Evaluation of Landslide Susceptibility of the Ya'an-Linzhi Section of the Sichuan-Tibet Railway based on Deep Learning

Shibao Wang
Chang'an University
Jianqi Zhuang (jqzhuang@chd.edu.cn)
Chang'an University
Jiaqi Mu
Chang'an University
Jia Zheng
Chang'an University
Jiewei Zhan
Chang'an University
Jie Wang
Chang'an University
Yuting Fu
Chang'an University

Research Article

Keywords: Landslide susceptibility, CNN, DNN, DL, Sichuan-Tibet Railway

Posted Date: October 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-714294/v2

Version of Record: A version of this preprint was published at Environmental Earth Sciences on April 18th, 2022. See the published version at https://doi.org/10.1007/s12665-022-10375-z.

Evaluation of landslide susceptibility of the Ya'an-Linzhi section of the Sichuan-Tibet Railway based on deep learning

Shibao Wang[1], Jianqi Zhuang[1]*, Jiaqi Mu[1], Jia Zheng[1], Jiewei Zhan[1], Jie Wang[1], Yuting Fu[1]
1 College of Geological Engineering and Geomatics/Key Laboratory of Western China Mineral Resources and Geological Engineering, Chang'an University, Xi'an, Shaanxi, 710054, China.

ABSTRACT
The Qinghai-Tibet Plateau is one area with the most frequent landslide hazards due to its unique geology, topography, and climate conditions, posing severe threats to engineering construction and human settlements. The Sichuan-Tibet Railway that is currently under construction crosses the Qinghai-Tibet Plateau; there are frequent landslide disasters along the line, which seriously threaten the construction of the railway. This paper applied two deep learning (DL) algorithms, the convolutional neural network (CNN) and deep neural network (DNN), to landslide susceptibility mapping of the Ya'an-Linzhi section of the Sichuan-Tibet Railway. A geospatial database was generated based on 587 landslide hazards determined by Interferometric Synthetic Aperture Radar (InSAR) Stacking technology, field geological hazard surveys, and 18 landslide influencing factors were selected. The landslides were randomly divided into training data (70%) and validation data (30%) for the modeling training and testing. The Pearson correlation coefficient and information gain method were used to perform the correlation analysis and feature selection of 18 influencing factors. Both models were evaluated and compared using the receiver operating characteristic (ROC) curve and confusion matrix. The results show that better performance in both the training and testing phases was provided by the CNN algorithm (AUC = 0.88) compared to the DNN algorithm (AUC = 0.84). Slope, elevation, and rainfall are the main factors affecting the occurrence of landslides, and the high and very high landslide susceptibilities were primarily distributed in the Jinsha, Lancang, and Nujiang River Basins along the railway. The research results provide a scientific basis for the construction of the Ya'an-Linzhi section of the Sichuan-Tibet Railway within the region, as well as the disaster prevention and mitigation work during future safe operations.

Keywords: Landslide susceptibility; CNN; DNN; DL; Sichuan-Tibet Railway

Landslides are one of the most destructive natural disasters and are caused by a combination of natural and human factors[1]. Under the influence of extreme climatic events, an increasing number of landslides are occurring worldwide, which results in significant economic and human losses[2]. The Sichuan-Tibet Railway was built on the eastern margin of the Qinghai-Tibet Plateau, where the plates collide and are structurally active. It is one of the regions where crustal deformation and tectonic activities are extremely intense today[3,4]. The Sichuan-Tibet Railway traverses the most complex geological, topographic, and topographical areas in the world, and the plate structures along the line are the most active. The active faults are dense within this region, topographic changes are significant, and natural disasters such as landslides, collapses, and mudslides are the most developed[5]. Preliminary investigation results indicate that a total of 3043 geological hazards such as collapses, landslides, and debris flows have been discovered on the Ya'an-Linzhi section of the Sichuan-Tibet Railway[6]. In 2000, a substantial high-speed landslide occurred on the Yigong

Zangbo River in Bomi County, Tibet, which blocked the river and formed a barrier lake of 3.0×108 m^3; the debris flow caused by the dam break washed away the Tongmai Bridge about 17 km downstream, resulting in huge casualties and property losses[7]. Two large sequential landslides formed a dam and the resulting lake along the Jinsha River on October 11 and November 3, 2018. About 24 and 9 × 106 m^3 of material collapsed and rushed into the river[8]. Due to the large inflow rates at the time of damming, the barrier lake level rose rapidly, destroying the Jinsha River Bridge and other downstream coastal transportation facilities, posing considerable risks to the downstream residents and properties[9]. All of these disasters result in enormous damage to roads and the surrounding environment, cause serious threats and impacts to the construction and safe operation of major cross-river transportation projects such as the Sichuan-Tibet Railway. Therefore, it is imperative to evaluate the geological disaster susceptibility of the Sichuan-Tibet Railway Ya'an-Linzhi section, which can provide a scientific basis for the construction of the Sichuan-Tibet railway as well as disaster prevention and mitigation in future safe operations.

Recently, with the rapid development of remote sensing (RS) and geographic information system (GIS) technologies, various machine learning methods have been applied to assess landslide susceptibility mapping, including naïve Bayes[10,11], logistic regression[12,13], artificial neural network[14,15], decision trees[16,17], random forest[18,19], and support vector machines [20,21]. Compared with the subjective and heuristic models, the machine learning models can successfully handle non-linear data with different scales in the fields of remote sensing and disaster mitigation[22]. However, with the continuous in-depth research on machine learning, these models only have a shallow learning structure with one or zero hidden layers; thus, they have shortcomings such as limited training time, unstable convergence, and local optimal[23,24]. To address this problem, the DL framework has recently received more attention. DL has significant advantages over traditional models; it has more non-linear operation levels than the single hidden layer neural network, support vector machine, and other "shallow learning" methods. In addition, the ability to build advanced features encourages the discovery of the deepest connection between the parameters, which generally obtain a robust performance for non-linear processing[25,26]. DL models, especially convolutional neural network (CNN) models and deep neural network (DNN) models, have been successfully used in a wide range of applications and are optimal for the handling of large data sets[27-30]. However, the accuracy of susceptibility is related to the method itself, and relates to the input training dataset including the historical landslides and landslide predisposing factors. In this study, InSAR Stacking technology was used to identify landslide hazards in the study area and use the landslide data to evaluate landslide susceptibility. The InSAR Stacking technology overcomes the limitations of time incoherence and avoids the long temporal separation, spatial incoherence, and atmospheric effects in traditional interferometry methods, producing land deformation results that are more continuous in time and space[31,32]. Therefore, this method has been widely used in landslide identification and deformation monitoring[33,34].

In this study, InSAR Stacking technology was used to identify landslides. The CNN and DNN models were used to evaluate the landslide susceptibility of the Ya'an-Linzhi section of the Sichuan-Tibet Railway. This study proposes: (1) applying the landslides identified by InSAR Stacking technology in the assessment of landslide susceptibility; (2) using deep learning to evaluate the landslide susceptibility of the Ya'an-Linzhi section of the Sichuan-Tibet Railway and compare the prediction performance of the CNN and the DNN models; (3) provide new ideas and valuable information for landslide related research, and provide the government with better use of

land resources in order to achieve economic development.

Study Area

This study area is identified by the 25 km buffer of the Ya'an-Linzhi section of the Sichuan-Tibet Railway. The area is 50957 km^2, which is about 1011 km long (Fig. 1). This area is primarily affected by the warm and humid air currents of the Pacific and Indian Oceans. The regional differentiation of climate along the route is extremely apparent. Along the line, it transitions from the mid-subtropical climate zone in the Sichuan Basin to the plateau sub-temperate humid-sub-temperature-humid zone and the plateau temperate sub-humid-sub-arid zone. The annual average temperature and annual rainfall decrease from east to west as the altitude increases. The vertical zoning characteristics of the climate zone of the Qinghai-Tibet Plateau are obvious, with significant temperature differences between winter and summer, day and night, and strong freeze-thaw weather.

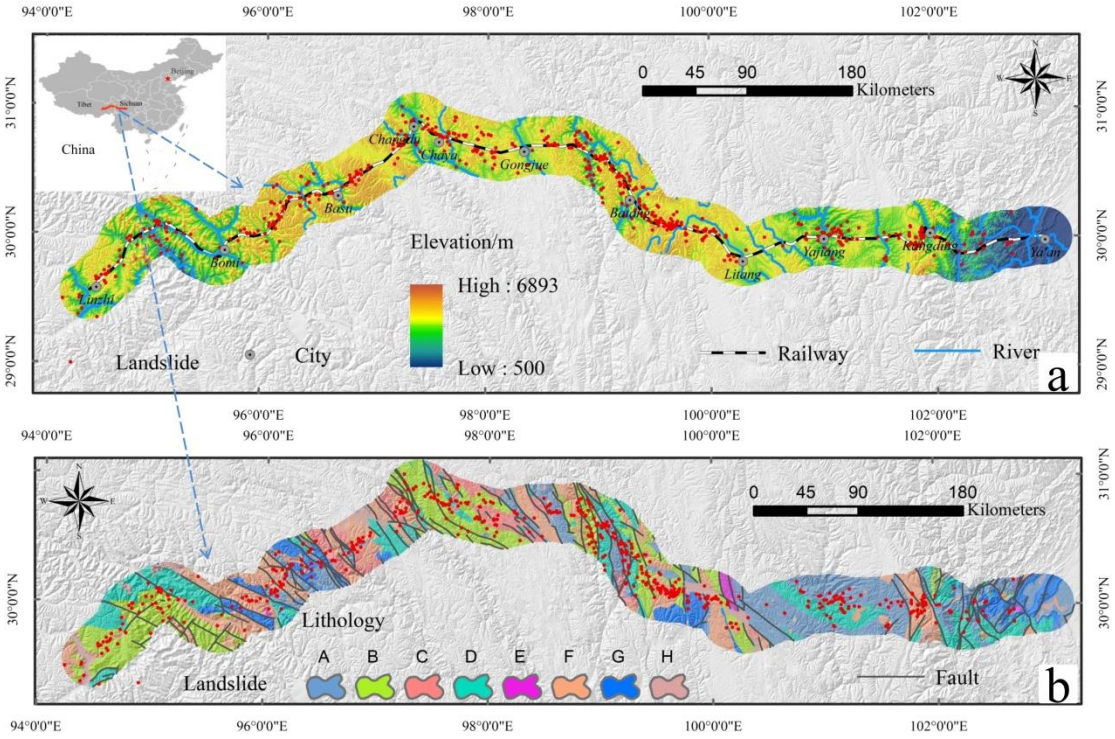

Fig. 1 General situation of the study area and landslide distribution.

