Landslide Detection of Hyperspectral Remote Sensing Data Based on Deep Learning With Constrains

Chengming Ye, Yao Li, Peng Cui, Li Liang, Saeid Pirasteh, José Marcato, Jr., *Member, IEEE*, Wesley Nunes Gonçalves, *Member, IEEE*, and Jonathan Li, *Senior Member, IEEE*

***Abstract*—Detecting and monitoring landslides are hot topics in remote sensing community, particularly with the development of remote sensing technologies and the significant progress of computer vision. To the best of our knowledge, no study focused on deep learning-based methods for landslide detection on hyperspectral images. We proposes a deep learning framework with constraints to detect landslides on hyperspectral image. The framework consists of two steps. First, a deep belief network is employed to extract the spectral–spatial features of a landslide. Second, we insert the high-level features and constraints into a logistic regression classifier for verifying the landslide. Experimental results demonstrated that the framework can achieve higher overall accuracy when compared to traditional hyperspectral image classification methods. The precision of the landslide detection on the whole image, obtained by the proposed method, can reach 97.91%, whereas the precision of the linear support vector machine, spectral information divergence, and spectral angle match are 94.36%, 84.50%, and 86.44%, respectively. Also, this article reveals that the high-level feature extraction system has a significant potential for landslide detection, especially in multi-source remote sensing.**

***Index Terms*—Deep belief network (DBN), deep learning, feature extraction, hyperspectral data, landslide.**

Manuscript received December 8, 2018; revised October 3, 2019 and October 23, 2019; accepted October 26, 2019. Date of publication November 24, 2019; date of current version February 4, 2020. This work was supported in part by the Strategic Priority Research Program of the Chinese Academy of Sciences [XDA23090203], in part by the Key Research Program of Frontier Sciences, CAS Grant QYZDY-SSW-DQC006, and in part by the Key Program of Sichuan science and technology department under Grant 2018SZ0350. (*Corresponding author: Yao Li.*)

C. Ye is with the Key Laboratory of Earth Exploration and Information Technology of Ministry of Education, Chengdu University of Technology, Chengdu 610059, China (e-mail: rsgis@sina.com).

Y. Li and P. Cui are with the Key Laboratory of Mountain Hazards and Earth Surface Process, Chinese Academy of Sciences, Chengdu 610041, China, with the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China, and also with the University of the Chinese Academy of Sciences, Beijing 100049, China (e-mail: yaolicd@hotmail.com; pengcui@imde.ac.cn).

L. Liang is with the College of Management Science, Chengdu University of Technology, Chengdu 610059, China (e-mail: liangl@cdut.edu.cn).

S. Pirasteh is with the Faculty of Geosciences & Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China (e-mail: s2pirast@uwaterloo.ca).

J. Marcato, Jr. is with the Faculty of Engineering, Architecture and Urbanism, and Geography, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Brazil (e-mail: jose.marcato@ufms.br).

W. N. Gonçalves is with the Faculty of Computer Science and the Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Brazil (e-mail: wesley.goncalves@ufms.br).

J. Li is with the Departments of Geography and Environmental Management and Systems Design Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada (e-mail: junli@uwaterloo.ca).

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

Digital Object Identifier 10.1109/JSTARS.2019.2951725

I. Introduction

Because of casualties and loss of goods, landslides seriously affect the social and economic order [1]–[4]. In China, 7403 landslides occurred in 2016 (an ever-increasing number) with 405 people killed or missing and 209 injured [5]. With complex geological conditions and human activities (including deforestation, mining of minerals, and intensive exploitation of land for construction), landslides were easily induced by extreme natural events in large-scale area. Quickly and accurately extracting landslide information can increase the efficiency of disaster mitigation, especially in response to emergency situations [3], [6]–[9]. Nevertheless, because of the risks in a field survey and the vastness of a disaster area, it is impossible to extract landslide information by means of a man-made investigation requiring a large number of human and financial resources [10]–[13]. Therefore, remote sensing, with its characteristics of macro-scale, rapidity, and noncontact detection, is widely used to landslide mapping [14]–[16]. Over the last three decades, landslide detection and mapping by remote sensing have been categorized into three general classes.

1) Analysis of the image features of a landslide with optical image data (including space-borne and air-borne remote sensing data), to recognize the extent and location of the landslide by visual interpretation or by automatic extraction method [17]–[21].
2) Detection of surface deformation and deposition resulted from landslides using radar data (such as synthetic aperture radar (SAR), InSAR, LiDAR, DInSAR) [22]–[26].
3) Mapping of landslides combining radar and optical image data [27]–[30]. Significant progress has been made in radar remote sensing [31]–[34].

On the other hand, because of the less bands leading to shortage of image features in optical remote sensing, automatic interpretation of landslides fails to provide high detection accuracy [35], [36]. To promote the accurate mapping of landslides, we applied hyperspectral remote sensing to detect a weak spectral change after a landslide.

By combining imaging and spectroscopy technology, hyperspectral remote sensing provides spatially and spectrally continuous data for earth objects [37]. Therefore, hyperspectral imaging is widely applied in many fields such as precision agriculture, mineralogy, environmental science, and forestry [38]–[41]. A fundamental aspect of research in these applications is the classification of objects in an image. Many conventional supervised machine learning methods (such as *k*-nearest neighbors, spectral angle mapper, target detection, spectral un-mixing, neural networks and multiple kernel learning) have been introduced to improve accuracy of classification [42]–[49]. However, Hughes phenomenon effect, caused by the high-dimensional datasets and the limited training samples, restrict the classification of hyperspectral data [50]. Support vector machine (SVM) which can solve the problem of classifying high-dimensional data was introduced to improve hyperspectral image classification [51]. In addition, derived algorithms based on SVM, which consider spatial information, composite kernels, or active learning, outperform it [52], [53]. Therefore, SVM has been a long-time state-of-the-art method. As the other traditional machine learning algorithms, SVM is a single-layer classifier, which interprets hyperspectral image data by extracting the targets shallow features, but ignores the invariant deep features of the data [54].

Previous study focused in the detection of landslide in multispectral images using machine learning and deep learning methods [55].

To the best of our knowledge, no study focused thus far on deep learning-based methods for landslide detection on hyperspectral images. In this study, with hyperspectral remote sensing data, a deep learning framework with constraints (DLWC) is used to detect a landslide. Possessing the ability to extract the deep features of complex data, a deep belief network (DBN) obtains the spectral–spatial features of the data, which play an important role in landslide detection. [57]. Then, by logistic regression (LR), the extracted image features and imposed constraints [including the digital elevation model (DEM), fault zone, earthquake, soil, river, road, rainfall, and vegetation coverage] are combined input to improve the detection result.

II. Deep Learning, RBM, DBN, and Landslide Detection Frameworks

A. Advantages of Deep Learning

A key procedure in the early stage of image classification is feature extraction (FE), whose performance greatly affects classification precision [57]. Consequently, to improve the efficiency of classification, an immense amount of effort has been made to extract features from an image. However, the conventional option is to design model architecture exclusively based on engineering skills and domain expertise, which could be easily affected by artificial factors [58]. And the features extracted by the "shallow" machine learning method, which has only one nonlinear feature transformation, are hard to apply to reveal the intricate structures of large datasets [59].

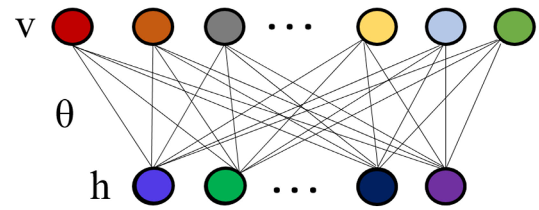

Fig. 1. Illustration of RBM. h, θ, v represent hidden units, parameter, and visible units, respectively.

As a subfield of machine learning, deep learning with more than two mapping layers has gained wide attention, because it can hierarchically extract features from an original dataset. By learning high-level semantic features from low-level visual features layer by layer, the deep learning architecture yields more abstract and useful representations which have fewer relationships with domain expertise [60]. In processing of images classification, the higher layers of feature representation expand the differences between categories, thereby improving classification accuracy [61].

Basic deep learning models include DBN, stacked auto-encoder (SAE), and deep convolutional neural networks (CNNs) [62]–[68]. Many derivative models have been proposed to improve the performance in various applications. For example, de-noising auto-encoder and contractive auto-encoder can learn robust and useful representations for a dataset [69], which improve the SAE by adding some restrictions. On the other hand, many modified CNNs are applied to image recognition. Among these modified architectures, AlexNet is regarded as the beginning of the deep convolution network, which famously won the ILSVRC-2012 competition [70], [71]. Then, many research teams also structure deep convolution network models that mainly include Network in Network, GoogLeNet, ResNet [72], [73]. Although the CNN nets previously mentioned have excellent performance in image recognition, they require massive training data to trigger their powers.

DBN overcomes overfitting resulted from small-sized samples, especially in the classification of hyperspectral remote sensing [74]. In this study, a framework based on DBN is applied to detect landslides caused by the earthquake.

B. Restricted Boltzmann Machine (RBM)

The restricted Boltzmann machine (RBM) is a basic composition of DBN, adapted from the Boltzmann machine. As shown in Fig. 1, the standard type of RBM has two units, hidden and visible (h, v). With a sufficient number of hidden units, any kinds of discrete distributions can be simulated.

Thus, RBM is used heavily to extract data features. A joint configuration of the units has an energy given by the following equation:

$$E\left(v, h; \theta\right) = -\sum_{i=1}^{n} a_i.v_i - \sum_{j=1}^{m} b_j.h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}.v_i.h_j \quad (1)$$

where $\theta = \{a_i, b_j, w_{ij}\}$ is a model parameter; $v_i = \{1 \text{ or } 2\}$ and $h_j = \{1 \text{ or } 2\}$ represent the states of visible unit, i, and hidden unit, j respectively; a_i, b_j are the biases for visible and hidden units; w_{ij} is the weight between visible units i, and hidden units j.

