Landslide susceptibility mapping based on convolutional neural network and conventional machine learning methods

Rui Liu a,b, Xin Yang a,b *, Chong Xu c, Luyao Li a,b and Xiangqiang Zeng b

aKey Laboratory of Earth Exploration and Information Technology of Ministry of Education, Chengdu University of Technology, Chengdu, Sichuan 610059, China

bCollege of Geophysics, Chengdu University of Technology, Chengdu, Sichuan 610059, China

*cNational Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China

*Corresponding author - (Xin Yang) College of Geophysics, Chengdu University of Technology, Chengdu, 610059, China. Permanent address: No.1, Er Xian Qiao Dongsan Road, Chenghua District, Chengdu, Sichuan, China.

Email: xinyang@stu.cdut.edu.cn. Tel: +86-184-2830-8950. ORCID: 0000-0002-4599-530X

Email address: lr@cdut.edu.cn (Rui Liu), xinyang@stu.cdut.edu.cn, kkshiney@qq.com (Xin Yang), xc11111111@126.com (Chong Xu), liluyao@stu.cdut.edu.cn (Luyao Li), 632695399@qq.com (Xiangqiang Zeng)

Declarations

Founding: Bureau of Geology and mineral resources exploration and development of Sichuan Province under Grant No. 20170612-0413; China Geological Survey under Grant No. 1512007402438.

Conflicts of Interest: The authors declare no conflict of interest.

Availability of data and material: The data in this study subject to third party restrictions. The data that support the findings of this study are available when required.

Code availability: The original code is available when required.

Author’s contributions: Rui Liu conceived the presented idea. Xin Yang wrote the first draft of the manuscript. Chong Xu collected and processed the data. Xin Yang, Xiangqiang Zeng and Luyao Li performed the computation. All authors revised and edited the final version.
Landslide susceptibility mapping based on convolutional neural network and conventional machine learning methods

Abstract: Landslide susceptibility mapping (LSM) is a useful tool to estimate the probability of landslide occurrence, providing a scientific basis for natural hazards prevention, land use planning, and economic development in landslide-prone areas. To date, a large number of machine learning methods have been applied to LSM, and recently the advanced Convolutional Neural Network (CNN) has been gradually adopted to enhance the prediction accuracy of LSM. The objective of this study is to introduce a CNN based model in LSM and systematically compare its overall performance with the conventional machine learning models of random forest, logistic regression, and support vector machine. Herein, we selected the Jiuzhaigou region in Sichuan Province, China as the study area. A total number of 710 landslides and 12 predisposing factors were stacked to form spatial datasets for LSM. The ROC analysis and several statistical metrics, such as accuracy, root mean square error (RMSE), Kappa coefficient, sensitivity, and specificity were used to evaluate the performance of the models in the training and validation datasets. Finally, the trained models were calculated and the landslide susceptibility zones were mapped. Results suggest that both CNN and conventional machine-learning based models have a satisfactory performance (AUC: 85.72% - 90.17%). The CNN based model exhibits excellent good-of-fit and prediction capability, and achieves the highest performance (AUC: 90.17%) but also significantly reduces the salt-of-pepper effect, which indicates its great potential of application to LSM.

Keywords: Landslide susceptibility mapping; Convolutional neural network; Machine learning; GIS; Jiuzhaigou region

1 Introduction

Landslides is one of the most serious hazards, which are driven by geomorphology, geology, hydrology, and human activities (Jaafari et al. 2019). When a big landslide occurs, substantial casualties and infrastructure destruction can be caused. Prediction and management of landslides are thus necessary for prevention and mitigation of the losses caused by such a hazard. However, due to the lack of reliable precursory data, it is generally difficult to predict landslides in real-time at a good precision, i.e. forecasting their time, place, and magnitude. Another method similar to the landslide prediction is susceptibility mapping, which predicts “where” landslides are more likely to occur in the future (Jaafari et al. 2019). The “susceptibility” means the likelihood of landslide occurrence in a targeted region on a basis of local conditions and predisposing factors (Reichenbach et al. 2018).
prediction, landslide susceptibility mapping (LSM) is commonly considered as the first stage in the hazard management and can provide a scientific guideline for further disaster management, as it calculates the probability of landslide occurrence and identify the landslide zone of a specific area.

Over the past years, various methods have been proposed and applied for dealing with LSM, including qualitative and quantitative methods. Qualitative methods are knowledge-driven based methods that dependent on the experience of a geomorphologist (Lucà, Conforti, and Robustelli 2011; Tien Bui et al. 2016). These methods are gradually being replaced by quantitative methods. The development of GIS technology, statistical algorithms, and machine learning has introduced various powerful quantitative methods for LSM. Of them, the widely used models for LSM include, but are not limited to Analytical Hierarchy Process (Abedini and Tulabi 2018; Pourghasemi and Rossi 2017), Weight of Evidence (Dahal et al. 2008; Hasekioğulları and Ercanoglu 2012), Discriminant Analyses (Dong et al. 2009; He et al. 2012), Naïve Bayes (NB) (Lee et al. 2019; Tsangaratos and Ilia 2016), Frequency Ratio (Jana, Sekac, and Pal 2019; Yilmaz 2009), Index-of-Entropy (Constantin et al. 2011; Zhang et al. 2019), and Multivariate Regression (Felicísimo et al. 2013). In particular, some of the popular machine learning based-models such as Random Forests (RF), Logistic Regression (LR), Artificial Neural Network (ANN), and Support Vector Machine (SVM) are widely used and have achieved satisfactory results in LSM (Hong et al. 2016; Reza et al. 2020; Zare et al. 2013). In addition, hybrid models integrated by statistical learning algorithms and machine learning provide more options for LSM. These models include, but are not limited to ANN-fuzzy, fuzzy weight of evidence, EBF-fuzzy logic, Adaptive neuro-fuzzy inference system, and rotation forest (Chen et al. 2018; Zhang et al. 2019).

More recently, in addition to the above mentioned methods, the convolutional neural network (CNN) based model has been applied to LSM (Pham et al. 2020; Sameen, Pradhan, and Lee 2020; Wang, Fang, and Hong 2019; Yi et al. 2020). As a powerful deep learning technique, CNN has shown excellent performance in image classification and identification. The attractive capability of CNN is that it can automatically extract robust and general features from convolutional layers and pooling layers without excessive parameters (Zhao, Du, and Emery 2017). As LSM is essentially a binary problem that mines the relationship between landslide predisposing factors and landslides from the dataset and classifies the corresponding pixels, CNN could also be appropriate for LSM. To the best of our knowledge, according to the dimension of data representation, current implementation of CNN based models in the LSM-related literature can be divided into two categories: (1) converting landslide predisposing factors into 1D or 2D...
datasets to fit the need of CNN, and (2) using 3D datasets for LSM. The 1D or 2D data represents the environmental information of landslide locations and do not exploit the spatial information, while 3D data contains the information of landslides and surroundings (Yi et al. 2020). A detail comparison study of these techniques can be found in the reference (Wang et al. 2019).