(a) geographic location, (b) geological background.

The topography and geomorphology along the Sichuan-Tibet Railway are complex and highly variable. It passes through 5 geomorphological units, namely the Sichuan Basin, West Sichuan Alpine Canyon, West Sichuan High Mountain Plain, Hengduan Mountains in Southeast Tibet, and Southern Tibet. The railway traverses the Hengduan, Nyainqentanglha, and Himalayan Mountains, as well as other mountains, across the Dadu, Jinsha, Nu, and Yarlung Zangbo Rivers. Active faults and strong earthquakes along the line are frequent, such as the Longmenshan, Xianshuihe, Jinshajiang, Lancangjiang, and Nujiang fault zones. The active fault zone controls this area's topography and geomorphology and plays an essential role in controlling the distribution of earthquakes. The formations along the route are diverse and are controlled by geological structures. Except for the Cambrian, it is distributed from the Quaternary to the Sinian.

The main lithologies are sedimentary and metamorphic rocks dominated by sandstone, slate, and phyllite, dominated by granite, and soluble rock dominated by limestone.

Results

Selection of landslide influencing factors

The Pearson correlation coefficient method was used to analyze the correlation of 18 landslide influencing factors in the study area, and the results are shown in Table 1. It can be observed that the correlation coefficient between the factor SCD and TSC is 0.94, the correlation coefficient between the factor TRI and slope is 0.92, and the correlation coefficient is greater than 0.7l; thus, there is a high correlation between the factors. Therefore, the SCD and TRI were removed from the initial factors to improve the data quality, and the remaining 16 factors with less correlation were evaluated for landslide susceptibility in the study area.

Using the information gain method to predict the contribution weight of 16 factors on landslide occurrence, the factors with higher weights are more significant to the prediction methods. In contrast, factors with weights of zero cannot contribute to the landslide susceptibility model and should be excluded from further analysis. Fig. 2 demonstrates the factor weight of each factor determined by the information gain. Based on these results, the AM values of all sixteen factors were greater than zero, implying that these factors contributed to the landslide susceptibility modeling in our study area. The slope factor has the highest AM value of 0.242, indicating that it is the dominant factor to induce landslide occurrence. Secondly, the AM values of elevation, rainfall, and topographic relief are between 0.1 and 0.2, which influence the landslide. The remaining AM values are between 0 and 0.1, indicating that they have little contribution to the occurrence of the landslide.

Table 1 Pearson correlation coefficient results. (Elevation (Ele); Slope (Slo); Aspect (Asp); Plan curvature (Pla); Profile curvature (Pro); Lithology (Lit); Distance to faults (Fau); Land use (Lan); NDVI (NDV); Average annual rainfall (Rai); Distance to rivers (Riv); distance to roads (Roa).

Factors	Ele	Slo	Asp	Pla	Pro	TSC	SCD	TRI	TWI	SPI	STI	Lit	Fau	Lan	NDV	Rai	Riv	Roa
Ele	1.00																	
Slo	-0.05	1.00																
Asp	-0.02	-0.04	1.00															
Pla	0.05	0.08	-0.10	1.00														
Pro	-0.07	0.02	0.07	-0.50	1.00													
TSC	-0.10	0.69	0.03	0.03	0.02	1.00												
SCD	-0.03	0.68	0.02	0.03	-0.02	0.94	1.00											
TRI	-0.08	0.92	-0.04	0.08	0.05	0.64	0.62	1.00										
TWI	0.07	-0.49	-0.12	0.03	-0.02	-0.31	-0.31	-0.40	1.00									
SPI	-0.03	0.48	0.03	-0.13	0.18	0.39	0.36	0.43	-0.16	1.00								
STI	-0.07	0.57	0.03	-0.11	0.17	0.49	0.45	0.52	-0.19	0.64	1.00							
Lit	0.00	-0.05	-0.01	0.03	-0.05	-0.02	-0.04	0.01	0.10	-0.04	-0.01	1.00						
Fau	0.07	-0.10	-0.01	0.02	-0.06	-0.12	-0.10	-0.10	0.07	-0.03	-0.04	0.04	1.00					
Lan	0.37	0.07	-0.03	0.09	-0.02	0.11	0.10	0.09	-0.01	-0.01	0.02	0.15	0.02	1.00				
NDV	-0.44	-0.01	-0.03	-0.02	0.01	-0.08	-0.09	-0.03	-0.17	-0.01	-0.01	-0.06	-0.03	-0.34	1.00			
Rai	-0.32	0.00	-0.02	-0.08	0.01	0.04	0.06	-0.01	-0.02	-0.03	-0.02	0.10	0.13	-0.09	0.04	1.00		
Riv	0.29	-0.01	-0.02	0.02	-0.08	0.00	0.08	-0.05	-0.01	-0.01	-0.04	-0.06	0.07	0.02	-0.11	0.15	1.00	
Roa	0.40	0.10	0.02	-0.04	0.03	0.14	0.20	0.03	-0.06	0.07	0.07	-0.09	0.08	0.16	-0.22	0.09	0.32	1.00

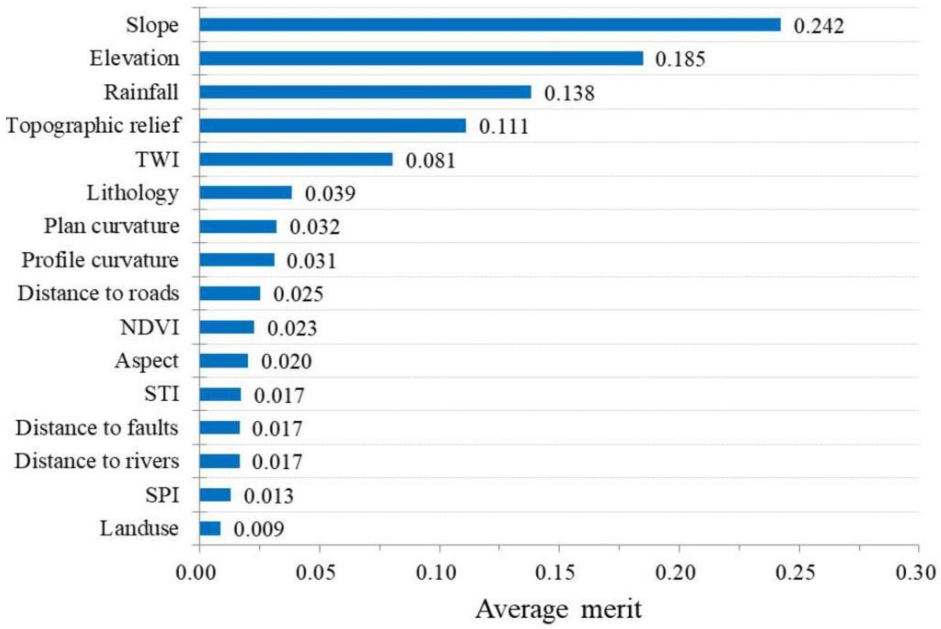

Fig. 2 Average merit (AM) of each landslide influencing factor.

Analysis of landslide influencing factors

The relationship between landslide occurrence and related influencing factors obtained by the FR model is shown in Fig. 3. If the FR value > 1, the corresponding area is more prone to landslide occurrences[35,36]. In the case of elevation, the highest FR of 1.57 is in the class of 4500-5500 m, indicating a high probability of landslide occurrence. For slope, the FR value between 20 and 40 is greater than 1, which positively affects the occurrence of landslides. For reference, the FR value of the east, north, and northeast classes is greater than 1, and the other slope directions are less than 1. The relationship between plan curvature and landslides demonstrates that the highest FR of 1.57 is in the class of > 1, followed by the class of 0.5-1 with an FR of 1.32. For profile curvature, the highest FR of 1.21 is in the class of 0.3-0.7. In the case of TSC, the FR value between 200-400 is greater than 1, promoting landslide occurrences. The relationship between landslide occurrence and TWI showed that the < 6 and > 12 classes have the highest and lowest FR values of 1.45 and 0.15. For the SPI factor, the FR value greater than 90 is greater than 1, and the value of 0-90 is less than 1. For STI, the highest and lowest landslide occurrence probabilities can be reached in the classes of 45-60 and < 15, respectively. In the case of lithology, the groups B and F are more prone to high landslide susceptibility. For distance to faults, the highest FR value of 1.41 is in the class of 3000-4000 m. For land use, the class of grassland and bare land are more prone to higher susceptibility of landslide occurrence. In the case of NDVI, the FR value between 0 and 0.1 is greater than 1, which has a positive effect on the occurrence of landslides. The relationship between rainfall and landslides demonstrates the highest FR of 1.91 and is in the 800-1000 mm class, followed by the class of 600-800 mm with an FR of 1.31. In the case of distance to rivers, the FR becomes smaller as the distance increases. For distance to roads, the < 500 and >2000 classes have the highest FR values of 1.0 and 1.06.