The probability of a configuration over both visible and hidden units is given by the following equation:

$$p(v,h;\theta) = \frac{1}{Z(\theta)}\exp(-E(v,h;\vartheta))$$
$$Z(\theta) = \sum_v\sum_h E(v,h;\theta) \quad (2)$$

where $Z(\theta)$ is the "partition function" that sums all the possible pairs of visible and hidden vectors. When visible and hidden units have low energy, their probability of distribution is high. The probability of a training vector can be enhanced by adjusting θ to a lower energy of the units.

We pay more attention to distribution of the observed data (v) in a practical issue. So the aim of learning process is to raise the probability of the training data (v), which is given as follows:

$$P(v|\theta) = \frac{1}{Z(\theta)}\sum_h e^{-E(v,h|\theta)}. \quad (3)$$

To obtain the optimal w_{ij} of θ, the log probability of a training vector is derived as follows:

$$\frac{\partial \log p(v)}{\partial w_{ij}} = \langle v_i h_j\rangle_{\mathrm{data}} - \langle v_i h\rangle_{j\,\mathrm{model}} \quad (4)$$

where "data" and "model" denote the distribution of $P(h|v^t;\theta)$ and $P(h,v|\theta)$, respectively; $\langle . \rangle_p$ denotes the expectations under the distribution of p; w_{ij} is updated by the following equation:

$$\Delta w_{ij} = \varepsilon\left(\langle v_i.h_j\rangle_{\mathrm{data}} - \langle v_i.h_j\rangle_{\mathrm{model}}\right) \quad (5)$$

where ε is the learning rate.

$\langle v_i.h_j\rangle_{\mathrm{data}}$ is easily determined, while there are no direct connections between hidden units in an RBM. The conditional distributions of hidden unit (h) and input vector (v) are given by logistic functions, as follows:

$$p(h_j = 1|v) = \sigma\left(b_j + \sum_i v_i w_{ij}\right)$$
$$p(v_j = 1|h) = \sigma\left(a_j + \sum_i h_i w_{ij}\right)$$
$$g(x) = \frac{1}{1+\exp(-x)}. \quad (6)$$

However, it is much more difficult to determine $\langle v_i.h_j\rangle_{\mathrm{model}}$. In this study, a contrastive divergence (CD) method is used. The change in weight is finally given by the following equation:

$$\Delta w_{ij} = \varepsilon\left(\langle v_i.h_j\rangle_{\mathrm{data}} - \langle v_i.h_j\rangle_{\mathrm{reconstruction}}\right) \quad (7)$$

where the reconstruction uses only the information in hidden units that is learned as features from the input. If the input data is recovered perfectly, the weights and biases are deemed as good measures of the input data.

C. *Deep Belief Network (DBN)*

Although some shallow features can be extracted from hyperspectral images (HIS) by a single hidden layer RBM, it is insufficient for the user to obtain high accuracy when those features are applied to classification. Accordingly, DBN is proposed and it is usually applied to speech and pattern recognition. After training the RBM, the output value regarded as a feature of the HIS is used as the input data for the second RBM. In this way, the RBM composes the DBN layer by layer. On the other hand, the final output of the DBN contains all the shallow features that the RBM extracts from the HIS. That is to say, the deep features are extracted gradually.

However, the deep feature extraction from the DBN is not the final step. When it comes to target detection or classification, using the differences of learned features between targets is the ultimate goal. Therefore, an LR layer regarded as a classifier is added to the end of a typical DBN. With the features extracted by the DBN and labeled data, the LR layer fine-tunes all the pretrained parameters with a back-propagation algorithm. In the fine-tune step, a likelihood function is used to construct a cost function, thereby calculating a membership value of the unlabeled data.

D. *Landslide Feature and Detection Frameworks Based on DBN*

A landslide, as showed on the remote sensing images, is generally divided into three areas: source area, transition area, and deposition area [75]. Fragments peeled off the source area are distributed across the transition and deposition areas from coarse to fine. As a result, a gradual texture structure in the remote sensing image is formed. Fig. 2 depicts two landslides in high-resolution images. After short air drying and sedimentation, the humidity and looseness of the fragments affect spectral reflectance which is an important characteristic for distinguishing landslides, bare soil, bench-land and other land cover, etc. Furthermore, as for the shape of a plane in a remote sensing image, a landslide is shaped like a tongue, an ellipse, or a horseshoe. Taken together, the abovementioned characteristics based on remote sensing information constitute the foundational condition for landslide monitoring.

Conventional machine-learning techniques for hyperspectral data classification are limited by the feature extractor, which requires abundant domain expertise to design, such as SVM, artificial neural network, and *K*-means. Although these shallow approaches require less time to be trained and have excellent performance in large-scale classification, they are ineffective for small and dim target detection owing to their lack of multi-scale representations [47], [49]. To take advantage of the high-level and intricate structure of landslides in a hyperspectral image, we propose a deep architecture for landslide detection. As far as we know, this is the first application of deep learning to landslide detection with hyperspectral images.

Deep architecture consists of DBN and LR. DBN with some prior knowledge is used to transform raw data into high-level representation. LR, using features learned from the raw data and

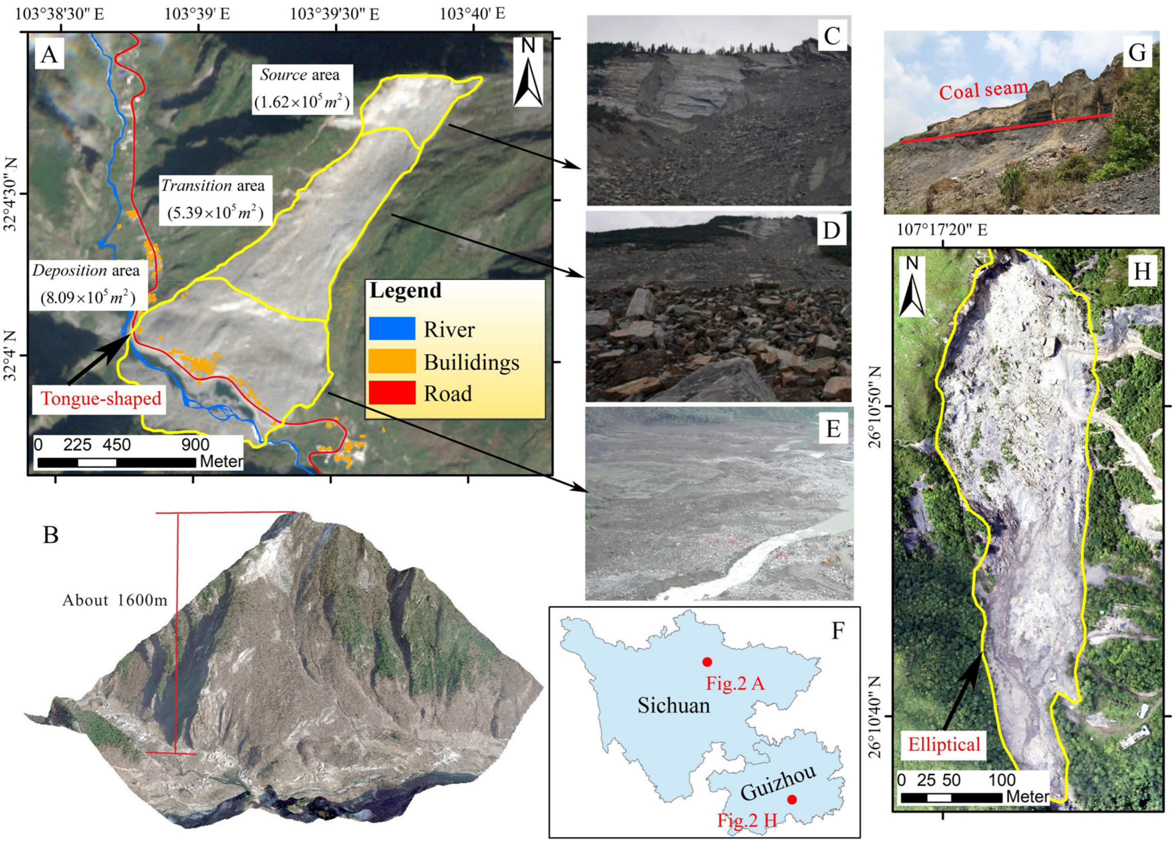

Fig. 2. Landslides on image. (a)–(e) huge landslide in Maoxian. (h) Madaling landslide in Guizhou province. (g) Source area of Madaling landslide.

fine-tuning of the whole architecture, is responsible for classifying the unlabeled pixels. To represent complicated features of a landslide, our framework considers spectral and spatial information that are widely used for classification with the development of remote sensing technology.

As shown in Fig. 3, the first step is to transform the raw data into the input format of DBN. A hyperspectral sensor detects the reflection or radiation intensity of hundreds spectral bands in different wavelengths from the target, which consists of a three-dimensional data cube, including spectral and spatial dimensions. Therefore, the pixel value obtained from the same spatial position in a cube constitutes a one-dimensional ($1 \times N$) vector as the input data of the first layer that denotes the initial spectral features. In terms of spatial information, a neighboring region of the pixel is extracted as a sub-data cube ($H \times H \times N$) of the image and stretched into a one-dimensional vector with ($1 \times H \times H \times N$) elements. However, what requires explaining is that principal component analysis (PCA) is always applied to reduce the data dimension before the transformation of the sub-data cube [76]. First, PCA reduces the high-dimensional raw data into several principal components without loss of spatial information. Thus, PCA prevents subsequent processes from training numerous parameters of the FE system (DBN) and over-fitting

Second, after being normalized, the visible units of the first layer are replaced by the real-valued vector, which is added Gaussian noise [77]. Then, following the CD method proposed by Hinton, the DBN starts training the parameters of every layer in an unsupervised manner. After this process, the output data of the top layer are regarded as a high-level feature vector that is inserted into a classifier to discriminate among different objects.

Finally, our framework, with features extracted by the DBN, employs LR to detect landslides. In addition, the LR layer attached to the top layer of the DBN is responsible for another function: fine-tuning the whole framework with labeled pixels.