However, literature review shows that previous research focused on the comparison of machine learning and statistical methods, while only a few studies have attempted to compare the CNN with conventional machine learning methods. To fill the gap, it is necessary to investigate these techniques and make a quantitative-systematic comparison.

This work makes a comparison of CNN and conventional machine learning methods for LSM. Three typical machine learning models, i.e. random forests, logistic regression, and support vector machine are selected as the comparison pairwise. These three models have been proven to have a good performance in LSM. A total of 12 landslide predisposing factors in term of three-dimension (3D) tensors are fed into the CNN based model. A series of indices including ROC curve, RMSE, Accuracy, and Kappa coefficients are used to perform model evaluation on the training and validation set. The landslides collected after the 8th August 2017 $M_w$ 6.5 Jiuzhaigou earthquake are used to test the resultant landslide susceptibility maps. In recent years, the Jiuzhaigou area has become a landslide prone area and has attracted a great deal of research work.

2 Study area

As shown in Fig. 1, the study area is a region in Jiuzhaigou county on the eastern edge of the Tibetan Plateau, located in Sichuan Province, southwest China. This area covers approximately 738.6 km$^2$ spanning 103°39′E and 103°57′E, and 33°06′N and 33°21′N. Two major geomorphic units, the Tibetan Plateau and the Sichuan Basin, transitioned here and formed a typical valley zone with large undulations and steep slopes. The elevation of the area ranges from 1852 to 4475 m. About half of the study area has a slope degree $>$30.58°. The majority lithological units (87%) are (Carboniferous) sandy argillaceous limestone and dolomitic Limestone, (Triassic) limestone dolomite with argillaceous limestone, and (Quaternary) glacial gravel and proluvium.

In this area, the tectonic activity has led to frequent earthquakes. A total of 49 earthquakes with $M_w$ $>$5.0 have occurred around the region in the last 100 years (Dai et al. 2017). The latest one is the 8th
August 2017 $M_w$ 6.5 Jiuzhaigou Earthquake, triggering a large number of co-seismic landslides. These landslides caused severe damage to the natural landscape and disruption to the local traffic.

**Fig. 1** Location of the study area

### 3 Materials and methodology

In this study, modeling and mapping landslide susceptibility in the Jiuzhaigou region involve four steps: (1) integrating landslide inventory and predisposing factors to establish tensor datasets for CNN based model and vector datasets for conventional machine learning methods. (2) modeling landslide susceptibility in terms of CNN, random forests, logistic regression, and support vector machine. (3) assessing the good-of-fit and predictive capability of models in training datasets and validation datasets. And (4) mapping the landslide susceptibility of the study area.

#### 3.1 Landslide inventory

Landslide inventory is fundamental to the research of landslide susceptibility. A complete landslide inventory map (LIM) records the location, size, and distribution of landslides, which can serve as basic data for further landslide studies.

Tian et al. (2019) established a co-seismic landslide inventory around the epicenter of the 2017 Jiuzhaigou earthquake. In their work, a series of high-resolution images before and after the earthquake close to the seismic origin time, including the pre-seismic images on the Google Earth (GE) Platform and 0.5 m-resolution Geoeye-1 post-seismic images were chosen for visual interpretation. And combined with the filed investigation, a total of 4834 landslides were identified, dominated by small size rockfalls and debris flow. In these landslides, the largest, smallest, and average area are approximately 236336.2 m$^2$, 7.7 m$^2$, and 1993.2 m$^2$ respectively. In this study, considering the image resolution and the research purpose, we selected 710 locations from the original landslide inventory. In addition, to ensure that the distribution pattern of selected landslides was consistent with the original inventory, we used the Subset tool in ArcGIS to established a sub-landslide inventory. Note that all landslide locations were converted into raster data with a grid size of 30×30 m. Subsequently, we divided the landslide inventory into two groups, including training (70%) and validation (30%) data for construction and validation of the
landslide susceptibility model, respectively. As the landslide susceptibility assessment is a binary
classification problem which needs positive and negative samples, 710 non-landslide locations with the
same proportional as landslide data were randomly selected to balance the landslide inventory.

3.2 Landslide predisposing factors

The occurrence of landslide is closely related to the combination of various factors including
topography, geology, and other environmental indexes. A proper combination of predisposing factors
can make the model more competitive (Chen and Chen 2021). Because of the complexity of landslides
and the diversity of triggering sources, the predisposing factors should be chosen based on the
circumstances of the specific area. In this study, 12 landslide predisposing factors namely peak ground
acceleration (PGA), topographic wetness index (TWI), distance to rivers, distance to roads, distance to
faults, normalized difference vegetation index (NDVI), land use, lithology, elevation, slope, aspect, and
topographic ruggedness index (TRI) were chosen for landslide susceptibility modeling.

PGA is an important indicator that reflects the relationship between co-seismic landslide density
and the earthquake, and also a necessary factor in landslide susceptibility mapping (Meunier, Hovius,
and Haines 2007). In this study, the PGA data were adapted from USGS (https://earthquake.usgs.gov/),
which range from 0.12 to 0.26 g, being classified into five groups: (<0.12, 0.12-016, 0.16-0.20, 0.20-
0.24, >0.24) (Fig.2(a). In LSM, TWI is also a frequently-used factor derived from the digital elevation
model (DEM) that quantifies topographic control on hydrological processes (Sørensen, Zinko, and
Seibert 2006). As shown in Fig. 2(b), TWI used in this work was divided into five categories, with the
values of <5.88, 5.88-7.56, 7.56-10.32, 10.32-15.51, 15.51-24.38. The stability of a slope around a river
will be significantly affected by the fluctuation of water in river (Çevik and Topal 2003; Saha, Gupta,
and Arora 2002). The roads can affect the spread and size of landslides. Therefore, the distances to rivers
and roads derived from a topographical map are considered as predisposing factors (Fig. 2(c)-(d)). NDVI
is a crucial factor concerned with the slope stability, especially in mountain areas. Fig. 2(e) shows the
NDVI is divided into five categories, with values ranging from -0.50 to 0.88. Land use is another common
factor contributing to landslides. Using the supervised classification method, the land use was classified
into six categories namely water, construction land, bare land, dense forests, sparse forests, grass land,
and others (Fig. 2(f)). In this study, NDVI and land use were derived from the pre-seismic Sentinel-2
data on the Google Earth Engine (GEE). Geological factors play a significant role in landslide susceptibility, including lithology and distance to faults. The original lithology and faults from a geological map were provided by local authorities with 1:500000 scale. The lithology of the study area mainly contains five groups, including group 1: medium thick bedded sandy argillaceous limestone, dolomitic limestone (Carboniferous), group 2: sandy slate, argillaceous limestone and massive quartz sandstone (Devonian), group 3: clastic limestone (Lower Permian), group 4: glacial gravel and proluvium (Quaternary), and group 5: Medium-thick layered limestone dolomite with argillaceous limestone (Triassic) (Fig. 2(g)). The distance to faults was generated using a Buffer tool in ArcGIS, which was classified into <300 m, 300-600 m, 600-900 m, 900-1200 m, 1200-1500 m, >1500m (Fig. 2(h)). Morphological factors including elevation, slope angle, slope aspect, and TRI were derived from ASTER Global Digital Elevation Model (Fig. 2(i-l)). All predisposing factors were converted into raster data with a grid size of 30×30 m, in accord with the landslide inventory available.