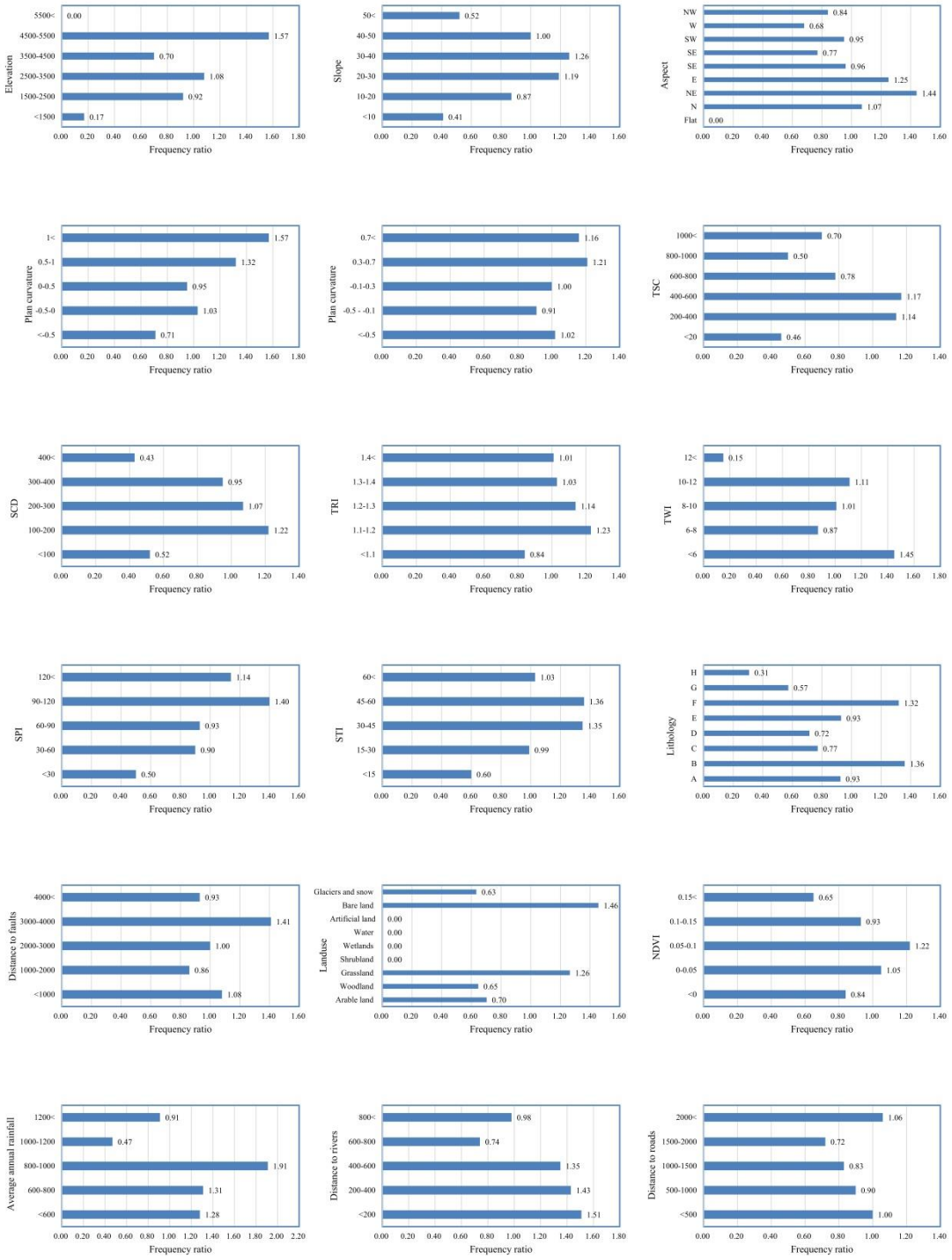

Fig. 3 Frequency ratio chart of the combination factors.

Evaluation of landslide susceptibility

In this study, a CNN model was constructed based on the TensorFlow framework of Python for landslide susceptibility evaluation. Input the selected sample data set into the constructed CNN model for training, select the mean square error (MSE) as the loss function to measure the difference between the output value of the model and the true dependent variable value, and use

AdamOptimizer as the optimizer. After approximately 1300 epochs, the CNN model converged, as illustrated in Fig. 4. Finally, the landslide susceptibility map was generated by the trained CNN model; based on the visual and easy interpretation and comparison of the areas, the susceptibility was classified into five categories: 10%, 10%, 10%, 20%, and 50% (from high to low), corresponding to very high, high, middle, low, and very low susceptibility regions, respectively (Pradhan and Lee 2010; Sun et al. 2020). Fig. 5 shows the landslide susceptibility map of Ya'an-Linzhi Section of the Sichuan-Tibet Railway, showing that the high susceptibility areas are primarily distributed in the Jinsha, Lancang, and Nujiang River Basins along the railway, which conforms with the distribution law of historical landslides.

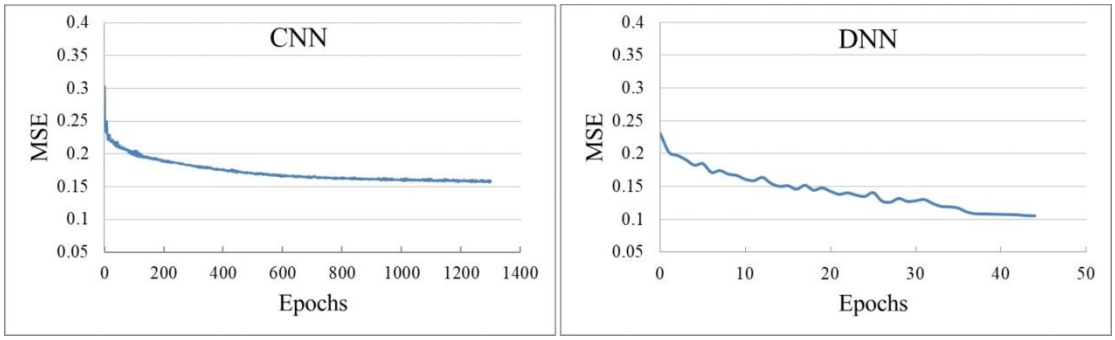

Fig. 4 The variation of MSE values during the training process.

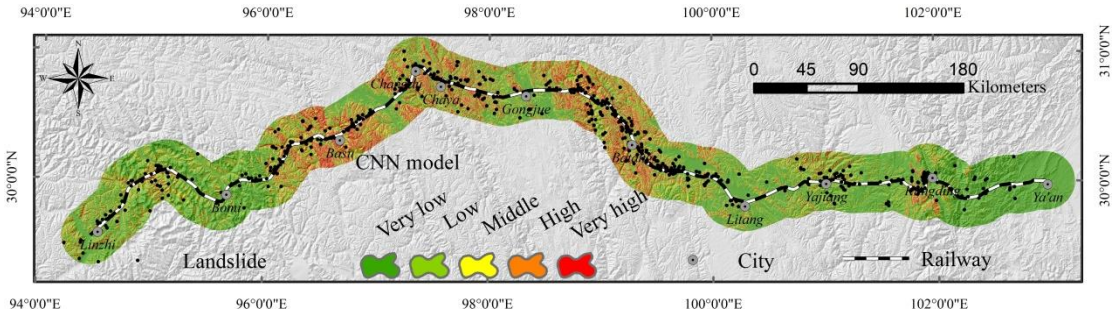

Fig. 5 Landslide susceptibility map generated by CNN model.

The DNN-based approach was implemented using the Python package Keras with Tensorflow as the backend. Similarly, input the training sample data set into the constructed DNN model for training, select the mean square error (MSE) as the loss function, Relu as the activation function, and RMSProp as the optimizer. After approximately 45 epochs, the DNN model converged, as illustrated in Fig. 4. Input 16 factors into the trained DNN model for prediction and obtain the landslide susceptibility index in the study area; the susceptibility was classified into five categories: 10%, 10%, 10%, 20%, and 50%. Fig. 6 shows the landslide susceptibility map produced by the DNN model for the Ya'an-Linzhi Section of the Sichuan-Tibet Railway, overlaid with landslides. The distribution results of high and very high susceptibility areas are more consistent with those predicted by the CNN model. This indicates that the landslide susceptibility map matched well with the distribution of the actual historical landslides.

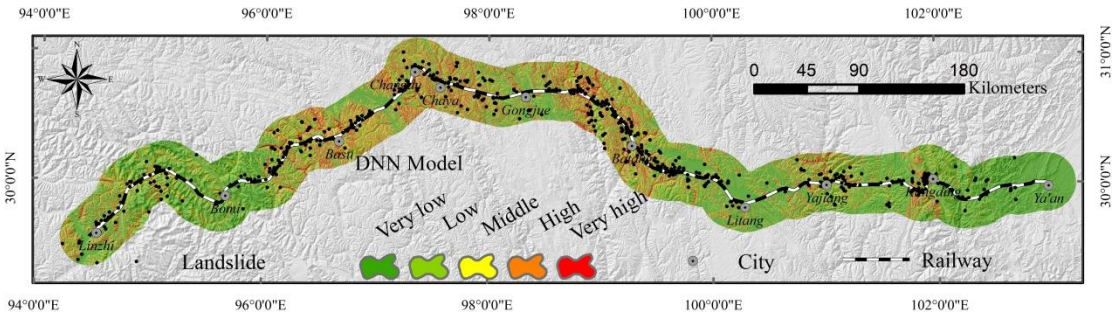

Fig. 6 Landslide susceptibility map generated by DNN mode

Model validation and comparison

Evaluate the goodness-of-fit and the prediction performance of the two models, the results of the ROC curve are shown in Fig. 7, and the confusion matrix is shown in Table 2. It can be observed that the AUC of the training data of the CNN and DNN models are 0.99 and 0.98, respectively. The two deep learning models have a much better performance on the goodness-of-fit to the training data (success rate); the CNN model has the best performance (99%), followed by the DNN (98%). Using test data to verify the prediction performance of the two deep learning models, the AUC of the CNN and DNN models are 0.88 and 0.84, respectively. It can be observed that both models show higher predictive power; however the predictive power of the CNN model is higher than that of the DNN model. In addition, the ACC, recall, precision, and F1 of the confusion matrix are used to validate the test data of the two models. The results are shown in Table 3; these metrics ACC, recall, precision, and F1 of the CNN model were 84.68, 84.48, 84.39, and 84.64, respectively, while those of the DNN model was 79.48, 81.88, 75.72, and 78.68, respectively. All the metrics revealed that, although both models demonstrated reasonable goodness of fit, the CNN model performed better in terms of the training and test datasets. Therefore, the CNN model had a better prediction than the DNN model in this case.