To improve the accuracy of landslide detection, the framework adds some constraints to the LR layer, which include soil erodibility, fault, river, road, rainfall erosivity, vegetation coverage, DEM, and slope. To keep the structure consistent, the LR method is also used to analyze the relationship between constrains and landslides. It should be noted that the regression coefficients of constraints would be trained by historical landslides before being input to the top layer of framework. Thus, those constrains could induce a probability of the landslide. Ultimately, a landslide would be identified when the probabilities decided by the image features and predisposing factors reach thresholds respectively.

III. Study Area and Experimental Data

A. Study Area

For the study area in our experiments, we selected a part of Yinxing country, Wenchuan County, Central Sichuan, China, located between longitudes 103.378–103.557°E, and latitudes

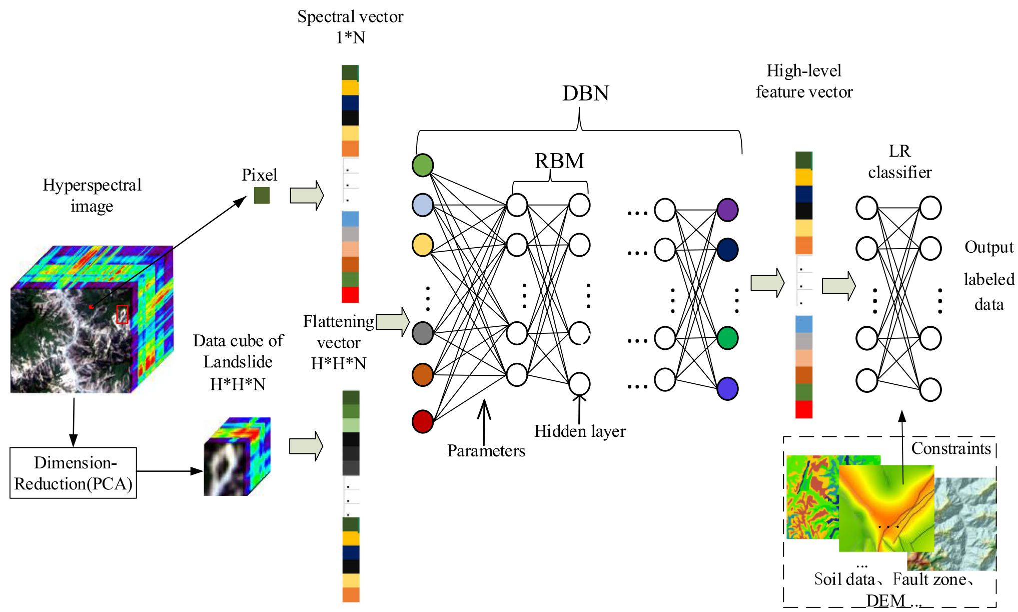

Fig. 3. Landslide detection using the DLWC framework.

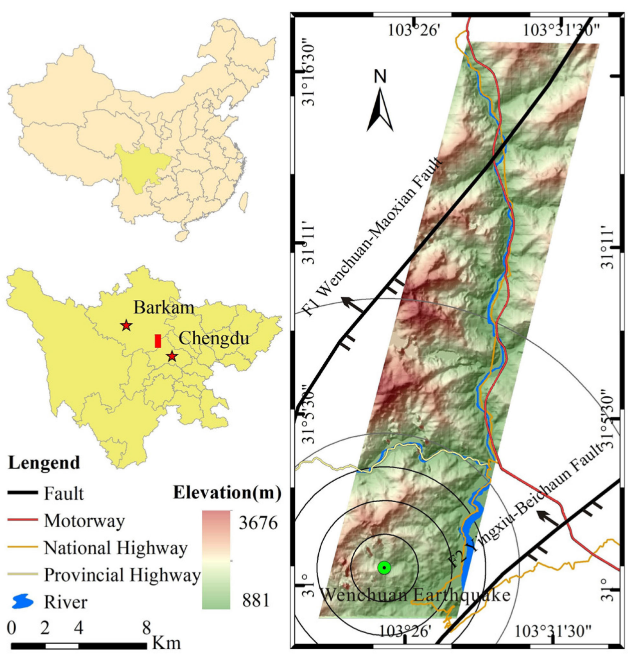

Fig. 4. Map of study area with general regional location.

30.983–31.293°N (see Fig. 4). This area is a typical steep mountain region with elevations from 861 to 3676 m and an area of ∼274 km^2 [78]. Also, the Longmenshan (LMS) fault belt crosses the study area with a strike direction at about N40° E, which is located at the eastern margin of the Tibetan Plateau [79]. Based on GPS measurements, it has been proved that the central Tibetan Plateau is moving eastward into the eastern plateau at a rate of 15–20 mm/yr, because of the continental collision between India and Eurasia in the Cenozoic Era. The extrusion of crustal material from the Tibetan Plateau against the rigid blocks of the Sichuan Basin induces deformation at about 1 mm/yr and accumulates stress in the Longmenshan regions. Consequently, the study area is one of the most tectonically active regions on Earth and seriously threatened by secondary geological disasters triggered by earthquakes.

On May 12, 2008, a catastrophic earthquake with Ms 8.0 struck the LMS region [80]–[82]. The earthquake was triggered by a sudden massive crust displacement along the Yingxiu–Beichuan Fault which is one of the three most active fault zones in LMS, including Wenchaun–Maowen and Pengguan faults. As of January 5, 2009, more than 15 000 geo-hazards were detected by using high-resolution color aerial photographs, satellites, and field investigation [83]–[86]. Among these geohazards, landslides are clustered along two rupture zones. One extending about 250 km along the Yingxiu-Beichuan fault, the other for about 72 km along the Pengguan fault.

B. Experimental Data

A Hyperion sensor carried on an Earth Observing-1 (EO-1) satellite collected images in 242 contiguous bands sampled at approximately 10 nm intervals in a 356–2577 nm range. The spatial resolution is 30 m. Until now, this set of hyperspectral remote sensing data has been one of the most widely used. The data set used in this article was acquired on July 7, 2008, after the Wenchuan Earthquake (see Fig. 5). For our experiments, to avoid zero-value and low signal-to-noise ratio bands, we selected 180 bands as available data in the 436–2405 nm spectral region, from which 4641 pixels were labeled as samples based on field investigation and high-resolution images and, then, divided into training and testing sets in a ratio of 2:8. Detailed information is depicted in Table I.

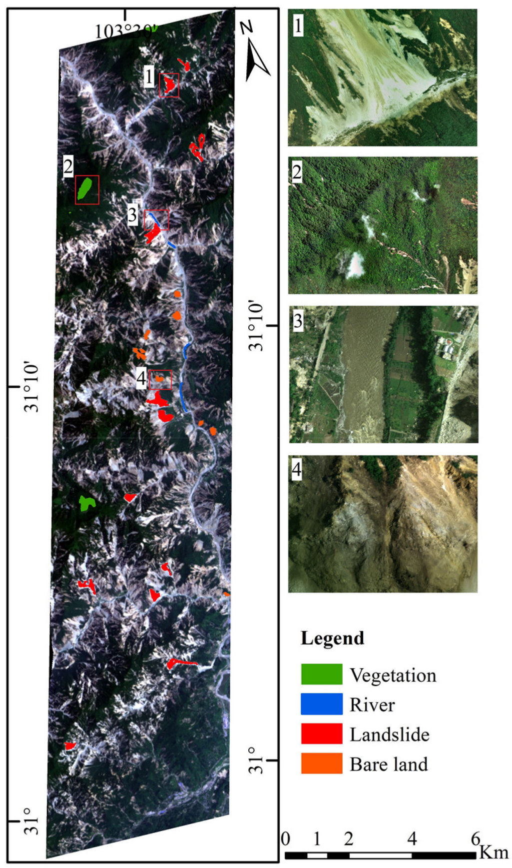

Fig. 5. Hyperion data of study area.

TABLE I
LAND-COVER CLASSES AND NUMBERS OF LABELED PIXELS IN HYPERION

Class code	Name	NO. of training samples	NO. of testing samples
1	Landslide	304	1218
2	Vegetation	222	889
3	River	26	102
4	Bare land	376	1504
Total		928	3713

Landslides are affected by predisposing factors that are commonly divided into natural and human factors. Therefore, besides the Hyperion data, a total of nine landslide predisposing factors for Sichuan province were used as constraints for the detection accuracy of landslides [87]–[90]. These data were extracted from different spatial databases (see Table II).

IV. EXPERIMENTS AND ANALYSES

A. Predisposing Data and Regression Parameters of LR

In this study, raw data of predisposing factors consist of vector and raster data with different ranges. We converted discrete points, lines, and polygons to a computable raster, based on correlation with landslides before being used as constraints of the classifier. Meanwhile, we collected historical landslide data to analyze predisposing factors contributing to landslide occurrence (see Fig. 6).

TABLE II
LIST OF DATA SOURCES USED AS CONSTRAINS IN THE STUDY

Spatial database	Data	GIS/RS data type	Data description
Geological disaster distribution map	Landslide	Point	Land and resources department of Sichuan province
Soil distribution map	Soil erodibility	Polygon	
Geological map	Fault zone	Polyline	
Earthquake database	Earthquake	Point	China Seismological Bureau
Drainage map	River	Polyline	Forestry department of Sichuan province
Traffic map	Road	Polyline	Department of transportation of Sichuan province
Rainfall monitoring data	Rainfall erosivity	Point	Sichuan meteorological station
Landsat 8	Vegetation coverage	Raster	USGS, 30m spatial resolution
DEM	slope	Raster	USGS,1 arc-second spatial resolution

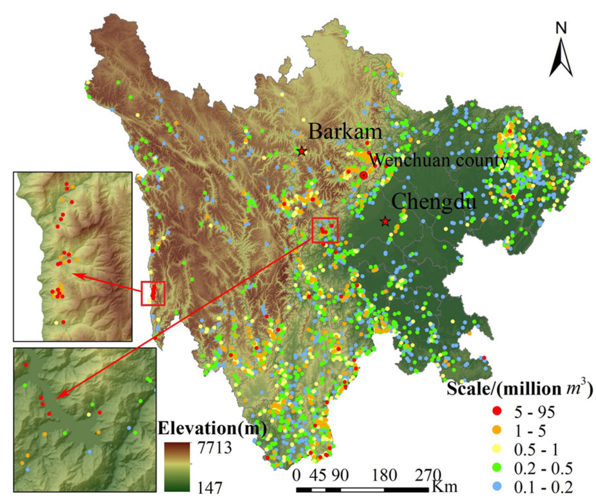

Fig. 6. Landslide Inventory Map of Sichuan Province.