Fig. 2 Maps showing landslide predisposing factors used in this study. (a) PGA, (b) TWI, (c) distance to rivers, (d) distance to roads, (e) NDVI, (f) land use, (g) lithology, (h) distance to faults, (i) elevation, (j) slope angle, (k) slope aspect, (l) TRI

3.3 Establishment of spatial datasets

Having the landslide inventory and predisposing factors, the next step is to format and integrate these data into datasets for further modeling. As shown in Fig. 3(a), all layers of 12 landslide predisposing factors were stacked together to form a tensor with the size of 12×w×h, where w, h represent the length and width of the entire study area, respectively. Then, by overlaying the landslide inventory with the factor tensor, the specific pixel corresponding to each landslide location was obtained. Note that the grid size of all factor layers and the landslide inventory should be the same to ensure that they can be pixel-by-pixel. As mentioned earlier, the grid size of all raster data in this study was 30 × 30 m. However, the factor tensor had different numerical ranges for each dimension. For instance, the slope aspect was divided into 9 groups while elevation divided into 6 groups. Therefore, it was essential to normalize each dimension of the factor tensor for improving the convergence speed and accuracy of the machine learning algorithm (Casella et al., 2013).
Fig. 3(b) shows the landslide and non-landslide locations extracted from the factor tensor which were used in the deep learning model. The cell corresponding to each landslide location is taken as the center and then expanded into a raster with a size of $n \times n$. Each cell is assigned a value that contains the data of all factor layers. In this way, more environmental information around the landslides can be considered for further modeling as opposed to just use one grid (Yi et al., 2020). Similarly, the raster of non-landslide location is extracted from the tensor data. The size of raster used for learning should be set according to specific demand (H. Wang et al., 2020). In this study, the size of landslide and non-landslide raster was $17 \times 17$ using trial and error approach. A total of 1420 12-dimensional training and validation tensors extracted from landslide inventory were generated. Table 1 summarizes the CNN datasets we produced in the study area.

In addition, Fig. 3(c) is the diagram of dataset production for the conventional machine learning method. The cell corresponding to each landslide location is extracted from all factor layers and converted to a one-dimension array with 12 elements. With the addition of non-landslide locations and the label column, we finally get a matrix with a size of $1420 \times 13$ for learning.

| Table 1. Number of pixels in dataset of the CNN based model |
|---------------|---------------|
| Fig. 3 Schematic diagram of spatial datasets generation process |

### 3.4 Convolutional neural network (CNN)

Convolutional neural network is perhaps the most popular and widely used deep learning algorithm, which is composed of three key components: the convolutional layer, down-sampling layer, and fully connected layer (Ye et al. 2020). As the key of CNN, convolutional layers contain multiple convolution kernels that extract finer feature information from the previous layer. Meanwhile, the shared weights strategy in convolution layer allows the entire network to be trained with fewer parameters than the fully connected network. The down-sampling layer, known as the pooling layer, is leveraged to reduce the size of the feature information as well as to improve the overfitting resistance of the model in face of different data. The fully connected layer actually acts as a classifier with the same structure as that of the conventional fully connected network. The input of the fully connected layer is the high dimensional...
features extracted by convolution and pooling operations. Using these basic layers, various extended
CNN architectures have been proposed and applied in many fields.

In landslide susceptibility mapping, the tensor dataset carrying information on all predisposing
factors are input to the convolutional layer. And a series of features related to the landslide event are
extracted by the convolution operation. Besides, the pooling layer after the convolutional layer can filter
the redundant information. Then the previous features match the label of input data in the fully connected
layer. In the training process, the parameters of CNN are constantly updated with the backpropagation
algorithm until an acceptable training accuracy is reached. In this study, the architecture of CNN was
designed as shown in Fig. 4.

Fig. 4 The CNN architecture

3.5 Conventional machine learning methods

3.5.1 Random forest

Random forest (RF) is a popular method of machine learning introduced by Breiman (2001). It is
an ensemble algorithm which generates multiple decision trees with different classification capabilities
to learn input data using bootstrap sampling strategy. The classification is achieved by voting among all
the independent decision trees in RF. Different from the single decision tree, the input data of each tree
in RF are randomly selected from all input dataset, and each node is split using the subset of predictive
features (i.e. landslide predisposing factors) that are randomly selected (Youssef et al. 2016). In such a
way, RF increases diversity among the decision trees and improves the capability to handle the redundant
data (over-fitting resistance). Thus, RF is a useful method for mining the useful but hidden association
between features and targets within large amounts of data (Chen et al. 2018).

3.5.2 Logistics regression

Logistic regression (LR), a generalized linear model, is widely used in mapping landslide
susceptibility. In this study, the dependent variable of LR was a binary variable that represented the
presence or absence of a landslide (1 means landslide and 0 mean non-landslide), and the independent
variables were the 12 landslide predisposing factors. The goal of LR is to describe the relationship
between binary-coded landslide locations and the various landslide predisposing factors and estimate the probability of landslide occurrence. The general expression of LR is as follows:

\[ z = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \cdots + b_n x_n \]  

(1)

\[ p = \frac{1}{1 + e^{-z}} \]  

(2)

where \( x_1, x_2, \ldots, x_n \) are the landslide predisposing factors, \( b_0 \) is the intercept of the model, parameters \( b_1, b_2, \ldots, b_n \) are the regression coefficients that must be determined, and \( p \) is the probability of landslide occurrence.