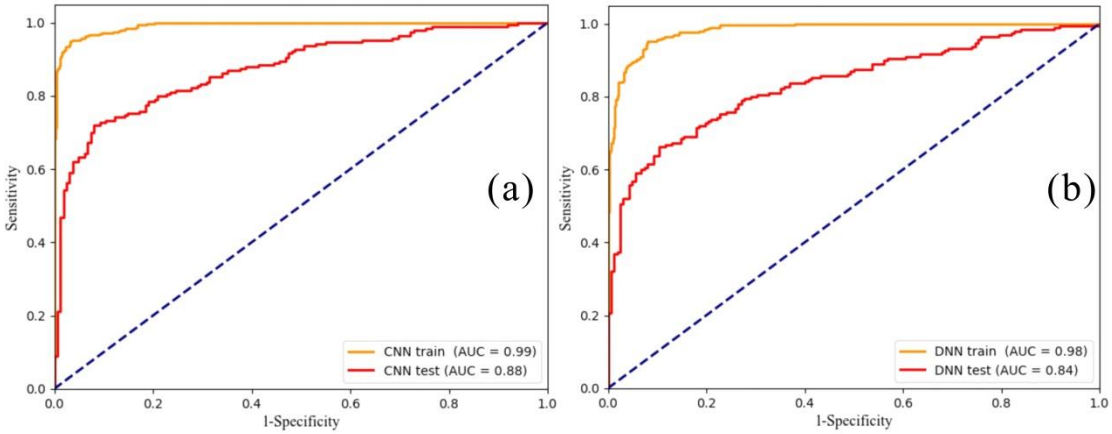

Fig. 7 ROC curves of the CNN (a) and the DNN (b) models.

Table 2 Confusion matrix of the CNN and the DNN models

Model\Parameters	TP	TN	FP	FN	ACC (%)	Recall (%)	Precision (%)	F1
CNN	146	147	27	26	84.68	84.88	84.39	84.64
DNN	131	144	42	29	79.48	81.88	75.72	78.68

The two models were compared to the landslide density (number/km^2) quantitative analysis based on the predicted landslide susceptibility zoning map and historical landslides. The results are shown in Table 3, it can be observed that in the landslide susceptibility maps generated by the two models, as the landslide susceptibility increases, the landslide density also increases, and the landslide density in extremely high-prone areas is the largest. The very high prone area (10%) of the CNN model is distributed with 204 historical landslides, which is 34.75% of the total number of landslides, while the DNN model distributed with 150 landslides (accounting for 25.55%), and 57.41% of the landslides are distributed in the CNN model in very high and high prone areas, and only 48.21% in the DNN model. In addition, there are 70 landslides in the very low prone area (50%) of the CNN model and 80 landslides in the DNN model. I was found that the landslide distribution results of the CNN model are more reasonable than that of the DNN model when comparing the historical landslide distribution of the landslide susceptibility map of the two models.

Table 3 Statistical results of landslide density of the CNN and the DNN models

Model	Landslide susceptible zones	Area of zones (%)	Number of landslides	Landslides percentage (%)	Landslide density (number/km^2)
CNN	Very high	10	204	34.75	0.0400
	High	10	133	22.66	0.0261
	Middle	10	84	14.31	0.0165
	Low	20	96	16.35	0.0094
	Very low	50	70	11.93	0.0027
DNN	Very high	10	150	25.55	0.0294
	High	10	133	22.66	0.0261
	Middle	10	93	15.84	0.0183
	Low	20	131	22.32	0.0129
	Very low	50	80	13.63	0.0031

Discussion

Landslide susceptibility maps are essential for decision-makers to formulate reasonable policies and reduce the impact of landslides[37]. Therefore, it is of great significance to obtain high-quality landslide susceptibility maps. However, with insufficient data, these machine learning models often suffer from generalizing to areas other than the training area. Especially in landslide susceptibility mapping, gathering inventory data is expensive, and it is difficult to collect a complete list of landslides. In order to solve this problem, this paper uses InSAR Stacking technology to identify early landslide hazards in the Ya'an-Linzhi section of the Sichuan-Tibet Railway and uses the identified landslide and historical landslide data as modeling data to evaluate the landslide susceptibility. Zhao et al.[38] used a combination of landslide data identified by InSAR Stacking technology and historical landslides to map landslide susceptibility. The study found that the optimized results of InSAR Stacking technology were more reliable than the results of only the historical landslides. The slopes deformation identified by the InSAR Stacking method is usually a precursor to the occurrence of landslides. In time series analysis, the slope deformation rate is an accelerating process. It usually indicates the occurrence of a landslide. Therefore, InSAR Stacking deformation monitoring results can provide an important basis for early identification and susceptibility evaluation of landslides and make the results of model predictions more reliable.

The choice of prediction model has an important influence on the results of landslide susceptibility evaluation. Some scholars have conducted comparative studies on the application of deep learning and traditional machine learning in the evaluation of landslide susceptibility, and found that deep learning has higher predictive capabilities[39-41]. Therefore, this study uses CNN and DNN models to evaluate landslide susceptibility in the study area and compares the two models. Comparing the training and prediction accuracy of the two models using AUC and a confusion matrix, the results show that in this study area, the CNN model has a higher success and prediction rate than the DNN model, and the distribution of the landslide hazards in the susceptible areas results in a more reasonable CNN model. Although the two models have good predictive capabilities, the key parameter values of the two models are determined by trial and error; thus, the parameters determined by this method may not be the best model. Therefore, to compare the performance of the two models more accurately, it is necessary to conduct numerous research on the determination of the models' parameters, and choose different research areas, that is, different geological environments and sample data for the comparative research. In the future, our research will further explore the application potential of deep learning techniques in the evaluation of landslide susceptibility.

Conclusions

In this study, two well-known deep learning algorithms, namely CNN and DNN based models, were applied to generate a landslide susceptibility map of the Ya'an-Linzhi section of the Sichuan-Tibet Railway, and simultaneously, combined with the application of InSAR Stacking technology to identify hidden danger points of landslides. A complete list of landslides improves the accuracy of landslide susceptibility evaluation. The results show that the two models have a higher success rate and prediction performance in this study area, but the CNN algorithm showed a 4% higher performance than DNN. According to the analysis of the landslide influencing factors in the study area, it was found that slope, elevation, and rainfall are the main influencing factors that affect the occurrence of landslides. High and very high landslide susceptibility were primarily distributed in the Jinsha, Lancang, and Nujiang River Basins along the railway, which can better reflect the distribution of landslide susceptibility in the study area, providing a scientific basis for the disaster prevention and mitigation work of the Ya'an-Linzhi section of the Sichuan-Tibet Railway.

Methods

There are four main stages using the CNN and DNN models for landslide susceptibility mapping: (1) the establishment of a spatial database, including InSAR Stacking technology and field survey to generate a list of landslides, as well as selecting the landslide impact factors; (2) assessing data accuracy and removing noisy data with null prediction power; (3) use CNN and DNN models to generate landslide susceptibility maps; (4) validation and comparison of the two models (Fig. 8).

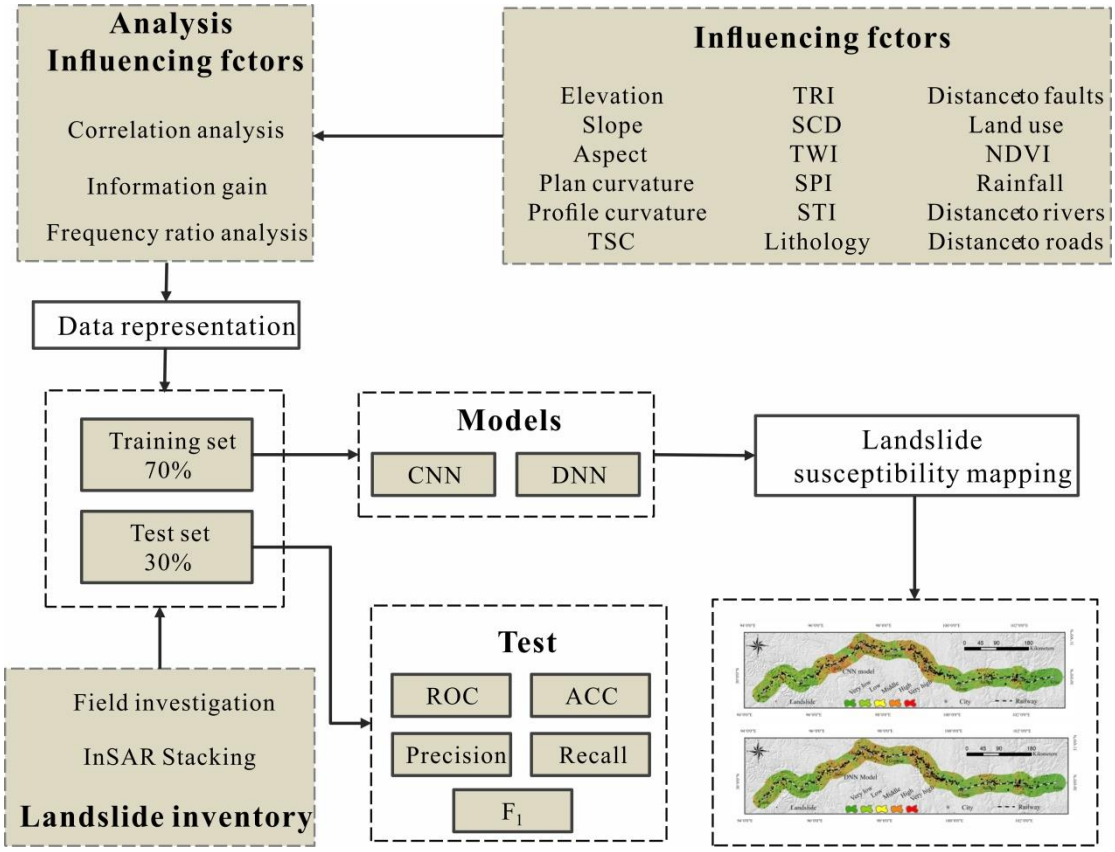

Fig. 8 Methodological flow chart of the study.