Because it is easily destroyed by heavy rainfall and earthquakes, soil is the main source of landslides. Soil erodibility is calculated as follows [91]:

$$
\begin{aligned}
K_{\mathrm{EPIC}} &= (-0.01383 + 0.51575 \times K_0) \times 0.1317 \\
K_0 &= \{0.2 + 0.3 \times \exp\left[-0.0256 \times m_s\left(1 - m_{\mathrm{silt}}/100\right)\right]\} \\
&\quad \times \left[m_{\mathrm{silt}}/\left(m_c + m_{\mathrm{silt}}\right)\right]^{0.3} \\
&\quad \times \{1 - 0.25 \times \mathrm{org}C/[\mathrm{org}C + \exp\left(3.72 - 2.95\mathrm{org}C\right)]\} \\
&\quad \times \left\{1 - 0.7\left(1 - m_s/100\right) / \left\{\begin{array}{l}(1 - m_s/100) + \\ \exp[-5.51 + 22.9 \\ (1 - m_s/100)]\end{array}\right\}\right\}
\end{aligned} \tag{8}
$$

where K_{EPIC} is the soil erodibility; K_0 is the uncorrected soil erodibility; $m_s, m_{\mathrm{silt}}, m_c$, and $\mathrm{org}C$ are the percentages of sand, silt, clay, and organic matter in soil, respectively.

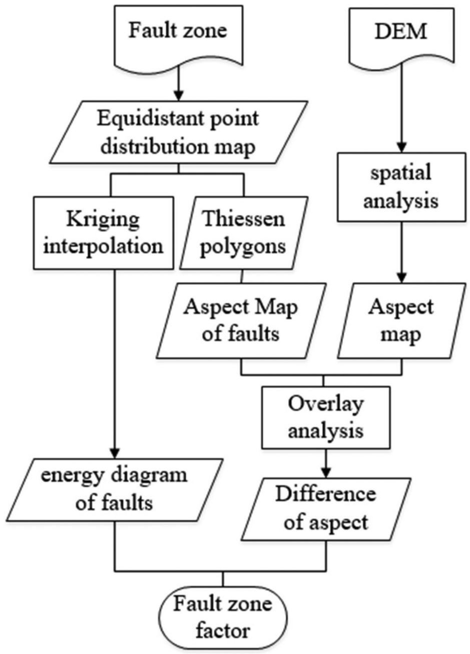

Fig. 7. Induced-factors data processing flowchart based on fault zone.

A fault zone, where crustal movement is very active, affects the stability of the geological environment. To acquire the range and intensity of the fault zone for an unstable geological body and to confirm the main impact zone of every fault in the study area, equidistant points on the fault zone were used to build a Thiessen polygon. Then, based on fault length and difference of direction between slop and fault, the results of the Kriging interpolation were multiplied. Ultimately, we use the result to reflect the impact of faults. Details of this step are illustrated in Fig. 7.

As one of the main predisposing factors for landslides, the influence of the earthquake is evaluated from the results of density analysis and interpolations based on historical earthquake data.

Geological structures adjacent to rivers or roads are easily destroyed. Besides, along with rivers and roads, landslides and other geological disasters are easily triggered when the slopes of rivers and roads two sides are too sharp. Therefore, Kriging interpolation is used to generate a river factor based on river width. Same as the river, the road factor is based on road level.

In term of rain factor, the intensity of erosion caused by rain on a surface is represented by rainfall erosivity and is calculated as follows [92], [93]:

$$R = \sum_{i=1}^{12} \left\{1.735 \times 10 \left(1.5 \times \lg(P_i^2/P\right) - 0.818\right\} \tag{9}$$

where P_i stands for monthly average rainfall; P stands for yearly average rainfall; the unit of rainfall erosivity is the metric unit $(MJ.mm/(hm^2.ha))$.

Vegetation has an anchoring effect to unstable geological bodies. In this study, we represented vegetation coverage by an empirical equation with a vegetation factor as follows [94]:

$$\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{RED}) / (\mathrm{NIR} + \mathrm{RED})$$
$$C = [(1 - \mathrm{NDVI}) /2]^{1+\mathrm{NDVI}} \tag{10}$$

where C stands for vegetation coverage; NDVI stands for normalized vegetation index; RED and NIR stand for the spectral reflectance measurements acquired in the red (visible) and near-infrared regions respectively.

Predisposing factors for landslides are shown in Fig. 8. To guarantee that data from a diverse database can be used to overlap the analysis, we converted all data into a raster format with a resolution of 30 m and transformed in the WGS-84 coordinate system.

In this article, the LR approach is used to analyze the relationship between landslide-occurrence and the predisposing factors. In this step, 3683 large-scale landslides, collected by the Land and Resources Department of Sichuan Province, were used as positive examples. To maintain an equal proportion between landslides and non-landslides, the same number (3683) of points were randomly selected from the non-landslide area as negative examples. Then, the feature data of the examples, obtained from the predisposing factors in the grid format, were put into LR to calculate regression coefficients. The parameters of the optimization problem are estimated by the maximum-likelihood estimation and the gradient descent method. The results (see Table III) show that all the predisposing factors have a *P*-value less than 0.1, indicating a statistical correlation between factors and the susceptibility of landslides at the 90% confidence level. And, those coefficients demonstrate that landslides in Sichuan province are more sensitive to the river and slope factors, less sensitive to the vegetation factor. On the other hand, the standard error of vegetation also indicates that the vegetation indexes extracted from the landslides area are spread out in a distribution. In the model proposed in this paper, the product of predisposing factors and regression coefficients, which serve as constraints, is input to the last layer-LR classifier to improve the accuracy of landslide detection.

B. Detection Frameworks: Characteristic and Analysis

After proposing a deep learning framework with constrains which are conducted from predisposing factors of landslides, we repeat the training process to analyze the landslides detection efficiency with different framework setting. In this article, five factors including the number of hidden units, the depth of layers, the number of principle components and the size of input cubes, are considered. For finding the optimal setting, we used the grid search method with 500 and 200 epochs for pretraining and fine-tunes, respectively. In the end, we applied the framework with the best performance in classification to validate its effect in landslides detection.

1) Number of Hidden Units in RBM: Although the RBM can extract the spectral feature, which is the key factor for classification. There is unavoidable information loss between original data and output data of RBM. As a basic composition of DBN, RBM further decides the representation ability of the whole framework. Therefore, determining the optimal number

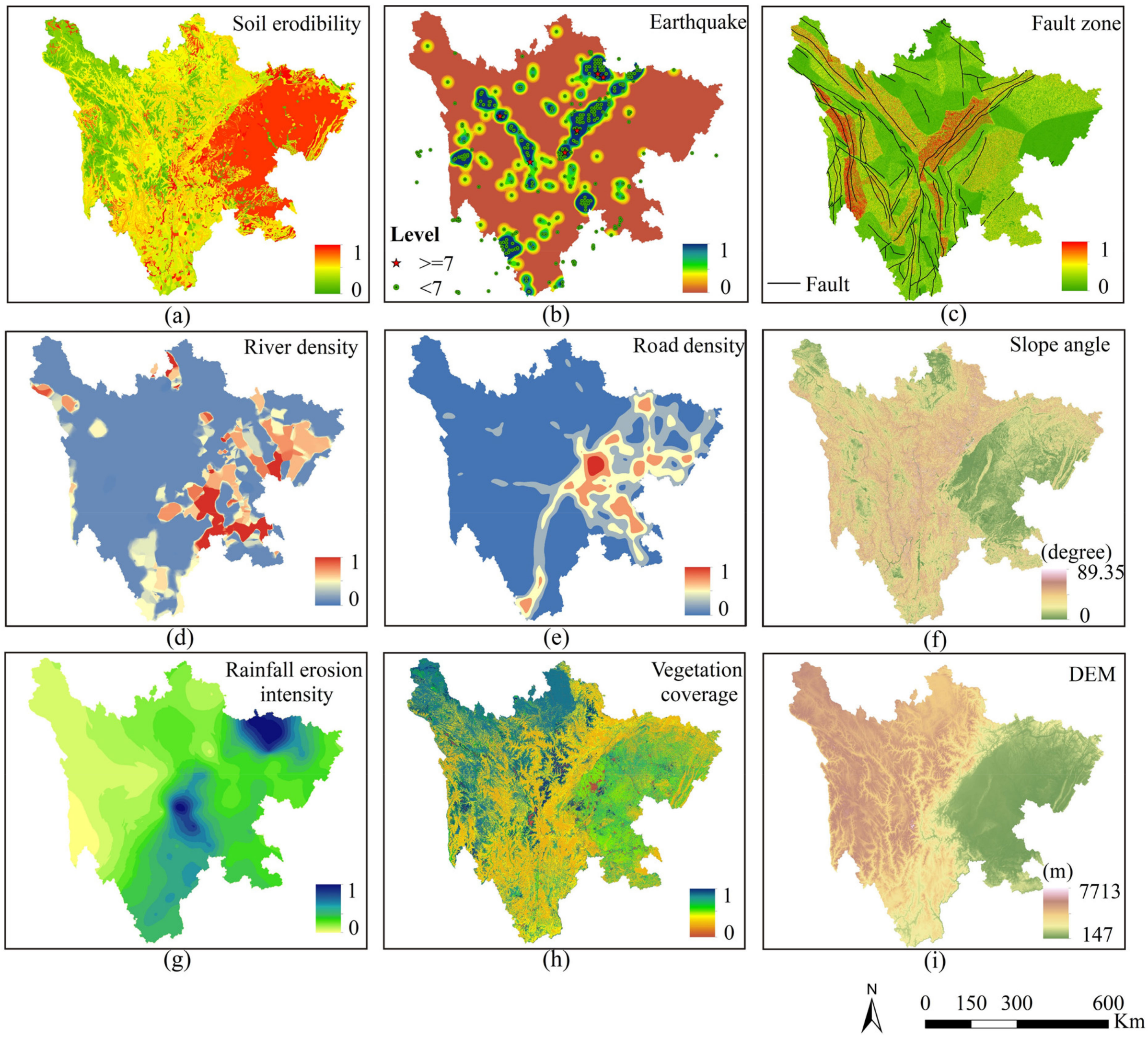

Fig. 8. Landslide-predisposing factors. (a) Soil erodibility. (b) Earthquake factor. (c) Fault zone factor. (d) River density. (e) Road density. (f) Slope angle. (g) Rainfall erosion intensity. (h) Vegetation coverage factor. (i) DEM.