In a regression model, the term multicollinearity implies that a perfect linear relationship exists among more than two variables, which makes estimates of the model be not uniquely computed (Ozdemir and Altural 2013). There is thus a need to implement multicollinearity diagnosis before modeling. The variance inflation factor (VIF) and tolerances methods are commonly used to quantify the multicollinearity among predisposing factors in landslide studies. When the VIF is greater than 10 or the tolerance is small than 0.1, there is potential multicollinearity in the datasets (Tien Bui et al. 2016). The result of multicollinearity diagnosis using SPSS software is shown in Table 2, in which the highest VIF is 2.247 and the lowest tolerance is 0.445, implying no potential multicollinearity among landslide predisposing factors.

**Table 2.** Multicollinearity diagnosis

3.5.3 Support vector machine

As a machine learning method based on structural risk minimization (SRM) strategy, the support vector machines (SVM) has been widely used in various classification and regression process. For classification of this tool, similar to other algorithms, a boundary needs to be found to delineate the data space. The SVM algorithm achieves the goal of classification by finding a hyperplane with the maximum distance from classes (Cortes and Vapnik 1995).

Given a training dataset \( D(x_i, y_i) \) with \( n \) samples, where \( x_i \in \mathbb{R}^n \), the labels \( y_i \in \{1, -1\} \), a separating hyperplane is defined as:

\[ \mathbf{w} \cdot \mathbf{x} + b = 0 \]  

(3)
where \( w \) is the coefficient vector that determines the direction of hyperplane, and \( b \) is generally referred to as the bias. Parameters \( w \) and \( b \) affect the hyperplane synthetically. The optimization of \( w \) and \( b \) can be considered as a quadratic programming problem:

\[
\text{Minimize} \frac{\|w\|^2}{2} + C \sum_{i=1}^{l} \xi_i
\]

Subject to \( y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \)

where \( C \) is the penalty, \( \xi \) is the slack factor which is used for the soft-margin classifier, and \( l \) is the number of support vector. The optimal hyperplane is determined after solving the optimization problem with Lagrange multiplier \( \alpha_i \). The classification function is expressed as:

\[
f(x) = \text{sgn}\left( \sum_{i=1}^{l} \alpha_i y_i x_i \cdot x + b \right)
\] (4)

For nonlinear classification, it is common to use nonlinear kernels to map data in a high-dimensional feature space and then classify it using separating hyperplane. By introducing kernel functions, Eq.4 becomes:

\[
f(x) = \text{sgn}\left( \sum_{i=1}^{l} \alpha_i y_i K(x_i \cdot x) + b \right)
\] (5)

In this study, the radial basis kernel was chosen to be the kernel function \( K(x_i \cdot x) \) as

\[
K(x_i \cdot x) = \exp(-\gamma \|x_i - x\|^2)
\] (6)

### 3.6 Generation of landslide susceptibility maps

According to previous studies, there are two main approaches to generate landslide susceptibility maps (LSM). One is to first calculate the susceptibility values of random points in the study area and then perform inverse distance weighted (IDW) spatial interpolation which is an effective method to assign values to the entire area using limited points. Then the landslide susceptibility values are classified into various risk levels using the natural breaking method. Another approach is to calculate the susceptibility values of all pixels of the study area and then classify the values into five risk levels: very low, low, moderate, high, and very high. The choice between random point-based and pixel-based methods is driven primarily by considerations of study area size and computer performance. LSM generated by these methods can reflect the distribution of potential landslide hazards with corresponding probability. However, the pixel-based LSM may require more computer resources and code to implement.
when the study area is large. In this study, nevertheless, the pixel-based method was used to generate landslide susceptibility maps.

### 3.7 Model performance evaluation

In this study, landslide locations are overlaid with the landslide susceptibility map for model evaluation. The receiver operating characteristic (ROC) curve is considered as an effective metric for assessing the performance of predictive models, especially developed for classification. A ROC curve is a graph with two axes, corresponding to “1-specificity” and “sensitivity”. Specificity is the proportion of landslide pixels that are correctly predicted as landslides, Sensitivity is the proportion of non-landslide pixels that are correctly predicted as non-landslide (Youssef et al. 2016). Specifically, the ROC curve obtained on the training dataset is called a success rate curve (SRC) which explains how well the model fits, and the ROC obtained on validation dataset is the prediction rate curve (PRC) which measures the predictive capability of the model. The area under the curve of ROC indicated by the value of area under curve (AUC) (Schneider and Gorr 2015) is another index to test the model performance. The AUC values are between 0.5 and 1. A larger value of AUC generally indicates that the corresponding model can achieve better performance.

Furthermore, several statistical indices including root mean square error (RMSE), accuracy, and Kappa index are adopted for model validation. A full description about Kappa index can be found in the literature (Cohen and J. 1960). Following are the mathematical expressions of these indices:

\[
\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{n} (y_{pred.} - y_{act.})^2}
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FN}
\]

\[
\text{Specificity} = \frac{TN}{TN + FP}
\]

\[
\text{Kappa} = \frac{P_{obs} - P_{exp}}{1 - P_{exp}}
\]

\[
P_{obs} = \frac{TP + TN}{TS}
\]
\[ P_{\text{exp}} = \frac{(TP + FP) \times (TP + FN) + (FN + TN) \times (FP + TN)}{TS} \]  

Note: Accuracy, sensitivity, and specificity can be derived from the confusion matrix, which is a table presenting the percentages of false positive (FP), false negative (FN), true positive (TP), and true negative (TN) observations (e.g. Chen and Zhang 2014). In RMSE: \( y_{\text{pred.}} \) and \( y_{\text{act.}} \) represent the predictive and actual values, respectively. In Kappa: \( TS \) is the total number of landslide samples in the study area.

4 Results

4.1 Construction of models

For the construction of CNN, hyperparameter tuning is the fundamental step to obtain an expected model with high predictability. In this study, the related hyperparameters including convolutional kernel size, pooling size, loss function, optimizer, epoch, batch size, learning rate, and activation function were determined by plenty of experiments. Table 3 shows the final hyperparameters used in the CNN model.

In the training process, the CNN operation on the dataset is as follows: the tensor datasets containing comprehensive environmental information with a size of 12×17×17 are first fed into two convolutional layers and two down sampling layers for convolution and pooling operations generating a feature map with the size of 156×4×4, then the feature map is flattened into the fully connected and matched the landslide label. Finally, the entire model can learn the pattern of landslide attribution and surrounding environment. In terms of conventional machine learning methods, the process is to input the landslide factor vector with size of 12 × 497 (representing 12 predisposing factors and 497 training samples) into the model and match the landslide label for training.

The CNN and conventional machine learning models were developed by Pytorch and sklearn, respectively. The hardware environment of this study is that a personal computer with 6GB graphic card GTX1660Ti, a 2.6 GHz Intel(R) Core (TM) i7-9750H CPU, and 16 GB of RAM.