Landslide inventory map

Landslide inventory maps are prepared for multiple scopes, which is the first step toward modeling landslide susceptibility[42,43]. This study, combines the results of InSAR Stacking deformation[44], using Google Earth satellite images and field surveys to prepare a landslide inventory map (Fig. 1). Consequently, a total of 587 landslides were identified in the inventory map. The area of landslides in the study area is 691 km^2, and the largest and smallest landslides are 19688941 m^2 and 1152 m^2, respectively. The scale distribution of landslides is mainly small shallow surface landslides. Fig. 9 shows the hidden danger points of landslides near the Jinsha River Bridge identified by InSAR Stacking technology.

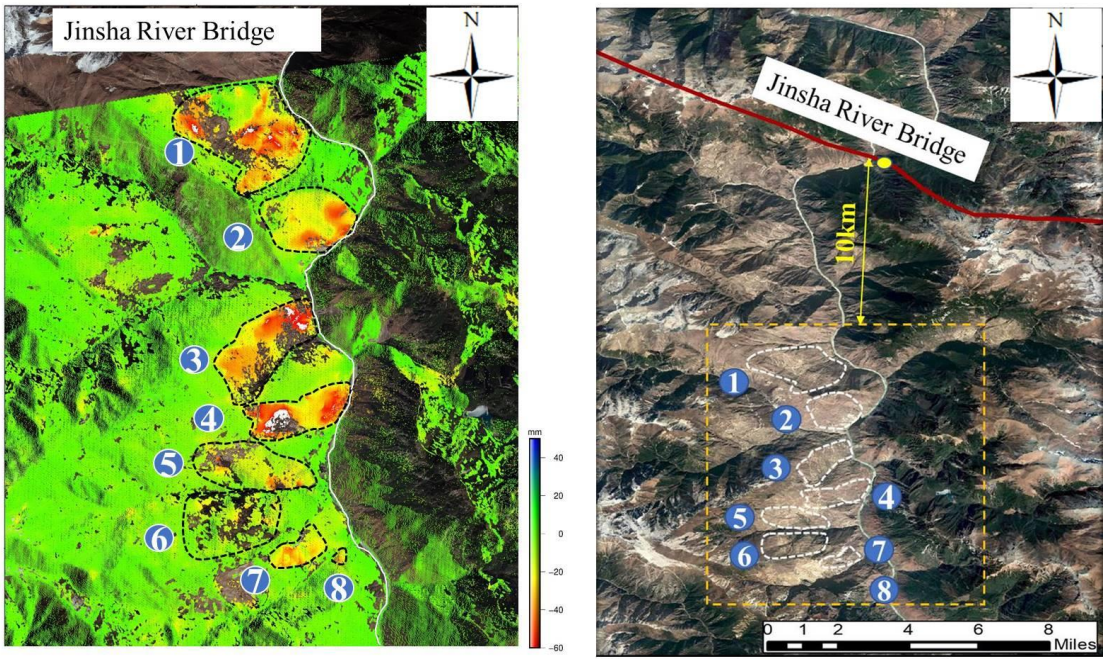

Fig. 9 Hidden danger points of landslide identified by InSAR Stacking technology

Using machine learning methods to model landslides is typical a binary classification[45]. Therefore, it is necessary to use positive samples (landslides) and negative samples (non-landslides) for modeling. Within the inventory map, 587 pixels of landslide occurrences have been extracted. In this study, in order to avoid the error rate of non-landslide selection, an equal number of non-landslide points were randomly selected out of landslide buffer area. The landslide and non-landslide samples were randomly selected 410 (70%) for model training, and the remaining 177 (30%) were used for model testing.

Landslide influencing factors

The causes of landslide occurrences are complex, and their mechanisms remain under debate. Generally, landslides result from a combination of internal geological conditions and external environmental factors[46,47]. Internal factors include topography, stratigraphic lithology, geological structure, and tectonic movement[48]. The external factors of landslides can be divided into human and natural factors. Natural factors include meteorological hydrology, hydrogeology, weathering, and new tectonic movements[49]. Human factors refer to human engineering activities, including constructing roads, buildings, factories, and mining of minerals. In general, the external factors are the inducing factors of landslide occurrence[50,51].

In this study, according to the data availability, geo-environmental conditions, as well as landslide occurrence mechanisms of the study area, eighteen landslide conditioning factors including elevation, slope, aspect, plan curvature, profile curvature, terrain surface convexity (TSC), terrain ruggedness index (TRI), surface cutting degree (SCD), topographic wetness index (TWI), stream power index (SPI), sediment transport index (STI), lithology, distance to faults, land use, normalized difference vegetation index (NDVI), rainfall, distance to rivers, and distance to roads (Fig. 10). Since the 18 factors are represented on different intervals or scales, all factors are converted into a grid with DEM resolution (30m×30m) for unification. Furthermore, all factor datasets can be divided into either continuous or discrete datasets. Continuous dataset of each factor was reclassified into discrete subclasses with data in specific intervals using a manual

method; discrete datasets of the rest factors were classified using the original natural grouping. The detailed information of the classes of each landslide conditioning factor is shown in Table 4.

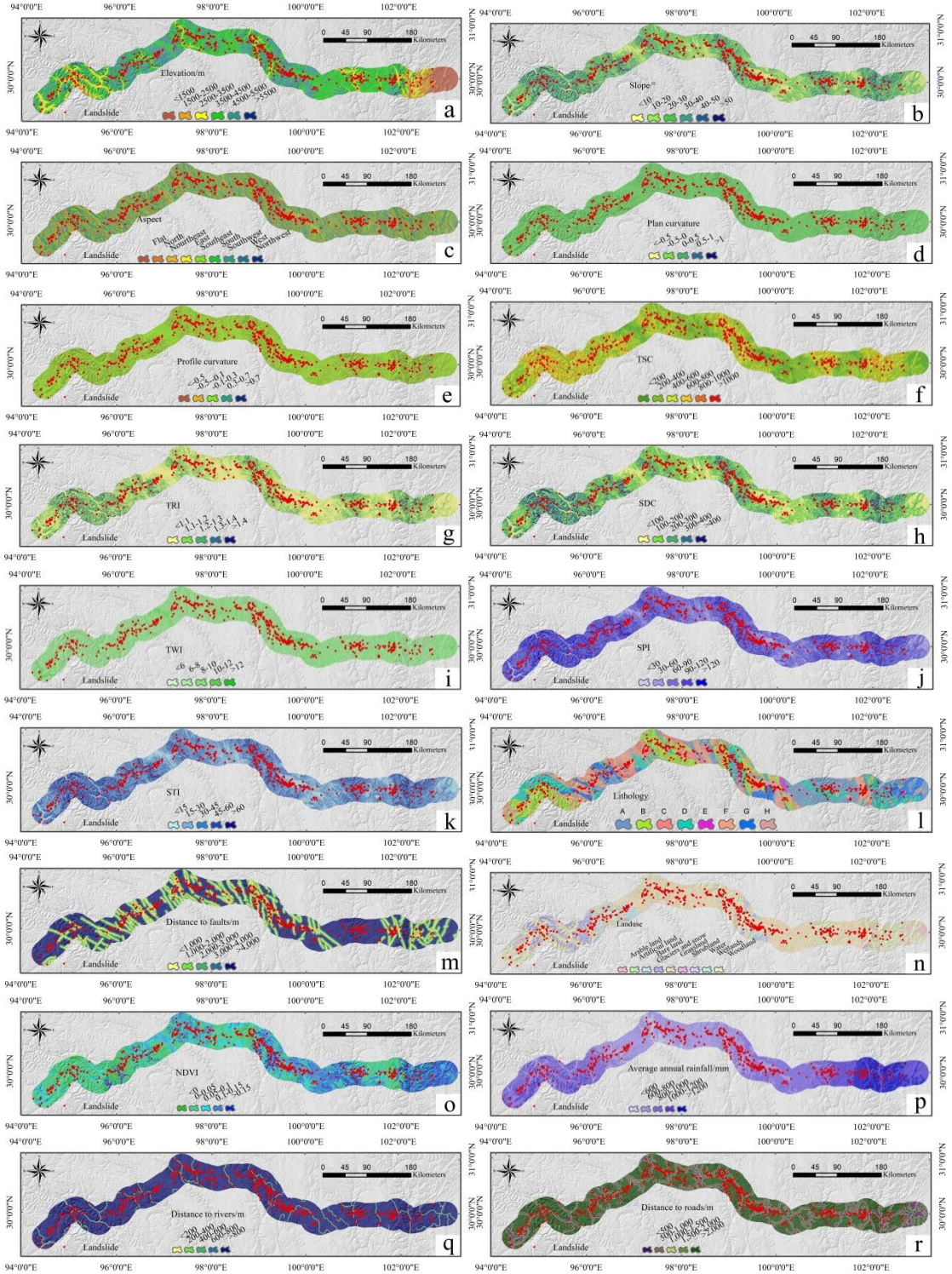

Fig. 10 Landslide influencing factor maps. (a) Elevation, (b) Slope, (c) Aspect, (d) Plan curvature, (e) Profile curvature, (f) TSC, (g) TRI, (h) SCD, (i) TWI, (j) SPI, (k) STI, (l) Lithology, (m) Distance to faults, (n) Land use, (o) NDVI, (p) Average annual rainfall, (q) Distance to rivers, (r) Distance to roads