TABLE III
COEFFICIENTS AND STATISTICS OF THE PREDISPOSING FACTORS USED IN THE LOGISTIC REGRESSION EQUATION

Predisposing factors	**Coefficients**	**Standard error**	**P-value**	**Exp(B)**
Soil erodibility	1.511	0.208	0	4.529
Fault	0.529	0.155	0.001	1.697
River	1.691	0.14	0	5.426
Road	–0.546	0.078	0.054	0.579
Rainfall erosivity	0.428	0.129	0.010	1.498
Vegetation coverage	0.283	0.839	0.065	1.326
DEM	0.957	0.238	0	3.46
Slope angle	1.213	0.189	0	3.782
Earthquake	–1.009	0.126	0.006	0.0364
Constant	–0.66	0.141	0	0.517

of hidden units in RBM will not only control the information loss, but also improve the accuracy of landslide detection. First, with the same conditions, the RBM is set with different numbers of hidden units ranged from 2 to 200 at 2 interval, and trained with training dataset. Next, based on the parameters, the square error between four original spectral curves [see Fig. 9(a)] and its reconstruction is calculated.

In Fig. 9(b), the total square error between original spectral curves and reconstruction become smaller with increase of the number of hidden units. On the other hand, the overall accuracy (OA) of classification increases then decreases. Although, in Fig. 9(c), the spectral curves reconstructed by RBM with 40 hidden units can represent the approximate shape of original curves, there exist relatively large square errors on some bands whose reflectance value have significant changes from adjacent bands, as shown in Fig. 9(d). Therefore, considering the square errors and OA, we selected 40 as the optimal number of hidden units for the remote sensing dataset used in this article.

2) *Depths of DBN:* Depth which has a great influence on the performance of classification is directly related to the availability and feature levels of objects. Based on an RBM with 40 hidden neurons, the depth of the framework is increased layer by layer.

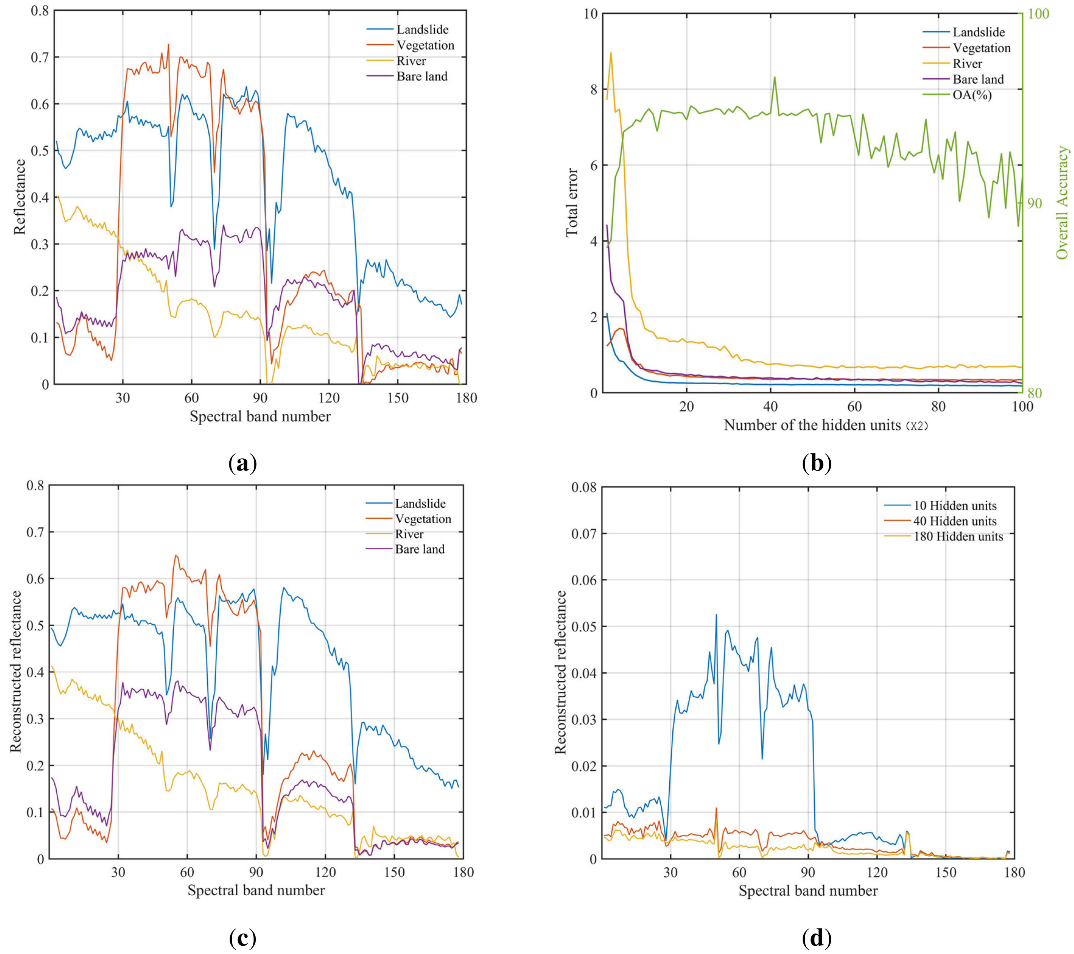

Fig. 9. Reconstruction and error of RBM. (a) Original curve. (b) Total error obtained by RBM with different hidden units ranged from 2 to 200. (c) Reconstruction by RBM with 40 hidden units. (d) Error in every band resulted from RBM with different hidden units.

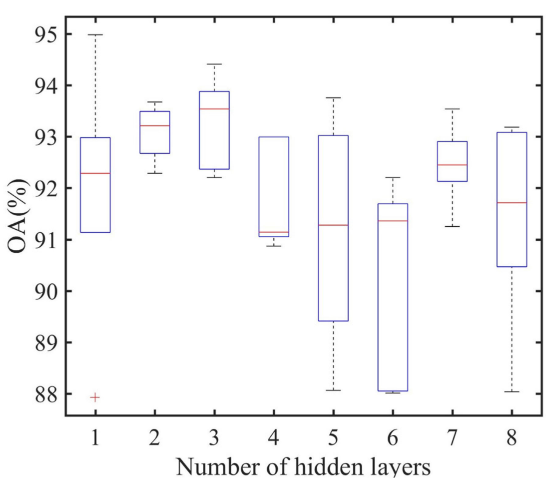

Fig. 10. Influence of depths.

In this process, the DBN is pretrained in 2000 epochs, and the whole framework is fine-tuned in 4000 epochs. Fig. 10 demonstrates the influence of hidden layers by the classification accuracy. As shown in Fig. 10, for the data in this article, the optimal depth is three.

TABLE IV
OA (%) OF DLWC WITH DIFFERENT NUMBER OF PRINCIPLE COMPONENTS

The number of principle components	Overall accuracy (%)	Cumulative variance contribute rate (%)
1	79.12±0.23	90.05
2	91.90±0.41	98.52
3	89.39±0.31	99.02
4	76.40±1.24	99.23
5	71.48±1.43	99.35
6	66.63±1.62	99.46

3) Principle Components: The spatial-dominated feature is one of the most important role in landslide detection based on remote sensing, because it could represent the shape and texture information of the landslide. But, the redundancies in hyperspectral remote sensing which make the spatial information extraction more difficult will decrease the accuracy of landslide detection.

Therefore, before applying to data cubes, the image is processed by PCA which is efficient dimensionality reduction method. Then, we analyze the overall accuracy of classification based on different input data cubes which are generated from selected principle components. In Table IV, the results, from the same framework setting except for number of principle of

TABLE V
OA (%) OF DLWC WITH DIFFERENT INPUT SIZE OF DATA CUBE

The input size of the data cube	Overall accuracy (%)	Accuracy of the landslide detection (%)
3 x 3 x3	92.29±0.54	96.45±0.34
5 x 5 x3	92.64±0.42	96.78±0.45
7 x 7 x3	93.32±0.35	97.23±24
9 x 9 x3	92.58±0.24	95.96±34

components, reveal that the two is optimal number of principle components.

4) Input Size of Data Cube: The resolution of remote sensing image decides the target's spatial detail information and how much pixels the objective contains on the image. Therefore, the spatial feature extraction procedure is directly influenced by the input size of data cube. In Table V, the overall accuracy as well as the accuracy of landslide detection are obtained in different input spatial size, while other framework setting stays the same. It is concluded the results increase with the input size of data cube, which is a consequence of the more discriminative spatial features extracted from a data cube. On the other hand, the spatial feature becomes more complicated coupled with the size of input data, resulting in the decrease of both kinds of accuracy. Considering the consequence of classification, we decide 7×7 as the optimal input size of data cube.