Table 3. Hyperparameters of the CNN
4.2 Model validation and comparison

The goodness-of-fit of the susceptibility models was evaluated using the training dataset, SRC-AUC, and statistical measures. The results are shown in Fig. 5 (a) and Table 4. It can be seen that the values of SRC-AUC vary from 83.43% to 98.65%, and the CNN model achieves the highest values (98.65%), followed by RF (89.83%), SVM (87.58%), and LR (83.43%). Notably, the SRC-AUC of CNN is significantly larger than other three conventional machine learning models, almost close to 1. The similar pattern is also observed in terms of accuracy, RMSE, kappa, sensitivity, and specificity, which indicates that there is a good degree of fit of the CNN based model with the training dataset.

Model performance assessment only on the training dataset cannot be a guarantee that the models overfit or underfit. The performance of a good model should be balanced between the training dataset and validation dataset. For that reason, indicators consistent with the training phase were used to evaluate the predictive ability of models. As shown in Fig. 5 (b) and Table 4, CNN model has the highest PRC-AUC value (90.17%), followed by RF (88.53%), SVM (88.43%), and LR (85.72%). The highest probability to correctly predict the non-landslide pixels is the CNN model (specificity = 0.83), whereas the RF has the highest probability to correctly predicted landslide pixels as landslides. The kappa coefficient varies from 0.55 to 0.62, which satisfies the strength of agreement given magnitude proposed by Landis and Koch (1977): 0.4-0.6, and 0.6-1 is moderate and almost perfect, respectively. In addition, the values of RMSE and Accuracy indicate a reasonable agreement between the predicted landslides and the real ones.

Overall, all of these susceptibility models achieved an acceptable performance, and the CNN model outperforms other four conventional machine learning models with excellent good-of-fit and predictive ability.

Table 4. Model evaluation and comparison on the training and validation datasets

Fig. 5 ROC curves of models on (a) training dataset and (b) validation dataset

4.3 Landslide susceptibility mapping

In this study, the CNN, RF, LR, and SVM models were used to generate the landslide susceptibility maps. All pixels in the study area were fed into these trained models for calculating the landslide
susceptibility index (LSI), and then the LSI was divided into five susceptibility levels, namely very low (VLS), low (LS), moderate (MS), high (HS), very high (VHS), using the natural break approach in ArcGIS10.6 (Fig. 6). The landslide susceptibility zones that indicate the ratio of each susceptibility level to the whole study area were used to qualitatively analyze the landslide susceptibility maps (Fig. 7).

In landslide susceptibility map of the CNN model, it was observed that VHS area was relatively concentrated with 17% of the study area, mainly distributed in southeast and northeast, and most landslide points accurately fall into it. The same distribution of VHS area was observed in the models of LR and SVM. Whereas, the VHS and LS area were lower than those of CNN models, and more than 50% of the area were calculated as three susceptibility levels ranging from low to high. Results of RF models are nearly similar to LR and SVM but this model is not sensitive to VHS area, which accounts only about 2.5% of the study area.

**Fig. 6** Landslide susceptibility maps of Jiuzhaigou region using (a) convolutional neural network, (b) random forest, (c) logistic regression, and (d) support vector machine

**Fig. 7** Landslide susceptibility zones

5 **Discussion**

5.1 **Computational efficiency**

In recent years, machine learning methods have received extensive attention in LSM. In this paper, CNN was applied to generate the LSM and the comparative analysis results show that the CNN model achieves satisfactory performance than conventional machine learning methods. The performance of models is a vital issue as well as the goal of LSM. But the computational efficiency also needs to be considered. As shown in Table 5, although there is a GPU to accelerate the computation, CNN based model consumes much more time than conventional machine learning models in both training and prediction phase. Accordingly, the performance of conventional machine learning methods is acceptable when considering the computational efficiency. But on the other hand, the CNN based approach used in this study extends the LSM from the point-based processing to the image-based processing with greater potential operating space in terms of the model construction. Compared to conventional machine learning
methods, the deep learning approach can be more flexible in changing the network and dataset structure to suit the specific condition. For instance, Fang et al. (2020) integrated CNN with three conventional machine learning classifiers to assess the landslide susceptibility. The features extracted from the convolutional layers were input into the conventional machine learning classifiers and obtained satisfactory results. Therefore, the decision to use the CNN method or conventional machine learning methods needs to be determined according to the specific conditions such as experimental equipment, model performance, and disaster emergency degree.

**Table 5.** Time consumption of different prediction models

## 5.2 Analysis of salt-and-pepper effect in LSM

The salt-and-pepper effect often appears in image segmentation and classification. The reason for this effect is that single pixels with different values cannot form homogeneous areas (Blaschke, T., Lang, S., Lorup, E., Strobl, J., & Zeil 2000). Nevertheless, none of studies analyzed the presence of salt-and-pepper effect in landslide susceptibility maps. The salt-and-pepper in the image segmentation may make the results difficult to distinguish, which for LSM means that this effect may destroy the coherence of susceptibility assessment in the study area and is not conductive to the subsequent hazards management. Therefore, it is necessary to eliminate the impact of the salt-and-pepper effect in LSM.

Fig. 8 shows the landslide susceptibility maps of different models for the same region, which demonstrate that the CNN based model significantly reduces the salt-and-pepper effect compared with the conventional machine learning-based models. The reason for this should be the difference in the establishment of the datasets. As stated in Section 3.3, the training data of the traditional machine learning model consists of single points and the corresponding predisposing factors, however, the CNN training data is expanded as 17×17 raster data with points, where more pixels are given class labels and more knowledge is learned by the model, thus creating more homogeneous areas. Therefore, the CNN based model is more suitable for LSM when considering the salt-and-pepper effect.

**Fig. 8** Local magnification of Fig. 7 (a), (b), (c) and (d)
6 Conclusions

This study compared and analyzed the performance of the CNN based model and conventional ML models for landslide susceptibility mapping (LSM) in the case of the Jiuzhaigou region where a M6.5 earthquake occurred in 2017. A total of 710 landslides by this event and 12 landslide predisposing factors, including distance to rivers, TWI, TRI, slope aspect, slope angle, PGA, NDVI, land cover, lithology, elevation, distance to roads, and distance to faults were stacked to form spatial datasets, and then fed into the LSM models. As a result, the CNN based model achieved superior performance compared to the conventional RF, LR, and SVM models. The following conclusion can be drawn:

1. Among four landslide susceptibility models (i.e. CNN, RF, LR, and SVM), CNN exhibits the most excellent good-of-fit and predictive capability for LSM on training and validation datasets.

2. Different from the datasets of conventional ML methods, the 3-D dataset allows more spatial information to be considered and learned by CNN models. The LSM generated by the CNN model is not only sensitive to the landslide high-risk zone but also significantly reduces the salt-and-pepper effect, which guarantees the consistency of susceptibility assessment.