Table 4 Influencing factors categories of landslides

Factions	Classification standard	Type
Altitude/m	<1500; 1500-2500; 2500-3500; 3500-4500; 4500-5500; 5500<	Continuous
Slope/°	<10; 10-20; 20-30; 30-40; 40-50; 50<	Continuous
Aspect	F (–1); N (0–22.5; 337.5–360); NE (22.5–67.5); E (67.5–112.5); SE (112.5–157.5); S (157.5–202.5); SW (202.5–247.5);W (247.5–292.5); NW (292.5–337.5)	Categorical
Plan curvature	<-0.5; -0.5-0; 0-0.5; 0.5-1; 1<	Continuous
Profile curvature	<-0.5; -0.5- -0. 1; -0.1- 0.3; 0.3-0.7; 0.7<	Continuous
TSC	<200; 200-400; 400-600; 600-800; 800-1000; 1000<	Continuous
TRI	<1.1; 1.1-1.2; 1.2-1.3; 1.3-1.4; 1.4<	Continuous
SCD	<100; 100-200; 200-300; 300-400; 400<	Continuous
TWI	<6; 6-8; 8-10; 10-12; 12<	Continuous
SPI	<30; 30-60; 60-90; 90-120; 120<	Continuous
STI	<15; 15-30; 30-45; 45-60; 60<	Continuous
Lithology	A; B; C; D; E; F; G; H	Categorical
Distance to faults	<1000; 1000-2000; 2000-3000; 3000-4000; 4000<	Continuous
Land use	Arable land; Artificial land; Bare land; Glaciers and snow; Grassland; Shrubland; Water; Wetlands; Woodland	Categorical
NDVI	<0; 0-0.05; 0.05-0.10; 0.10-0.15; 0.15<	Continuous
Average annual rainfall/mm	<600; 600-800; 800-1000; 1000-1200; 1200<	Continuous
Distance to rivers /m	<200, 200-400; 400-600; 600-800; 800<	Categorical
Distance to roads /m	<500, 500-1000; 1000-1500; 1500-2000; 2000<	Categorical

Each factor has a different effect on the occurrence of landslides. Elevation has a significant impact on landslide development and determines the potential energy of the landslide; it also affects the movement characteristics of the landslides[52]. An increasing slope angle will cause the increasing size of the free face and shear strength on the potential slide surface, resulting in slope failure[53]. The slope aspect determines the illumination time received by the slope surface. There are differences in surface humidity, vegetation coverage, and different slope aspects, which affect the distribution of pore water pressure and the physical and mechanical characteristics of rock and soil masses[54]. Plan curvature affects convergence and divergence of flow. Profile curvature has great significance on the acceleration and deceleration of flow providing valuable information about erosion and deposition[55,56]. TSC describes the relief characteristics of the terrains surface[57]. TRI is a measure of the roughness and brokenness of the ground. The larger the roughness means the ground is broken, and the loose deposits are richer, which is conducive to the occurrence of a landslide[58]. SCD refers to the difference between the average and minimum value of the elevation of a point on the ground in a specific area, reflecting the degree to which the ground surface is cut[59]. TWI comprehensively analyzes the influence of topographical features on the spatial distribution of soil moisture[60]. SPI indicates the erosion power of streams which might affect landslide occurrences[61]. STI describes topographic variables of water and sediment transport in landslides[52]. Lithology is one of the basic factors affecting the occurrence of landslides[62,63].

According to the hardness and type of lithology, it is divided into the following eight groups, A (Harder sandstone, siltstone), B (Weaker gneiss, phyllite, mudstone), C (Soft and hard limestone interbedded with sandstone), D (Harder quartz sandstone, feldspar quartz sandstone), E (Hard basalt, ophiolite, syenite), F (Hard granite, diorite), G (Soft and hard silty slate, conglomerate sandstone), and H (Weak loose deposits). Faults have an important influence on the strength of the rock mass, the development of the terrain structure and the slope's stability[64]. Land use influences slope stability by changing land use and disturbing the slope stability conditions[65]. NDVI represents vegetation coverage and groundwater content, which may affect the development of landslides[66]. Rainfall causes a large amount of rainwater to infiltrate, saturating the soil layer on the slope, increasing the weight of the sliding body, thus causing the occurrence of landslides[67]. When building roads, natural slopes must be excavated and repaired, which will inevitably interfere with the balanced conditions of the original slope, often leading to unstable slopes and landslides[68]. Distance to rivers is one of the conditioning factors that has an effective role in landslide stability. The wet saturated water of the river acting on the sliding area and part of the sliding body may reduce the shear strength of the soil and weaken the layers, thus, reducing the stability of the landslide[69,70].

Influencing factor evaluators

Correlation analysis. When evaluating landslide susceptibility, it is essential that the influencing factors need to maintain mutual independence. If there appears to be a strong linear correlation in the aforementioned factors, then the predisposing factors are assumed to exist within a multicollinearity problem. The multicollinearity problem will affect the accuracy of the training model and may lead to errors in the prediction results[71,72]. In this paper, a Pearson correlation coefficient is used to analyze the correlation between the influencing factors. Its value ranges from -1 to 1. -1 means that the two variables are completely negatively correlated, 1 means that the two variables are completely positively correlated, and 0 means that they are not correlated. When the absolute value of the correlation coefficient between two factors is greater than 0.7, it is considered to have a high correlation[73,74].

Information gain. In this study, the feature selection method of information gain (IG) was used to select an optimal subset to improve the prediction performance in the evaluation of landslide susceptibility[75]. The information gain is determined by calculating the entropy reduction of the output category y to which the input factor x_i corresponds.

$$IG(y, x_i) = E(y) - E(y|x_i) \quad (1)$$

where $E(y|x_i)$ is the conditional entropy and $E(y)$ is a priori Shannon entropy, and they are calculated as follows:

$$E(y) = -\sum_{i=1}^{n} y_i \log_2(y_i) \quad (2)$$

$$E(y|x_i) = -\sum_{i=1}^{n} y_i E(y) \quad (3)$$

The average merit (AM) derived from this method uncovers the importance between conditioning factors and landslide occurrence. The greater the weight, the greater the contribution of the corresponding factors to the occurrence of landslides. If this value is less than or equal to 0, then this influencing factor has nothing to do with the occurrence of landslides and should be excluded when making predictions.

Frequency ratio analysis

The frequency ratio (FR) method can be employed to evaluate the correlation between landslide occurrence and influencing factors. In landslide susceptibility analysis, it is perceived that future landslides will occur under the same conditions as past landslides. The FR can be calculated as follows[55]:

$$FR = \frac{N/N\prime}{A/A\prime} \qquad (4)$$

where N is the number of each factor's landslide; $N'$is the number of total landslides; A is the number of pixels in a particular class; and A' is the number of total pixels.

Landslide susceptibility models

CNN. A CNN model exhibiting robust performance in visual image analysis is a class of feed-forward neural network whose artificial neurons respond to a portion of the surrounding elements[76]. The general CNN model structure includes an input layer, convolutional layers, maximum pooling, fully connected layers, and an output layer[77], as shown in Fig. 11. The convolutional layer uses a sliding convolution window method to extract features from the input layer. The first convolutional layer usually extracts some low-level features, and more layers of the convolutional layers can iteratively extract higher-level features from low-level features. The output of the convolutional manipulation is defined as follows:

$$C_j = \sum_i^N f(w_j * v_i + b_j), j = 1,2,3 \ldots, k \qquad (5)$$

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad (6)$$

where N is the number of factors affecting the landslide, f represents a nonlinear activation function and * denotes the convolutional operator, k is the number of convolutional kernels, and w_j and b_j denotes the weight and bias, respectively.

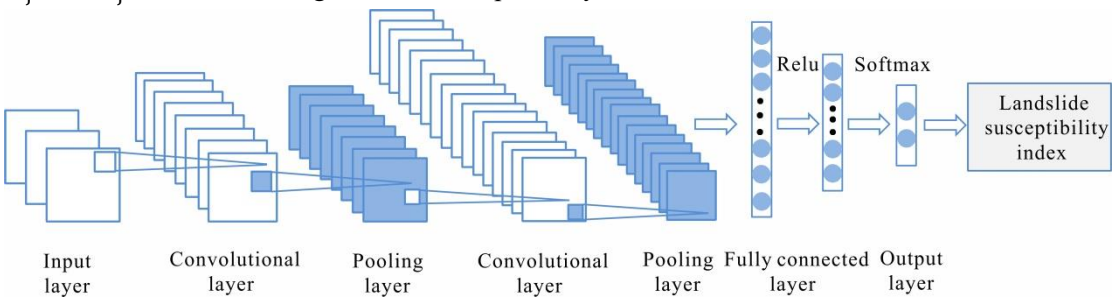

Fig. 11 Convolutional neural network (CNN) structure.

The pooling layer is used to realize the sample processing of the feature map, which can reduce the amount of data while retaining useful information, preventing over-fitting and improving the generalization ability of the model. Next, these local representations extracted by the convolutional and pooling operations are reorganized through the fully connected layers. Finally, the fully connected layer is connected to the output layer, which consists of two neurons representing landslide and non-landslide. The parameters in the CNN layer are optimized using the back-propagation algorithm.

DNN. DNN is the basic algorithm of deep learning[78]. The general DNN model structure includes an input layer, several hidden layers, and an output layer. In this architecture, the neurons (nodes) in the previous layer are completely connected to all the neurons in the next dense layer. Afterward, more dense layers are added to extract hidden information in the learning process. Fig.

12 presents the architecture of DNN. The basic processes of deep NN mechanisms are as follows: (1) the network correctly assigns inputs to their associated targets, (2) introduce a loss function to calculate the prediction and true target of the network, and (3) the training loop is repeated enough times to generate weights that minimize the loss function.