5) Comparison With Other Conventional Methods: In this article, three conventional methods for hyperspectral remote sensing classification [Spectral Angle Match (SAM), Spectral Information Divergence (SID), and linear support vector machine (linear-SVM)] were compared with DLWC for OA, Kappa coefficient, and the accuracy of detecting landslides [95]–[99]. To make a fair comparison, we set input vectors to the same size for all methods and tune methods to their optimal settings. For instance, the penalty parameter of linear-SVM is set to 0.1, and the threshold of SAM and SID are 0.4 and 0.3, respectively. Then, we randomly select 20% samples as training groups to validate different methods.

The results for the four methods are shown in Fig. 11. Overall, because the DLWC method uses spatial information, it eliminates "salt and pepper" phenomena, which is characterized by isolated and spurious pixels in the object boundaries leading to increased noise on resulting map.

Currently, the method based on spatial information is an advanced method that is used to classify remote sensing images, especially in high-resolution remote sensing fields. Furthermore, a landslide predisposing factor is used in the DLWC method to add constraints to landslide detection, thus making DLWC as an effective method for identifying images of wash and bare lands.

Additionally, constraints reduce the error rate of landslide detection. Therefore, as seen from Table VI, overall classification based on the DLWC has the highest OA, which illustrates that the deep features extracted by the deep learning framework benefit for the improvement of classification accuracy.

C. Performance of Landslide Detection on the Whole Image

To validate the capability of the DLWC method to detect landslides in a whole image, the landslide areas were analyzed separately. Fig. 12(b) and (c) display two landslides adjacent to a river; Fig. 12(a) shows the landslide detection results for the entire image.

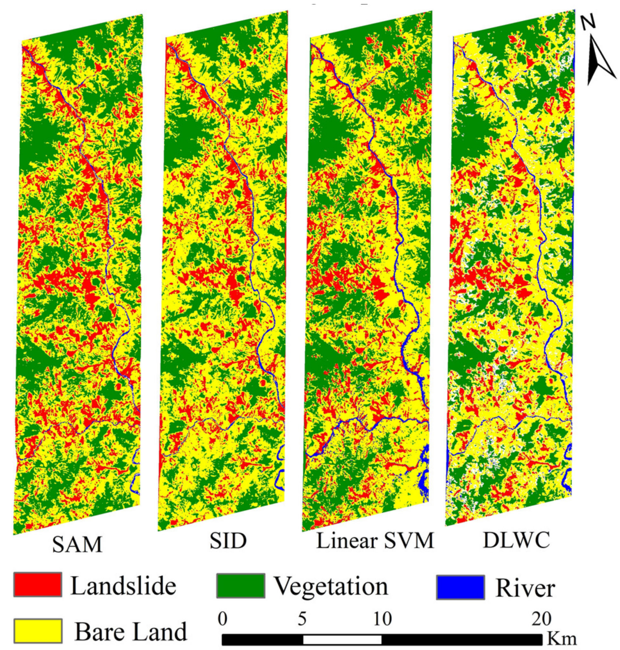

Fig. 11. Classification Maps Obtained by SAM, AID, linear SVM, and DLWC.

TABLE VI
CLASSIFICATION BY DLWC, LINEAR SVM, SID, SAM

Methods	DLWC	Linear-SVM	SID	SAM
Overall Accuracy(%)	97.09± 0.70	94.61±1.01	83.25± 1.14	84.35± 0.83
Accuracy of landslide detection(%)	97.91± 0.76	94.36±0.82	84.50± 0.93	86.44± 0.96
Kappa Coefficient(%)	96.12± 0.95	93.24±0.97	78.38± 0.78	79±1.10

As shown in Fig. 12, our proposed method not only detects landslides, but also effectively reduces noise disturbance, while simultaneously maintaining the basic shapes of the landslides. From high-resolution images and field surveys made after the disasters, 142 landslides were randomly selected to verify our proposed method, indicating that our method missed six landslides and made eight detection errors. In general, detection precision reached 94.4%. Statistics for the landslides are listed in Table VII. Furthermore, we use fragmentation index ranged from 0 to 1 to evaluate fragmentation of results. It presents the more fragmentized when the value of index is closer to 1 [100], [101]. Compared with the three methods mentioned above (SVM, SID, and SAM), the DLWC method achieves a minimal fragmentation index for the extracted results, which is of great help in judging the number of landslides. Thus, the DLWC method effectively avoids error detection caused by the problem of salt and pepper classification, resulting in a reduction of the error rate.

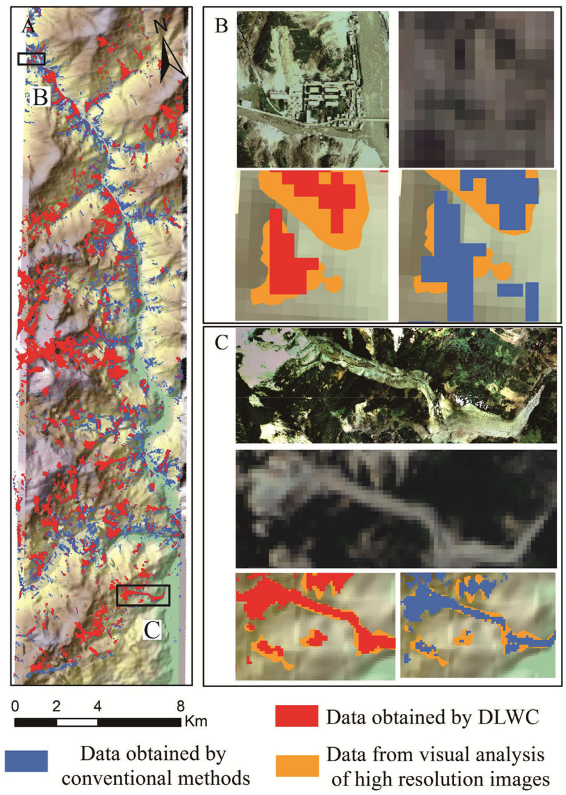

Fig. 12. Comparison of landslide mapping. (a) Distribution diagram of landslide obtained by DLWC and conventional method. (a) and (b) Zoom in the two landslides (the high-resolution images come from airborne remote sensing).

TABLE VII
STATISTICS OF LANDSLIDES DETECTED BY DLWC, SVM, SID, SAM

Methods	DLWC	SVM	SID	SAM
Number of patches	1636	3107	2245	2773
minimum patch size(m^2)	900	900	900	900
maxmum patch size (hm^2)	85.41	131.7	145.98	174.06
Fragment indexes	0.52	0.8	0.64	0.76

V. CONCLUSION

A DLWC was proposed for landslide detection in HIS. Then, we assessed the DLWC performance using HSI data acquired after the M8.0 Wenchuan Earthquake. The experiment results demonstrated that the designed DLWC exhibited better performance in landslide detection, when compared to other conventional methods. Furthermore, the DLWC has the ability to retain morphological information of landslide on the resulting map, especially for giant landslides. It is worth remarking that the proposed framework not only utilizes the image features but also considers the predisposing factors of landslides, which substantially lower the false alarm rate in landslide detection. And, because the uniformity of the input data and the variability of output results, DLWC can be easily applied to other kinds of remote sensing data.

The basic deep learning model is known as its feature extraction which could represent the complex information of original data and enhance performance in target detection compared to other conventional method. However, the result of target detection heavily depend on the framework setting, such as the number of hidden units, the depth of the DBN and the data cube input size. Based on the result of experiments with HIS, it is conducted that the optimal size of data cube and the number of hidden units are 7×7 and 40, respectively. On the other hand, it is not wise to set too many hidden layers in the framework, if the framework is without the ability to compensate the loss information after every hidden layer. Therefore, in DLWC, we set three hidden layers to guarantee the best accuracy of landslide detection.

Generally, in past two decades, many studies focused on how to improve the accuracy of landslide interpretation based on remote sensing, because it is a key role for disaster management and emergency response. But, due to the diversity and regionally of landslides, it is extremely difficult to design the handcrafted features for the representation of landslides. To get the better performance of landslides detection, the DLWC utilizes the deep leaning model to extract the high-level features on hyperspectral images, and detect the landslide combining the predisposing factors. Finally, the experimental results prove that deep learning model could extract the discriminative features of landslide and be successfully applied for landslide detection. Considering the regionally of landslides, we believe the DLWC can make better performance in landslide detection with developments of remote sensing technology. In the future, using other multi-source remote sensing (including high-resolution, SAR, and LiDAR), we will explore more effective deep architectures to detect landslides.

REFERENCES

[1] J. S. Kargel *et al.*, "Geomorphic and geologic controls of geohazards induced by Nepal's 2015 Gorkha earthquake," *Science*, vol. 351, no. 6269, Dec. 2016, Art. no. aac8353.
[2] H. Shahabi and M. Hashim, "Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment," *Sci. Rep.*, vol. 5, no. 3, Apr. 2015, Art. no. 9899.
[3] F. Calò *et al.*, "Enhanced landslide investigations through advanced DInSAR techniques: The Ivancich case study, Assisi, Italy," *Remote Sens. Environ.*, vol. 142, pp. 69–82, Feb. 2014.
[4] S. Pirasteh and J. Li, "Landslides investigations from geoinformatics perspective: Quality, challenges, and recommendations," *Geomat. Nat. Haz. Risk*, vol. 8, pp. 448–465, Oct. 2016.
[5] *Geological Disaster Circular of China*, China Geological Environ. Monitoring Inst., Beijing, China, 2017.
[6] C. Zhao and Z. Lu, "Remote sensing of landslides—A review," *Remote Sens.*, vol. 10, no. 2, Feb. 2018, Art. no. 279.
[7] N. Casagli *et al.*, "Landslide mapping and monitoring by using radar and optical remote sensing: Examples from the EC-FP7 project SAFER," *Remote Sens. Appl.: Soc. Environ.*, vol. 4, pp. 92–108, Oct. 2016.
[8] B. Di *et al.*, "Quantifying the spatial distribution of soil mass wasting processes after the 2008 earthquake in Wenchuan, China: A case study of the Longmenshan area," *Remote Sens. Environ.*, vol. 114, no. 4, pp. 761–771, Apr. 2010.
[9] Y. Dong *et al.*, "Extracting damages caused by the 2008 Ms 8.0 Wenchuan earthquake from SAR remote sensing data," *J. Asian Earth Sci.*, vol. 40, no. 4, pp. 907–914, Mar. 2011.
[10] D. Đurić *et al.*, "Using multiresolution and multitemporal satellite data for post-disaster landslide inventory in the Republic of Serbia," *Landslides*, vol. 14, no. 4, pp. 1467–1482, Aug. 2017.
[11] J. Dou *et al.*, "Optimization of causative factors for landslide susceptibility evaluation using remote sensing and GIS data in parts of Niigata, Japan," *PLoS One*, vol. 10, Jul. 2015, Art. no. e0133262.