3. Although the CNN model achieved significant results, it consumed more time than traditional ML models in both the training and prediction phase. When assessing landslide susceptibility for large areas, time efficiency is an issue that must be considered. Therefore, the choice of the LSM model should be a trade-off between time efficiency and performance.

4. The results of LSM would assist in disaster management and policy making in the Jiuzhaigou region. Also, this study adds value to the literature of landslide susceptibility mapping through a comparative study of CNN and conventional ML models.

Additionally, the limitation of this study is that we only used the typical models and network architecture, and did not combine engineering geology analysis methods into the model. In future research, a sophisticated and specific CNN architecture should be designed for dealing with the LSM, which may lead to better performance.

Acknowledgements

This work was supported by the Bureau of Geology and mineral resources exploration and development of Sichuan Province under Grant No. 20170612-0413 and China Geological Survey under Grant No. 1512007402438.
Conflicts of Interest

The authors declare no conflict of interest.

References

Abedini, M., and S. Tulabi. 2018. “Assessing LNRF, FR, and AHP Models in Landslide Susceptibility Mapping Index: A Comparative Study of Nojian Watershed in Lorestan Province, Iran.” Environmental Earth Sciences 77(11):1–13. doi: 10.1007/s12665-018-7524-1.

Blaschke, T., Lang, S., Lorup, E., Strobl, J., & Zeil, P. 2000. “Object-Oriented Image Processing in an Integrated GIS/Remote Sensing Environment and Perspectives for Environmental Applications. Environmental Information for Planning, Politics and the Public.” Environmental Information for Planning, Politics and the Public 2(March 2019):555-570.

Breiman, Leo. 2001. “Random Forests.” Machine Learning 45(1):5–32. doi: 10.1023/A:1010933404324.

Çevik, E., and T. Topal. 2003. “GIS-Based Landslide Susceptibility Mapping for a Problematic Segment of the Natural Gas Pipeline, Hendek (Turkey).” Environmental Geology 44(8):949–62. doi: 10.1007/s00254-003-0838-6.

Chen, Wei, Jianbing Peng, Haoyuan Hong, Himan Shahabi, Biswajeet Pradhan, Junzhi Liu, A. Xing Zhu, Xiangjun Pei, and Zhao Duan. 2018. “Landslide Susceptibility Modelling Using GIS-Based Machine Learning Techniques for Chongren County, Jiangxi Province, China.” Science of the Total Environment 626:1121–35. doi: 10.1016/j.scitotenv.2018.01.124.

Chen, Xi, and Wei Chen. 2021. “GIS-Based Landslide Susceptibility Assessment Using Optimized Hybrid Machine Learning Methods.” Catena 196(July 2020):104833. doi: 10.1016/j.catena.2020.104833.

Cohen, J. and J. 1960. “A Coefficient of Agreement for Nominal Scales.” Educational & Psychological Measurement 20(1):37–46.

Constantin, Mihaela, Martin Bednarik, Marta C. Jurchescu, and Marius Vlaicu. 2011. “Landslide Susceptibility Assessment Using the Bivariate Statistical Analysis and the Index of Entropy in the Sibicu Basin (Romania).” Environmental Earth Sciences 63(2):397–406. doi: 10.1007/s12665-010-0724-y.

Cortes, Corinna, and Vladimir Vapnik. 1995. “Support-Vector Networks.” Machine Learning 20(3):273–97. doi: 10.1007/bf00994018.

Dahal, Ranjan Kumar, Shuichi Hasegawa, Atsuko Nonomura, Minoru Yamanaka, Takuro Masuda, and Katsushiro Nishino. 2008. “GIS-Based Weights-of-Evidence Modelling of Rainfall-Induced Landslides in Small Catchments for Landslide Susceptibility Mapping.” Environmental Geology 54(2):311–24. doi: 10.1007/s00254-007-0818-3.

Dai, L. X., Q. Xu, X. M. Fan, M. Chang, Q. Yang, F. Yang, and J. Ren. 2017. “A Preliminary Study on Spatial Distribution Patterns of Landslides Triggered by Jiuzhaigou Earthquake in Sichuan on August 8th, 2017 and Their Susceptibility Assessment.” Journal of Engineering Geology 25(4):1151–64.

Dong, Jia Jyun, Yu Hsiang Tung, Chien Chih Chen, Jyh Jong Liao, and Yi Wen Pan. 2009. “Discriminant Analysis of the Geomorphic Characteristics and Stability of Landslide Dams.” Geomorphology 110(3–4):162–71. doi: 10.1016/j.geomorph.2009.04.004.

Fang, Zhice, Yi Wang, Ling Peng, and Haoyuan Hong. 2020. “Integration of Convolutional Neural Network and Conventional Machine Learning Classifiers for Landslide Susceptibility Mapping.” Computers and Geosciences 139(October 2019):104470. doi: 10.1016/j.cageo.2020.104470.

Felcísimo, Àngel M., Aurora Cuartero, Juan Remondo, and Elia Quirós. 2013. “Mapping Landslide Susceptibility with Logistic Regression, Multiple Adaptive Regression Splines, Classification and Regression Trees, and Maximum Entropy Methods: A Comparative Study.” Landslides 10(2):175–89. doi: 10.1007/s10346-012-0320-1.

Hasekiogullari, Gökçe Deniz, and Murat Ercanoglu. 2012. “A New Approach to Use AHP in Landslide Susceptibility Mapping: A Case Study at Yenice (Karabuk, NW Turkey).” Natural Hazards 63(2):1157–79. doi: 10.1007/s11069-012-0218-1.

He, Sanwei, Peng Pan, Lan Dai, Haijun Wang, and Jiping Liu. 2012. “Application of Kernel-Based Fisher Discriminant Analysis to Map Landslide Susceptibility in the Qinggan River Delta, Three Gorges, China.” Geomorphology 171–172:30–41. doi: 10.1016/j.geomorph.2012.04.024.
Hong, Haoyuan, Biswajeet Pradhan, Mustafa Neamah Jebur, Dieu Tien Bui, Chong Xu, and Aykut Akgun. 2016. “Spatial Prediction of Landslide Hazard at the Luxi Area (China) Using Support Vector Machines.” Environmental Earth Sciences 75(1):1–14. doi: 10.1007/s12665-015-4866-9.

Jaafari, Abolfazl, Mahdi Panahi, Binh Thai Pham, Himan Shahabi, Dieu Tien Bui, Fatemeh Rezaie, and Saro Lee. 2019. “Meta Optimization of an Adaptive Neuro-Fuzzy Inference System with Grey Wolf Optimizer and Biogeography-Based Optimization Algorithms for Spatial Prediction of Landslide Susceptibility.” Catena 175(November 2018):430–45. doi: 10.1016/j.catena.2018.12.033.

