. - T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Operations Res.*, vol. 134, no. 1, pp. 19–67, 2005.
[35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[36] G. Máttyus, W. Luo, and R. Urtasun, "DeepRoadMapper: Extracting road topology from aerial images," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 3438–3446.
[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
[38] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
[39] X. Qin, Z. Zhang, C. Huang, C. Gao, M. Dehghan, and M. Jagersand, "Basnet: Boundary-aware salient object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 7479–7489.
[40] F. W. Y. Ko and F. L. C. Lo, "From landslide susceptibility to landslide frequency: A territory-wide study in Hong Kong," *Eng. Geol.*, vol. 242, pp. 12–22, 2018.
[41] P. Chen, S. Liu, H. Zhao, and J. Jia, "Gridmask data augmentation," 2020, *arXiv2001.04086*.
[42] R. C. Daudt, B. L. Saux, and A. Boulch, "Fully convolutional Siamese networks for change detection," in *Proc. IEEE 25th Int. Conf. Image Process.*, 2018, pp. 4063–4067.
[43] C. Zhang et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 166, pp. 183–200, 2020.
[44] S. Fang, K. Li, J. Shao, and Z. Li, "SNUNet-CD: A densely connected Siamese network for change detection of VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Feb. 2021, Art. no. 8007805.

Lukang Wang received the B.S. degree in surveying engineering from the China University of Mining and Technology, Xuzhou, China, in 2018, where he is currently working toward the Ph.D. degree in geodesy and surveying engineering at the School of Environment and Spatial Informatics.

His research interests include change detection, landslide mapping, and deep learning for remote sensing.

Min Zhang received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2020.

He is currently a Research Assistant Professor with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong. His research interests include spatial data quality, deep learning, artificial intelligence, change detection, and object recognition in remote sensing.

Xiaoqi Shen received the B.S. degree in surveying engineering from the China University of Mining and Technology, Xuzhou, China, in 2017, where he is currently working toward the Ph.D. degree in the School of Environment and Spatial Informatics.

His research interests include human mobility analysis, geographic data mining, and geospatial artificial intelligence.

Wenzhong Shi received the Ph.D. degree in geomatics from the University of Osnabrück, Vechta, Germany, in 1994.

He is currently the Director of PolyU-Shenzhen Technology and Innovation Research Institute (Futian), the Director of Otto Poon Charitable Foundation Smart Cities Research Institute, PolyU, Hong Kong, the Chair Professor in geographic information science (GISci) and remote sensing, and the Director of the Joint Research Laboratory on Spatial Information of PolyU and Wuhan University, Wuhan, China. He has authored/coauthored more than 300 research articles in journals indexed by SCI and 20 books. His research interests include urban informatics for smart cities, GISci and remote sensing, specifically, artificial-intelligence-based object recognition and change detection from satellite imagery, intelligent analytics and quality control for spatial data, mobile mapping and 3-D modeling based on light detection and ranging and imagery, and 3-D geographic information system models.

Dr. Shi is an Academician of the International Eurasian Academy of Sciences and a Fellow of Academy of Social Sciences (U.K.).