[12] F. Chen, B. Yu, and B. Li, "A practical trial of landslide detection from single-temporal Landsat8 images using contour-based proposals and random forest: A case study of national Nepal," *Landslides*, vol. 15, no. 3, pp. 453–464, Mar. 2018.
[13] B. Yu, F. Chen, and S. Muhammad, "Analysis of satellite-derived landslide at central Nepal from 2011 to 2016," *Environ. Earth Sci.*, vol. 77, no. 9, May 2018, Art. no. 331.
[14] R. N. Keyport *et al.*, "A comparative analysis of pixel- and object-based detection of landslides from very high-resolution images," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 64, pp. 1–11, Feb. 2018.
[15] 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, pp. 284–303, Oct. 2005.
[16] P. Lacroix *et al.*, "Use of sentinel-2 images for the detection of precursory motions before landslide failures," *Remote Sens. Environ.*, vol. 215, pp. 507–516, Oct. 2018.
[17] G. Rossi *et al.*, "Multitemporal UAV surveys for landslide mapping and characterization," *Landslides*, vol. 15, no. 2, pp. 1045–1052, May 2018.
[18] Z. Y. Lv *et al.*, "Landslide inventory mapping from bitemporal high-resolution remote sensing images using change detection and multiscale segmentation," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 11, no. 5, pp. 1520–1532, Mar. 2018.
[19] Q. Wang *et al.*, "Application of remote sensing for investigating mining geological hazards," *Int. J. Digit Earth*, vol. 6, no. 5, pp. 449–468, Sep. 2013.
[20] X. Yang and L. Chen, "Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 12, no. 6, pp. 487–495, Dec. 2010.
[21] W. Zhao *et al.*, "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 Obs. Remote Sens.*, vol. 10, no. 5, pp. 1758–1768, Feb. 2017.
[22] P. Frattini *et al.*, "Activity and kinematic behaviour of deep-seated landslides from PS-InSAR displacement rate measurements," *Landslides*, vol. 15, no. 6, pp. 1053–1070, Jun. 2018.
[23] W. Shi, S. Deng, and W. Xu, "Extraction of multi-scale landslide morphological features based on local Gi* using airborne LiDAR-derived DEM," *Geomorphology*, vol. 303, pp. 229–242, Feb. 2018.
[24] G. Herrera *et al.*, "Multi-sensor advanced DInSAR monitoring of very slow landslides: The Tena Valley case study (Central Spanish Pyrenees)," *Remote Sens. Environ.*, vol. 128, pp. 31–43, Jan. 2013.
[25] A. Rosi *et al.*, "The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: Geomorphological features and landslide distribution," *Landslides*, vol. 15, pp. 5–19, Jul. 2017.
[26] J. Dong *et al.*, "Detection and displacement characterization of landslides using multi-temporal satellite SAR interferometry: A case study of Danba County in the Dadu River Basin," *Eng. Geol.*, vol. 240, pp. 95–109, Jun. 2018.
[27] M. Mezaal, B. Pradhan, and H. Rizeei, "Improving landslide detection from airborne laser scanning data using optimized Dempster–Shafer," *Remote Sens.*, vol. 10, 2018, Art. no. 1029.
[28] F. Raspini *et al.*, "Continuous, semi-automatic monitoring of ground deformation using Sentinel-1 satellites," *Sci. Rep.*, vol. 8, no. 1, May 2018, Art. no. 7253.
[29] M. Mwaniki *et al.*, "Image enhancements of Landsat 8 (OLI) and SAR data for preliminary landslide identification and mapping applied to the central region of Kenya," *Geomorphology*, vol. 282, pp. 162–175, Jan. 2017.
[30] C. Zhao *et al.*, "Large-area landslide detection and monitoring with ALOS/PALSAR imagery data over Northern California and Southern Oregon, USA," *Remote Sens. Environ.*, vol. 124, pp. 348–359, Sep. 2012.
[31] F. Bovenga *et al.*, "Using COSMO/SkyMed X-band and ENVISAT C-band SAR interferometry for landslides analysis," *Remote Sens. Environ.*, vol. 119, no. 3, pp. 272–285, Apr. 2012.
[32] V. Tofani *et al.*, "Persistent Scatterer Interferometry (PSI) technique for landslide characterization and monitoring," *Remote Sens.*, vol. 5, no. 3, pp. 1045–1065, Mar. 2013.
[33] K. Gaidzik *et al.*, "Landslide manual and automated inventories, and susceptibility mapping using LIDAR in the forested mountains of Guerrero, Mexico," *Geomatics, Natural Hazards Risk*, vol. 8, pp. 1054–1079, Apr. 2017.
[34] B. A. Leshchinsky, M. J. Olsen, and B. F. Tanyu, *Contour Connection Method for Automated Identification and Classification of Landslide Deposits*. New York, NY, USA: Pergamon Press, Inc., 2015.
[35] W. Li *et al.*, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2016.
[36] C. M. Ye *et al.*, "A method for recognising building materials based on hyperspectral remote sensing," *Mater. Res. Innov.*, vol. 19, pp. S10-90–S10-94, Mar. 2016.
[37] R. Wang, F. Nie, and W. Yu, "Fast spectral clustering with anchor graph for large hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 2003–2007, Sep., 2017.
[38] X. Briottet *et al.*, "Military applications of hyperspectral imagery," *Proc. SPIE*, vol. 6239, May 2006, Art. no. 62390B.
[39] A. Ghiyamat and H. Z. M. Shafri, "A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment," *Int. J. Remote Sens.*, vol. 31, no. 7, pp. 1837–1856, Apr. 2010.
[40] A. O. Onojeghuo *et al.*, "Applications of satellite 'hyper-sensing' in Chinese agriculture: Challenges and opportunities," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 64, pp. 62–86, Feb. 2018.
[41] J. Pontius *et al.*, "Ash decline assessment in emerald ash borer-infested regions: A test of tree-level, hyperspectral technologies," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2665–2676, May 2008.
[42] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 5, no. 2, pp. 354–379, May 2012.
[43] B. Du and L. Zhang, "A discriminative metric learning based anomaly detection method," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6844–6857, Nov. 2014.
[44] Y. Gu *et al.*, "Multiple kernel learning for hyperspectral image classification: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6547–6565, Mar. 2017.
[45] E. Merényi *et al.*, "Classification of hyperspectral imagery with neural networks: Comparison to conventional tools," *Eurasip. J. Adv. Sig. Pr.*, vol. 2014, no. 1, May 2014, Art. no. 71.
[46] J. Sigurdsson, M. O. Ulfarsson, and J. R. Sveinsson, "Semi-supervised hyperspectral unmixing," in *Proc. IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 3458–3461.
[47] T. Wang *et al.*, "A sparse representation method for a priori target signature optimization in hyperspectral target detection," *IEEE Access*, vol. 6, pp. 3408–3424, Nov. 2017.
[48] F. Zhang *et al.*, "Hierarchical feature learning with dropout k-means for hyperspectral image classification," *Neurocomputing*, vol. 187, pp. 75–82, Dec. 2016.
[49] J. M. Haut *et al.*, "Cloud implementation of the K-means algorithm for hyperspectral image analysis," *J. Supercomput.*, vol. 73, no. 1, pp. 514–529, Oct. 2017.
[50] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4085–4098, Aug. 2010.
[51] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
[52] Z. Shao *et al.*, "A novel hierarchical semisupervised SVM for classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 9, pp. 1609–1613, Feb. 2014.
[53] L. Yang *et al.*, "Semi-supervised hyperspectral image classification using spatio-spectral laplacian support vector machine," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 3, pp. 651–655, Jul. 2014.
[54] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Mar. 2013.
[55] O. Ghorbanzadeh *et al.*, "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection," *Remote Sens.*, vol. 11, no. 2, Jan. 2019, Art. no. 196.
[56] Y. Chen, X. Zhao, and X. Jia, "Spectral–spatial classification of hyperspectral data based on deep belief network," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 8, no. 6, pp. 2381–2392, Jan. 2015.
[57] G. Hinton, "Where do features come from?" *Cogn. Sci.*, vol. 38, no. 6, pp. 1078–1101, Aug. 2014.
[58] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[59] H. Shao *et al.*, "Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet," *ISA Trans.*, vol. 69, pp. 187–201, May 2017.

[60] Z. Zhong *et al.*, "Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, Feb. 2017.

[61] Y. Chen *et al.*, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[62] S. Bernhard, P. John, and H. Thomas, "Greedy layer-wise training of deep networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2007, pp. 153–160.

[63] Y. Chen *et al.*, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Jul. 2016.

[64] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[65] Y. Lecun *et al.*, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Dec. 1998.

[66] Y. Li, W. Xie, and H. Li, "Hyperspectral image reconstruction by deep convolutional neural network for classification," *Pattern Recognit.*, vol. 63, pp. 371–383, Mar. 2017.

[67] Y. Li, H. Zhang, and Q. Shen, "Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, pp. 67–88, Jan. 2017.

[68] H. Zhang *et al.*, "Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network," *Remote Sens. Lett.*, vol. 8, no. 5, pp. 438–447, May 2017.

[69] S. Rifai *et al.*, "Contractive auto-encoders: Explicit invariance during feature extraction," in *Proc. 28th Int. Conf. Mach. Learning*, 2011, pp. 833–840.