Jana, Sujoy Kumar, Tingneyuc Sekac, and Dilip Kumar Pal. 2019. “Geo-Spatial Approach with Frequency Ratio Method in Landslide Susceptibility Mapping in the Busu River Catchment, Papua New Guinea.” Spatial Information Research 27(1):49–62. doi: 10.1016/j.sir.2014.0128-0121-x.

Landis, J. Richard, and Gary G. Koch. 1977. “Landis Amd Koch1977_agreement of Categorical Data.” Biometrics 33(1):159–74.

Lee, Sunmin, Moung Jin Lee, Hyung Sup Jung, and Saro Lee. 2019. “Landslide Susceptibility Mapping Using Naive Bayes and Bayesian Network Models in Umyeonsan, Korea.” Geocarto International 0(0):000. doi: 10.1080/10106049.2019.1585482.

Lucà, Federica, Massimo Conforti, and Gaetano Robustelli. 2011. “Comparison of GIS-Based Gully Susceptibility Mapping Using Bivariate and Multivariate Statistics: Northern Calabria, South Italy.” Geomorphology 134(3–4):297–308. doi: 10.1016/j.geomorph.2011.07.006.

Meunier, Patrick, Niels Hovius, and A. John Haines. 2007. “Regional Patterns of Earthquake-Triggered Landslides and Their Relation to Ground Motion.” Geophysical Research Letters 34(20):1–5. doi: 10.1029/2007GL031337.

Ozdemir, Adnan, and Tolga Altural. 2013. “A Comparative Study of Frequency Ratio, Weights of Evidence and Logistic Regression Methods for Landslide Susceptibility Mapping: Sultan Mountains, SW Turkey.” Journal of Asian Earth Sciences 64(Mar.5):180–97. doi: 10.1016/j.jseaes.2012.12.014.

Pap, Vu Dong, Quoc Huy Nguyen, Huu Duy Nguyen, Van Manh Pham, Van Manh Vu, and Quang Thanh Bui. 2020. “Convolutional Neural Network - Optimized Moth Flame Algorithm for Shallow Landslide Susceptible Analysis.” IEEE Access 8(February):32727–36. doi: 10.1109/ACCESS.2020.2973415.

Pourghasemi, Hamid Reza, and Mauro Rossi. 2017. “Landslide Susceptibility Modeling in a Landslide Prone Area in Mazandarn Province, North of Iran: A Comparison between GLM, GAM, MARS, and M-AHP Methods.” Theoretical and Applied Climatology 130(1–2):609–33. doi: 10.1007/s10346-015-069.8.

Reichenbach, Paola, Mauro Rossi, Bruce D. Malamud, Monika Mihir, and Fausto Guzzetti. 2018. “A Review of Statistically-Based Landslide Susceptibility Models.” Earth-Science Reviews 180(March):60–91. doi: 10.1016/j.earscirev.2018.03.001.

Reza, Hamid, Aiding Kornejady, Norman Kerle, and Farzin Shabani. 2020. “Investigating the Effects of Different Landslide Positioning Techniques, Landslide Partitioning Approaches, and Presence-Absence Balances on Landslide Susceptibility Mapping.” Catena 187(November 2019):104364. doi: 10.1016/j.catena.2019.104364.

Saha, A. K., R. P. Gupta, and M. K. Arora. 2002. “GIS-Based Landslide Hazard Zonation in the Bhagirathi (Ganga) Valley, Himalayas.” International Journal of Remote Sensing 23(2):357–69. doi: 10.1080/0143116010014260.

Sameen, Maher Ibrahim, Biswajeet Pradhan, and Saro Lee. 2020. “Application of Convolutional Neural Networks Featuring Bayesian Optimization for Landslide Susceptibility Assessment.” Catena 186(August 2018):104249. doi: 10.1016/j.catena.2019.104249.

Sørensen, Rasmus, U. Zinko, and J. Seibert. 2006. “On the Calculation of the Topographic Wetness Index: Evaluation of Different Methods Based on Field Observations.” Hydrology and Earth System Sciences 10(1):101–12. doi: 10.5194/hess-10-101-2006.

Tian, Yingying, Chong Xu, Siyuan Ma, Xiwei Xu, Shiyuan Wang, and He Zhang. 2019. “Inventory and Spatial Distribution of Landslides Triggered by the 8th August 2017 M 6.5 Jiuzhaigou Earthquake, China.” Journal of Earthquake 30(1):206–17. doi: 10.1007/s12583-018-0869-2.

Tien Bui, Dieu, Tran Anh Tuan, Harald Klempe, Biswajeet Pradhan, and Inge Revhaug. 2016. “Spatial Prediction Models for Shallow Landslide Hazards: A Comparative Assessment of the Efficacy of Support Vector Machines, Artificial Neural Networks, Kernel Logistic Regression, and Logistic Model Tree.” Landslides 13(2):361–78. doi: 10.1007/s10346-015-0557-6.
Tsangaratos, Paraskevas, and Ioanna Ilia. 2016. “Comparison of a Logistic Regression and Naïve Bayes Classifier in Landslide Susceptibility Assessments: The Influence of Models Complexity and Training Dataset Size.” *Catena* 145:164–79. doi: 10.1016/j.catena.2016.06.004.

Wang, Yi, Zhicé Fang, and Haoyuan Hong. 2019. “Comparison of Convolutional Neural Networks for Landslide Susceptibility Mapping in Yanshan County, China.” *Science of the Total Environment* 666:975–93. doi: 10.1016/j.scitotenv.2019.02.263.

Ye, Chengming, Xin Liu, Hong Xu, Shicong Ren, Yao Li, and Jonathan Li. 2020. “Classification of Hyperspectral Images Based on a Convolutional Neural Network and Spectral Sensitivity.” *Journal of Zhejiang University: Science A* 21(3):240–48. doi: 10.1631/jzus.A1900085.

Yi, Yanjing, Zhijie Zhang, Wanchang Zhang, Huihui Jia, and Jianqiang Zhang. 2020. “Landslide Susceptibility Mapping Using Multiscale Sampling Strategy and Convolutional Neural Network: A Case Study in Jiuzhaigou Region.” *Catena* 195(January):104851. doi: 10.1016/j.catena.2020.104851.

Yilmaz, Işik. 2009. “Landslide Susceptibility Mapping Using Frequency Ratio, Logistic Regression, Artificial Neural Networks and Their Comparison: A Case Study from Kat Landslides (Tokat-Turkey).” *Computers and Geosciences* 35(6):1125–38. doi: 10.1016/j.cageo.2008.08.007.