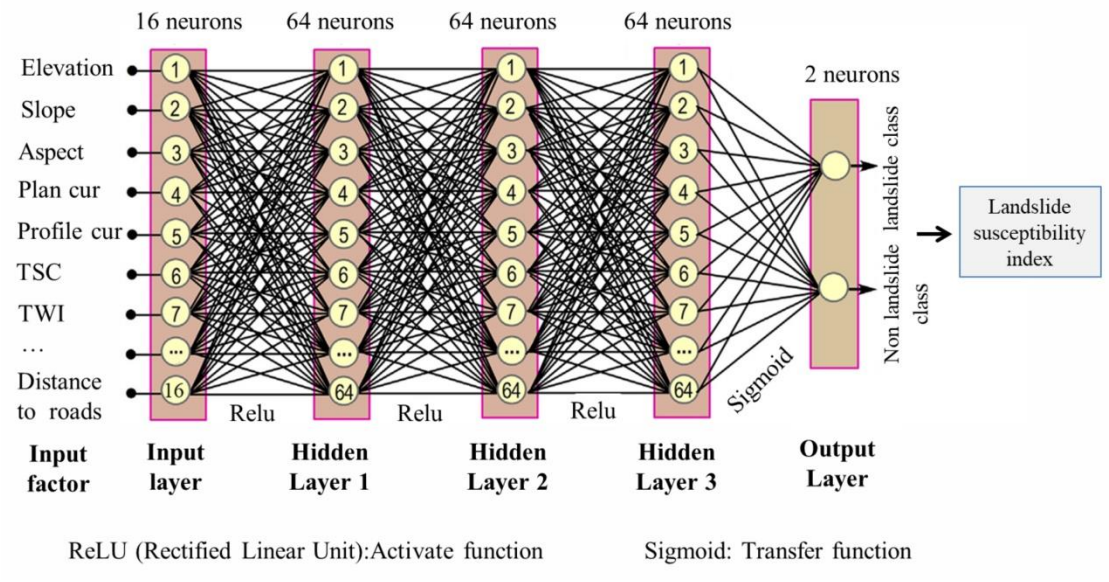

Fig. 12 Deep neural network (DNN) structure

In this study, the DNN model was applied to the evaluation of landslide susceptibility. The impact factor became the input signal received in the first layer and analyzed in the hidden layer. Finally, the prediction results are displayed in the output layer as landslide and non-landslide. The structure of the DNN model was determined through several trial and error methods, which consisted of a model of three hidden layers, including 16 neurons, two output neurons, and 3 hidden layers of 64 neurons.

Model evaluation methods

In landslide susceptibility mapping, it is essential and necessary to validate the model's performance. In this paper, the receiver operating characteristic curve (ROC) is used to evaluate the model's training and prediction accuracy. The ROC curve is an indicator of the continuous variables of data specificity and sensitivity. The area under the ROC curve (AUC) represents the accuracy of the model; the closer the AUC is to 1, the better the model performance[56,79]. Simultaneously, a confusion matrix was used to evaluate the performance of the two models. Statistical indices including accuracy (ACC), recall, precision, and F-measure (F1) were acquired from the confusion matrix. The calculation is as follows:

$$\mathrm{ACC} = \frac{\mathrm{TP+TN}}{\mathrm{TP+FP+TN+FN}} \qquad (7)$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (8)$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \qquad (9)$$

$$\mathrm{F_1} = \frac{2\times\mathrm{TP}}{2\times\mathrm{TP+FP+FN}} \qquad (10)$$

Where TP (True Positive) and TN (True Negative) are the numbers of correctly classified landslides, FP (false positive) and FN (false negative) is the numbers of landslides incorrectly classified. For ACC, recall, precision, and F1, these values are between 0 and 1. With increasing numbers, the model's performance improves.

Acknowledgement

This study was financially supported by the National Natural Science Foundation of China (No. 41941019), National Key Research and Development Program of China (No. 2020YFC1512000) and National Natural Science Foundation of China (No. 41922054）.

Author contributions

S. B W., J. Q Z. put forward the concept and method of research and wrote the paper. J. Q M., J. Z., J. W Z., J. W., Y. T F. participated in data analysis.

Competing interests

The authors declare no competing interests.

Reference

1. Kjekstad, O., Highland, L. Economic and Social Impacts of Landslides. *Springer Berlin Heidelberg.* https://doi.org/10.1007/978-3-540-69970-5_30 (2009).
2. Huang, R., Yan, R., Qu, K., Wang, K. Engineering geological assessment for route selection of railway line in geologically active area: A case study in China. *Journal of Mountain Science*. 10(4),495-508 (2013).
3. Peng, J., Cui, P., Zhuang, J. Challenges to engineering geology of Sichuan—Tibet railway. *Chinese Journal of Rock Mechanics and Engineering.* (12),2377-2389 (2020). (in Chinese)
4. Lan, H., Zhang, N., Li, L., Tian, N., Zhang, Y., Liu, S., Lin, G., Tian, C., Wu, Y., Yao, J., Peng, J., Zhou, C. Risk analysis of major engineering geological hazards for Sichuan—Tibet railway in the phase of feasibility study. *Journal of Engineering Geology*. 29(02),326-341 (2021). (in Chinese)
5. Guo, C., Zhang, Y., Jiang, L., Meng, W., Yuben, DU., Chuntian, MA. Discussion on the environmental and engineering geological problems along the Sichuan-Tibet railway and its adjacent area. *Geoscience.* 31(5),4-16 (2017). (in Chinese)
6. Guo, C., Wu, R., Jiang, W., Zhong, N., Wang, Y., Wang, D., Zhang, Y., Yang, Z., Meng, W., Li, X., Liu, G. Typical geohazards and engineering geological problems along the Ya'an-Linzhi section of the Sichuan-Tibet railway, China. *Geoscience.* 35(01),1-17 (2021).
7. Zhou, J., Cui, P., Hao, M. Comprehensive analyses of the initiation and entrainment processes of the 2000 Yigong catastrophic landslide in Tibet, China. *Landslides.* 13(1),39-54 (2016).
8. Ouyang, C., An, H., Zhou, S., Wang, Z., Su, P., Wang, D., Cheng, D., She, J. Insights from the failure and dynamic characteristics of two sequential landslides at Baige village along the Jinsha River, China. *Landslides.* 16,1397-1414 (2019).
9. Zhang, L., Xiao, T., He, J. et al Erosion-based analysis of breaching of Baige landslide dams on the Jinsha River, China, in 2018. *Landslides.* 16,1965-1979 (2019).
10. Shirzadi, A., Bui, DT., Pham, BT., Solaimani, K., Chapi, K., Kavian, A., Shahabi, H., Inge, R. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ Earth Sci.* 76,60 (2017).
11. Pham, BT., Bui, DT., Pourghasemi, HR., Indar, P., Dholakia, MB. Landslide susceptibility assesssment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theor Appl Climatol.* 128,255-273 (2017).
12. Ayalew, L., Yamagishi, H. The application of gis-based logistic regression for landslide susceptibility mapping in the

kakuda-yahiko mountains, central Japan. *Geomorphology.* 65(1/2),15-31 (2005).

13. Zhao, Z., Chen, W., Pourghasemi, HR. A gis-based comparative study of dempster-shafer, logistic regression and artificial neural network models for landslide susceptibility mapping. *Geocarto International.* https://doi.org/10.1080/10106049.2016.1140824 (2017).
14. Yilmaz, I. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from kat landslides (tokat—turkey). *Computers & Geosciences.* 35(6),1125-1138 (2009).
15. Lucchese, LV., Oliveira, G., Pedrollo, OC. Investigation of the influence of nonoccurrence sampling on landslide susceptibility assessment using artificial neural networks. *Catena.* 198,105067 (2021).
16. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using gis. *Computers & Geosciences.* 51(2),350-365 (2012).
17. Bui, DT., Pradhan, B., Lofman, O., Revhaug, I. Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and naive bayes models. *MATH PROBL ENG.* https://doi.org/10.1155/2012/974638 (2012).
18. Chen, W., Li, X., Wang, Y., Chen, G., Liu, S. Forested landslide detection using lidar data and the random forest algorithm: a case study of the three gorges, china. *Remote Sensing of Environment.* 152, 291-301 (2014).
19. Wang, S., Zhuang, J., Zheng, J., Fan, H., Kong, J., Zhan, J. Application of Bayesian Hyperparameter Optimized Random Forest and XGBoost Model for Landslide Susceptibility Mapping. *Front. Earth Sci*. 9:712240. https://doi.org/10.3389/feart.2021.712240 (2021).
20. Yao, X., Tham, L., Dai, F. Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. *Geomorphology.* 101(4), 572-582 (2008).
21. Yu, L., Cao, Y., Zhou, C., Wang, Y., Huo, Z. Landslide susceptibility mapping combining information gain ratio and support vector machines: a case study from wushan segment in the three gorges reservoir area, China. *Applied Sciences.* 9(22):4756 (2019).
22. Choubin, B., Zehtabian, G., Azareh, A., Rafiei-Sardooi, E., Sajedi-Hosseini F., Kisi O. Precipitation forecasting using classification and regression trees (cart) model: a comparative study of different approaches. *Environmental Earth Sciences.* 77(8):314 (2018).
23. Andrieu, C., De Freitas, N., Doucet, A., Jordan, MI., An introduction to MCMC for machine learning. *Machine Learning*. 50:5-43 (2003).
24. Huang, F., Yao, C., Liu, W., Li, Y., Liu, X. Landslide susceptibility assessment in the Nantian area of China: A comparison of frequency ratio model and support vector machine. Geomat. *Nat. Hazards Risk*. 9:919-938 (2018).
25. Hao, X., Zhang, G., Ma, S. Deep learning. *International Journal of Semantic Computing.* 10(03):417-439 (2016).
26. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J. Deep learning and process understanding for data-driven Earth system science. *Nature.* 566:195-204 (2019).
27. Yang, Z., Tao, D., Zhang, S., Jin, L. Similar handwritten chinese character recognition based on deep neural networks with big data. *Journal on Communications*. https://doi.org/10.3969/j.issn.1000-436x.2014.09.019 (2014).
28. Chen, Z., Li, C., Sanchez, RV. Gearbox fault identification and classification with convolutional neural networks. *Shock and Vibration*. 2015(PT.5):1-10 (2015).
29. Huynh, BQ., Li, H., Giger, ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *Journal of Medical Imaging*. 3(3):034501 (2016).
30. Wegmayr, V., Aitharaju, S., Buhmann, J. Classification of brain MRI with big data and deep 3D convolutional neural networks. *Computer-Aided Diagnosis.* https://doi.org/10.1117/12.2293719 (2018).
31. Hong, SH., Wdowinski, S., Kim, SW. Small Temporal Baseline Subset (STBAS): A New InSAR Technique for Multi-Temporal Monitoring Wetland's Water Level Changes. *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium. IEEE*. https://doi.org/10.1109/IGARSS.2008.4779406 (2009).
32. Luo, H., Zuo, L., Chen, Y., Chenal, Y., Chenal, J. Monitoring Surface Deformation of Transmission Corridors in Mountain Areas

Based on SBAS-INSAR. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. pp. 2058-2061. https://doi.org/10.1109/IGARSS.2019.8899307 (2019).