[70] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. 31st Int. Conf. Mach. Learning*, 2014, pp. 1097–1105.

[71] J. Donahue *et al.*, "DeCAF: A deep convolutional activation feature for generic visual recognition," vol. 50, no. 1, pp. I–647, 2013.

[72] S. Ren *et al.*, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. 28th International Conference on Neural Information Processing Syst.*, 2015, pp. 91–99.

[73] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.

[74] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[75] O. Hungr, S. Leroueil, and L. Picarelli, "The Varnes classification of landslide types, an update," *Landslides*, vol. 11, no. 2, pp. 167–194, Apr. 2014.

[76] I. Jolliffe, "Principal component analysis," in *International Encyclopedia of Statistical Science*, M. Lovric, Ed. Berlin, Germany: Springer, 2011, pp. 1094–1096.

[77] G. E. Hinton, "A practical guide to training restricted boltzmann machines," in *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Germany: Springer, 2012, pp. 599–619.

[78] C. Peng *et al.*, "The 12 May Wenchuan earthquake-induced landslide lakes: Distribution and preliminary risk evaluation," *Landslides*, vol. 6, no. 3, pp. 209–223, Sep. 2009.

[79] R. Huang *et al.*, "The characteristics and failure mechanism of the largest landslide triggered by the Wenchuan earthquake, May 12, 2008, China," *Landslides*, vol. 9, no. 1, pp. 131–142, Mar. 2012.

[80] C. M. Ye *et al.*, "GiT-based structural geologic feature analysis of the southern segment of Longmenshan fault zone for earthquake evidence," *J. Mountain Sci.*, vol. 13, no. 5, pp. 906–916, May 2016.

[81] L. Li *et al.*, "A novel genetic algorithm for optimization of conditioning factors in shallow translational landslides and susceptibility mapping," *Arabian J. Geosci.*, vol. 10, no. 9, pp. 8997–9012, May 2017.

[82] F. C. Dai *et al.*, "Spatial distribution of landslides triggered by the 2008 Ms 8.0 Wenchuan earthquake, China," *J. Asian Earth Sci.*, vol. 40, no. 4, pp. 883–895, May 2011.

[83] S. Qi *et al.*, "Spatial distribution analysis of landslides triggered by 2008.5.12 Wenchuan Earthquake, China," *Eng. Geol.*, vol. 116, no. 1, pp. 95–108, Oct. 2010.

[84] W. Yang *et al.*, "Spatial and temporal analyses of post-seismic landslide changes near the epicentre of the Wenchuan earthquake," *Geomorphology*, vol. 276, pp. 8–15, Jan. 2017.

[85] Y. Yin, F. Wang, and P. Sun, "Landslide hazards triggered by the 2008 Wenchuan earthquake, Sichuan, China," *Landslides*, vol. 6, no. 2, pp. 139–152, Jun. 2009.

[86] T. Gorum *et al.*, "Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake," *Geomorphology*, vol. 133, nos. 3/4, pp. 152–167, Oct. 2011.

[87] X. Chen *et al.*, "Weights-of-evidence method based on GIS for assessing susceptibility to debris flows in Kangding County, Sichuan Province, China," *Environ. Earth Sci.*, vol. 75, no. 1, pp. 70–86, Dec. 2016.

[88] M. Ding and K. Hu, "Susceptibility mapping of landslides in Beichuan County using cluster and MLC methods," *Natural Hazards*, vol. 70, no. 1, pp. 755–766, Jan. 2014.

[89] R. Niu *et al.*, "Susceptibility assessment of landslides triggered by the lushan earthquake, april 20, 2013, China," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 7, no. 9, pp. 3979–3992, May 2017.

[90] C. Xu *et al.*, "Comparison of different models for susceptibility mapping of earthquake triggered landslides related with the 2008 Wenchuan earthquake in China," *Comput. Geosci. UK*, vol. 46, no. 3, pp. 317–329, Sep. 2012.

[91] L. I. Yingjie *et al.*, "Mapping the hotspots and coldspots of ecosystem services in conservation priority setting," *J. Geogr. Sci*, vol. 27, no. 6, pp. 681–696, Jun. 2017.

[92] Y. Xiao *et al.*, "Optimizing hotspot areas for ecological planning and management based on biodiversity and ecosystem services," *Chin. Geogr. Sci.*, vol. 26, no. 2, pp. 256–269, Apr. 2016.

[93] B. Fu *et al.*, "Assessing the soil erosion control service of ecosystems change in the Loess Plateau of China," *Ecol. Complex*, vol. 8, no. 4, pp. 284–293, Dec. 2011.

[94] W.-T. Lin, C.-Y. Lin, and W.-C. Chou, "Assessment of vegetation recovery and soil erosion at landslides caused by a catastrophic earthquake: A case study in Central Taiwan," *Ecol. Eng*, vol. 28, no. 1, pp. 79–89, Nov. 2006.

[95] E. Hasan *et al.*, "Spectral angle mapper and aeromagnetic data integration for gold-associated alteration zone mapping: A case study for the Central Eastern Desert Egypt," *Int. J. Remote Sens.*, vol. 37, no. 8, pp. 1762–1776, Apr. 2016.

[96] R. Jain and R. U. Sharma, "Airborne hyperspectral data for mineral mapping in Southeastern Rajasthan, India," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 81, pp. 137–145, Sep. 2019.

[97] E. Zhang *et al.*, "Improving hyperspectral image classification using spectral information divergence," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 249–253, Jun. 2014.

[98] L. Xie *et al.*, "Hyperspectral image classification using discrete space model and support vector machines," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 374–378, Jan. 2017.

[99] K. Tan *et al.*, "GPU parallel implementation of support vector machines for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 8, no. 10, pp. 4647–4656, Aug. 2015.

[100] C. S. Reddy *et al.*, "National assessment of forest fragmentation in India: Landscape indices as measures of the effects of fragmentation and forest cover change," *Ecol. Eng.*, vol. 60, pp. 453–464, Nov. 2013.

[101] K. McGarigal, "Fragstats Help," 2015. [Online]. Available: http://www.umass.edu/landeco/research/fragstats/documents/fragstats_documents.html

Chengming Ye received the Ph.D. degree in earth exploration and information technology from the Chengdu University of Technology, Chengdu, China, in 2011.

He is currently an Associate Professor with the Department of Geophysical, Chengdu University of Technology. His main research interests comprise Geo-hazard remote sensing applications, ecological remote sensing, and LiDAR data processing.

Yao Li received the B.Sc. degree in saptial information and digital technology and the M.Sc. degree in math from the Chengdu Univeristy of Technology, Chengdu, China, in 2015 and 2018, respectively. He is currently working toward the Ph.D. degree in geotechnical engineering in the Institute of Mountain Hazards and Environment, Chengdu Univerisity of Technology.

His research interests include geo-hazards mechanism and risk management, machine learning, and synthetic aperture radar image processing.

Peng Cui received the Ph.D. degree in water and soil conservation from Beijing Forestry University, Beijing, China.

He is currently a Professor with the Institute of Mountain Hazards and Environment, Chengdu Univerisity of Technology, Chengdu, China; the Institute of Geographic Sciences and Natural Resources Research (Physical Geography), Chengdu Univerisity of Technology. He is also the Academician of Chinese Academy of Sciences, IRDR Science Committee Member. His research interests include geo-hazards (debris flow and landslide) for decades and obtained acknowledged achievements in debris-flow mechanism and risk management technology. He has published more than 300 papers and 19 patents, and authored four monographs plus one atlas. He also founded the *Journal of Mountain Science* in 2004 and made it an influential international journal in the field as the chief editor.

Li Liang received the M.Sc. degree in math from the Chengdu University of Technology, Chengdu, China, in 2008.

She is currently an Associate Professor with the Department of Management Science, Chengdu University of Technology, Chengdu, China. Her research interests include applied mathematics and mathematical optimization, and Geo-hazard risk assessment.

Saeid Pirasteh received the Ph.D. degree in geography (GIS, geoanalytics and LiDAR) from the University of Waterloo, Waterloo, ON, Canada, in 2018, and the Ph.D. degree in geology (remote sensing and GIS) from Aligarh Muslim University, Aligarh, India, in 2004.

He is currently an Associate Professor with the Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, China. His research interests include LiDAR data processing and applications in geology, environmental hazards, and disaster assessment toward implementing UN sustainable development goals (SDGs) 2030. He is the UN-GGIM Academic Network Member. His research interests include research and development of geospatial algorithms, models, software, mobile and web applications.

José Marcato, Jr. (M'17) received the Ph.D. degree in cartographic science from the Sao Paulo State University, Sao Paulo, Brazil.

He is currently a Professor with the Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil. His current research interests include UAV photogrammetry and deep neural networks for object detection, classification and segmentation. He has authored and coauthored more than 30 in refereed journals and over 70 in conferences, including papers published in *ISPRS Journal of Photogrammetry and Remote Sensing*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Wesley Nunes Gonçalves (M'19) received the Ph.D. degree in computational physics from the University of Sao Paulo, Sao Paulo, Brazil.

He is currently a Professor with the Faculty of Computer Science and the Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Campo Grande, Brazil. His current research interests include computer vision, machine learning, deep neural networks for object detection, classification and segmentation. He has coauthored more than 90 research papers, including *Pattern Recognition* and *Pattern Recognition Letters*.

Jonathan Li (M'00–SM'11) received the Ph.D. degree in geomatics engineering from the University of Cape Town, Cape Town, South Africa, in 2000.

He is currently a Professor with both the Departments of Geography and Environmental Management as well as Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada. His main research interests include mobile LiDAR, HD mapping, and artificial intelligence (AI)-based algorithms for information extraction from LiDAR point clouds and remotely sensed imagery. He has coauthored more than 400 publications, more than 200 of which are refereed journal papers.

Dr. Li is currently serving as an Associate Editor of IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Transactions on Intelligent Transportation Systems, *Canadian Journal of Remote Sensing*, and the Editor of *Measurement* published by Elsevier.