Youssef, Ahmed Mohamed, Hamid Reza Pourghasemi, Zohre Sadat Pourtaghi, and Mohamed M. Al-Katheeri. 2016. “Landslide Susceptibility Mapping Using Random Forest, Boosted Regression Tree, Classification and Regression Tree, and General Linear Models and Comparison of Their Performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia.” *Landslides* 13(5):839–56. doi: 10.1007/s10346-015-0614-1.

Zare, Mohammad, Hamid Reza Pourghasemi, Mahdi Vafakhah, and Biswajeet Pradhan. 2013. “Landslide Susceptibility Mapping at Vaz Watershed (Iran) Using an Artificial Neural Network Model: A Comparison between Multilayer Perceptron (MLP) and Radial Basic Function (RBF) Algorithms.” *Arabian Journal of Geosciences* 6(8):2873–88. doi: 10.1007/s12517-012-0610-x.

Zhang, Tingyu, Ling Han, Heng Zhang, Yonghua Zhao, Xin Li, and Lei Zhao. 2019. “GIS-Based Landslide Susceptibility Mapping Using Hybrid Integration Approaches of Fractal Dimension with Index of Entropy and Support Vector Machine.” *Journal of Mountain Science* 16(6):1275–88. doi: 10.1007/s11629-018-5337-z.

Zhao, Wenzhi, Shihong Du, and William J. Emery. 2017. “Object-Based Convolutional Neural Network for High-Resolution Imagery Classification.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10(7):3386–96. doi: 10.1109/JSTARS.2017.2680324.
Table and figure captions

Table 1. Number of pixels in dataset of the CNN based model

Table 2. The multicollinearity diagnosis

Table 3. Parameter setting of the CNN model

Table 4. Model evaluation and comparison on the training and validation datasets

Table 5. Time consumption of different prediction models

Fig. 1 Location of the study area

Fig. 2 Maps showing landslide predisposing factors used in this study. (a) PGA, (b) TWI, (c) distance to rivers, (d) distance to roads, (e) NDVI, (f) land use, (g) lithology, (h) distance to faults, (i) elevation, (j) slope angle, (k) slope aspect, (l) TRI

Fig. 3 Schematic diagram of spatial datasets generation process

Fig. 4 The CNN architecture

Fig. 5 The ROC curves of models on (a) training dataset, and (b) validation dataset

Fig. 6 Landslide susceptibility maps of Jiuzhaigou region using (a) convolutional neural network, (b) random forest, (c) logistic regression, and (d) support vector machine

Fig. 7 Landslide susceptibility zones

Fig. 8 The local magnification of the Fig. 6 (landslide susceptibility maps)

21
Table 1. Number of pixels in dataset of the CNN based model

|                  | Training set | Validation set |
|------------------|--------------|----------------|
| Landslide pixels | 497×17×17 = 143633 | 213×17×17 = 61557 |
| Non-landslide pixels | 497×17×17 = 143633 | 213×17×17 = 61557 |

Table 2. The multicollinearity diagnosis

| Number | Landslide predisposing factor | Tolerance | VIF  |
|--------|-------------------------------|-----------|------|
| 1      | PGA                           | 0.637     | 1.569|
| 2      | TWI                           | 0.688     | 1.454|
| 3      | Distance to rivers            | 0.501     | 1.996|
| 4      | Distance to roads             | 0.554     | 1.806|
| 5      | NDVI                          | 0.597     | 1.676|
| 6      | Land use                      | 0.767     | 1.303|
| 7      | Lithology                     | 0.840     | 1.191|
| 8      | Distance to faults            | 0.804     | 1.244|
| 9      | Elevation                     | 0.480     | 2.084|
| 10     | Slope angle                   | 0.499     | 2.002|
| 11     | Slope aspect                  | 0.975     | 1.025|
| 12     | TRI                           | 0.445     | 2.247|

Table 3. Hyperparameters of the CNN

| Number | Parameters                  | Values  |
|--------|-----------------------------|---------|
| 1      | Conventional kernel size    | (3,3)   |
| 2      | Pooling size                | (2,2)   |
| 3      | Loss function               | Cross entropy |
| 4      | Optimizer                   | SGD     |
| 5      | Epoch                       | 150     |
| 6      | Batch size                  | 32      |
| 7      | Learning rate               | 0.002   |
| 8      | Activation function         | ReLU    |

Table 4. Model evaluation and comparison on the training and validation datasets

| Data              | Metrics | Landslide susceptibility models |
|-------------------|---------|--------------------------------|
|                   |         | CNN   | RF    | LR    | SVM   |
| Training dataset  | ACC     | 0.99* | 0.83  | 0.76  | 0.80  |
|                   | RMSE    | 0.09* | 0.41  | 0.49  | 0.45  |
|                   | Kappa   | 0.98* | 0.67  | 0.52  | 0.59  |
|                   | Sensitivity | 0.99* | 0.88  | 0.77  | 0.780 |
|                   | Specificity | 0.98* | 0.79  | 0.74  | 0.794 |
| Validation dataset| ACC     | 0.81* | 0.80  | 0.77  | 0.78  |
|                   | RMSE    | 0.43* | 0.45  | 0.48  | 0.47  |
|                   | Kappa   | 0.62* | 0.60  | 0.55  | 0.57  |
|                   | Sensitivity | 0.79  | 0.80* | 0.76  | 0.75  |
|                   | Specificity | 0.83* | 0.81  | 0.79  | 0.81  |

Note: * denotes the best result
**Table 5.** Time consumption of different prediction models

| Method | Training time (s) | Prediction time (s) |
|--------|-------------------|---------------------|
| CNN    | 9.13              | 619.59              |
| RF     | 0.026             | 0.523               |
| LR     | 0.094             | 0.125               |
| SVM    | 0.153             | 15.039              |

**Fig. 1** Location of the study area

![Location of the study area](image)
Fig. 2 Maps showing landslide predisposing factors used in this study. (a) PGA, (b) TWI, (c) distance to rivers, (d) distance to roads, (e) NDVI, (f) land use, (g) lithology, (h) distance to faults, (i) elevation, (j) slope angle, (k) slope aspect, (l) TRI

Fig. 3 Schematic diagram of spatial datasets generation process
Fig. 4 The CNN architecture

Fig. 5 The ROC curves of models on (a) training dataset, and (b) validation dataset
Fig. 6 Landslide susceptibility maps of Jiuzhaigou region using (a) convolutional neural network, (b) random forest, (c) logistic regression, and (d) support vector machine.

Fig. 7 Landslide susceptibility zones

Fig. 8 The local magnification of the Fig. 6 (a), (b), (c) and (d)