33. Wang, G., Wang, Y., Zang, X., Zhu, J., Wu, W. Locating and monitoring of landslides based on small baseline subset interferometric synthetic aperture radar. *Journal of Applied Remote Sensing.* 13(4):1 (2019).
34. Guo, R., Sumin, L., I, Chen, Y., Li, X., Yuan, L. Identification and monitoring landslides in longitudinal range-gorge region with insar fusion integrated. *Landslides*. 18, 551–568 (2021).
35. Chen, W., Li, Y. Gis-based evaluation of landslide susceptibility using hybrid computational intelligence models. *Catena.* 195:104777 (2020).
36. Huang, F., Cao, Z., Guo, J., Jiang ,S., Li, S., Guo, Z. Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. *Catena.* 191:104580 (2020).
37. Wang, Y., Feng, L., Li, S., Ren, F., Du, Q. A hybrid model considering spatial heterogeneity for landslide susceptibility mapping in zhejiang province, china. *Catena.* 188:104425 (2020).
38. Zhao, F., Meng, X., Zhang, Y., Chen, G., Yue, D. Landslide susceptibility mapping of karakorum highway combined with the application of sbas-insar technology. *Sensors.* 19(12):2685 (2019).
39. Bui, DT., Tsangaratos, P., Nguyen, VT., Liem, NV., Trinh, PT. Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment. *Catena.* 188(2020):104426 (2020).
40. Dong, VD., Jaafari, A., Bayat, M. et al. A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. *Catena.* https://doi.org/10.1016/j.catena.2019.104451 (2020).
41. Mandal, K., Saha, S., Mandal, S. Applying deep learning and benchmark machine learning algorithms for landslide susceptibility modelling in rorachu river basin of sikkim himalaya, india. *Geoscience Frontiers.* https://doi.org/10.1016/j.gsf.2021.101203 (2021).
42. Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P . Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology.* 31(1):181-216(1999).
43. Rosi, A., Tofani, V., Tanteri L, Tacconi, S.C., Agostini, A., Catani, F., Casali, N. The new landslide inventory of Tuscany (Italy) updated with PSInSAR: geomorphological features and landslide distribution. *Landslides.* 15:5-19(2018).
44. Zhang, C., Li, Z., Yu, C . Landslide detection: Application of InSAR Stacking in the Jinsha River Basin with the aid of GACOS. *Geomatics and Information Science of Wuhan University.* https://doi.org/doi:10.13203/j.whugis20200675(2021).
45. Wang, Y., Fang, Z., Wang, M., Ling, P., Hong, H. Comparative study of landslide susceptibility mapping with different recurrent neural networks. *Computers & Geosciences.* 138:104445(2020).
46. Zezere, J. L., Ferreira, A. D. B., Rodrigues, M. L. The role of conditioning and triggering factors in the occurrence of landslides: a case study in the area north of lisbon (portugal). *Geomorphology.* 30(30):133-146 (1999).
47. Xiang, L., Cui, P., Zhang, J., Huang, D., Zhou, X. Triggering factors susceptibility of earthquake-induced collapses and landslides in wenchuan county. *Sichuan Daxue Xuebao (Gongcheng Kexue Ban)/Journal of Sichuan University (Engineering Science Edition).* 42(5):105-112(2010).
48. Yin, J., Chen, J., Xu, X., Wang, X., Zheng, Y. The characteristics of the landslides triggered by the wenchuan m-s 8.0 earthquake from anxian to beichuan. *Journal of Asian Earth Sciences*. 37(5-6):452-459 (2010).
49. Capitani, M., Ribolini, A., Federici, P. R. Influence of deep-seated gravitational slope deformations on landslide distributions: a statistical approach. *Geomorphology.* 201:127-134 (2013).
50. Huang, R. Large-scale landslides and their sliding mechanisms in china since the 20th century. *Chinese Journal of Rock Mechanics and Engineering.* 26(3):433-454 (2007). (in chinese)
51. Zhuang, J., Peng, J. A coupled slope cutting—a prolonged rainfall-induced loess landslide: a 17 october 2011 case study. *Bulletin of Engineering Geology & the Environment*. 73(4):997-1011(2014).
52. Pourghasemi, H. R., Mohammady, M., Pradhan, B. Landslide susceptibility mapping using index of entropy and conditional

probability models in gis: safarood basin, Iran. *Catena.* 97(15):71-84(2012).

53. Zhang, K., Wu, X., Niu, R., Yang, K., Zhao, L. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the three gorges reservoir area, China. *Environmental Earth Sciences.* 76(10):405(2017).
54. Pourghasemi, HR., Pradhan, B., Gokceoglu, C. Application of fuzzy logic and analytical hierarchy process (ahp) to landslide susceptibility mapping at haraz watershed, Iran. *Natural Hazards.* 63(2):965-996(2012).
55. Chen, W., Li, W., Hou, E., Zhao, Z., Deng, N., Bai, H., Wang, D. Landslide susceptibility mapping based on gis and information value model for the chencang district of baoji, China. *Arabian Journal of Geosciences.* 7(11):4499-4511(2014).
56. Wu, Y., Ke, Y., Chen, Z., Liang, S., Hong, H. Application of alternating decision tree with adaboost and bagging ensembles for landslide susceptibility mapping. *CATENA.* 187:104396(2020).
57. Tolga, G. Tectonic, topographic and rock-type influences on large landslides at the northern margin of the Anatolian Plateau. *Landslides.* 16:333-346(2019).
58. Korzeniowska, K., Pfeifer, N., Landtwing, S. Mapping gullies, dunes, lava fields, and landslides via surface roughness. *Geomorphology.* S0169555X16308108. https://doi.org/10.1016/j.geomorph.2017.10.011 (2018).
59. Zhang, Q., Ling, S., Li, X., Sun, C., Xu, J., Huang, T. Comparison of landslide susceptibility mapping rapid assessment models in Jiuzhaigou County, Sichuan province, China. *Chinese Journal of Rock Mechanics and Engineering*. 39(08):1595-1610 (2020). (in Chinese)
60. Mousa, A., Bahareh, G., Ataollah, S., Himan, S., Kamran, C., Thai, P. B., Baharin, B. A., Dieu, T. B. A novel hybrid approach of bayesian logistic regression and its ensembles for landslide susceptibility assessment. *Geocarto International.* 1-44(2018).
61. Moore, I. D., Burch, G. J. Physical basis of the length-slope factor in the universal soil loss equation. *Soil Science Society of America Journal.* 50(5) (1986).
62. He, S., Peng, P., Lan, D., Wang, H., Liu, J. Application of kernel-based fisher discriminant analysis to map landslide susceptibility in the Qinggan river delta, Three Gorges, China. *Geomorphology* . s171-172:30-41. https://doi.org/10.1016/j.geomorph.2012.04.024 (2012).
63. Nsengiyumva, JB., Luo, G., Hakorimana, E., Mind'Je, R., Gasirabo, A., Mukanyandwi, V. Comparative analysis of deterministic and semiquantitative approaches for shallow landslide risk modeling in rwanda. *Risk Analysis.* https://doi.org/10.1111/risa.13359 (2019).
64. Pham, BT., Bui, DT., Indra, P., Dholakia, MB. Landslide susceptibility assessment at a part of uttarakhand himalaya, India using gis – based statistical approach of frequency ratio method. *International Journal of Engineering & Technical Research.* V4(11):338-344 (2015).
65. Pham, BT., Bui, DT., Prakash, I., Dholakia, MB. Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using gis. *Nat Hazards.* 83:97-127 (2016).
66. Hong, H., Chen, W., Xu, C., Youssef, AM., Pradhan, B., Bui, DT. Rainfall-induced landslide susceptibility assessment at the chongren area (china) using frequency ratio, certainty factor, and index of entropy. *Geocarto International.* https://doi.org/10.1080/10106049.2015.1130086 (2017).
67. Zhuang, J., Peng, J., Wang, G., Lqbal, J., Zhu, X. Prediction of rainfall-induced shallow landslides in the loess plateau, yan'an, China, using the trigrs model. *Earth Surface Processes and Landforms.* 42(6) (2016).
68. Pham, BT., Bui, DT., Prakash, I. Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression and j48 decision trees methods: a comparative study. *Geotech Geol Eng.* 35:2597-2611 (2017).
69. Ding, Q., Chen, W., Hong, H. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. *Geocarto International.* 32(6):619-639 (2017).
70. Erener, A., Sivas, AA., Selcuk-Kestel, AS., Düzgün, HS. Analysis of training sample selection strategies for regression-based quantitative landslide susceptibility mapping methods. *Computers & Geosciences.* 104(JUL.):62-74 (2017).
71. Wang, Y., Fang, Z., Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in yanshan county, china. *Science of The Total Environment.* 666(MAY 20):975-993 (2019).

72. Chen, X., Chen, W. Gis-based landslide susceptibility assessment using optimized hybrid machine learning methods. *Catena.* 196:104833 (2021).

73. Martín. B., Alonso, JC., Martín, CA., Palacín, C., Magaña, M., Alonso, J. Influence of spatial heterogeneity and temporal variability in habitat selection: a case study on a great bustard metapopulation. *Ecological Modelling.* 228:39-48 (2012).

74. Hong, H., Liu, J., Zhu, A. Modeling landslide susceptibility using logitboost alternating decision trees and forest by penalizing attributes with the bagging ensemble. *Science of The Total Environment*. 718(18):137231 (2020).

75. Dash, M., Liu, H. Feature selection for classification. Intelligent Data Analysis. 1(1-4):131-156 (1997).

76. Girshick, R. Fast r-cnn. *Computer Science.* https://doi.org/10.1109/ICCV.2015.169 (2015).

77. Shin, HC., Roth, HR., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, RM. Deep convolutional neural networks for computer-aided detection: cnn architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging.* 35(5):1285-1298 (2016).

78. Schmidhuber, J. Deep learning in neural networks: an overview. *Neural Netw.* 61:85-117 (2015).

79. Sameen, MI., Pradhan, B., Lee, S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. *CATENA.* https://doi.org/10.1016/j.catena.2019.104249 (2020).