

Multi-scale convolutional neural networks (CNNs) for landslide inventory mapping from remote sensing imagery and landslide susceptibility mapping (LSM)

Baoyi Zhang, Jiacheng Tang, Yuke Huan, Lei Song, Syed Yasir Ali Shah & Lifang Wang

To cite this article: Baoyi Zhang, Jiacheng Tang, Yuke Huan, Lei Song, Syed Yasir Ali Shah & Lifang Wang (2024) Multi-scale convolutional neural networks (CNNs) for landslide inventory mapping from remote sensing imagery and landslide susceptibility mapping (LSM), *Geomatics, Natural Hazards and Risk*, 15:1, 2383309, DOI: [10.1080/19475705.2024.2383309](https://doi.org/10.1080/19475705.2024.2383309)

To link to this article: <https://doi.org/10.1080/19475705.2024.2383309>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 29 Jul 2024.



Submit your article to this journal



Article views: 291



View related articles



View Crossmark data

Multi-scale convolutional neural networks (CNNs) for landslide inventory mapping from remote sensing imagery and landslide susceptibility mapping (LSM)

Baoyi Zhang^a, Jiacheng Tang^a, Yuke Huan^b, Lei Song^c, Syed Yasir Ali Shah^a and Lifang Wang^d

^aKey Laboratory of Metallogenetic Prediction of Nonferrous Metals and Geological Environment Monitoring (Ministry of Education)/School of Geosciences and Info-Physics, Central South University, Changsha, China; ^bChina Railway Construction Heavy Industry Co. Ltd, Changsha, China; ^cHunan Institute of Geological Disaster Investigation and Monitoring, Changsha, China; ^dDepartment of Surveying and Mapping Geography, Hunan Vocational College of Engineering, Changsha, China

ABSTRACT

Accurate landslide susceptibility mapping (LSM) relies on a detailed landslide inventory and relevant influencing factors. In this study, Sentinel 2 remote sensing imagery is employed to establish a comprehensive landslide inventory leveraging attention U-Net backbone networks in Zhangjiajie City of Hunan Province, China. Subsequently, the refined landslide inventory, with more precise boundaries, is integrated into LSM process. A multi-scale sampling three-dimensional convolutional neural network (3D-CNN) is introduced into LSM, facilitating the extraction of multi-scale neighbourhood characteristics and deep information of relevant topographical, hydrological, meteorological, geological, and human activity factors. Experimental results demonstrate that this method achieves the highest accuracy and area under the receiver operating characteristic (ROC) curve. Moreover, its recall and F1-score significantly higher surpass those of other small-, medium-, and large-scale models, with the F1-score being more than 10% higher. This superior performance is attributed to its proficiency in discerning the nonlinear spatial correlation between landslide occurrences and influencing factors. The comprehensive consideration of scale characteristics through the multi-scale sampling strategy outperforms the single-scale CNN models across all evaluation metrics. This study furnishes a suite of high-precision methodologies for landslide hazard assessment in Zhangjiajie City, thereby offering invaluable support to decision-makers engaged in large-scale land use planning and geologic disaster prevention.

ARTICLE HISTORY

Received 12 October 2023
Accepted 18 July 2024

KEYWORDS

Landslide susceptibility mapping; deep learning; convolutional neural networks; multi-scale sampling; landslide inventory identification

1. Introduction

Landslides are complex and catastrophic natural disaster that results in huge human, economic, social, and environmental losses annually. These landslides are typically triggered by various influencing factors, including geomorphology, hydrology, meteorology, geology, and human activities (Guzzetti F et al. 2005). Anticipating future landslides based on past influencing factors is the premise behind landslide susceptibility mapping (LSM), which predicts the spatial occurrence probability of future landslides (Chung and Fabbri 1999; Guzzetti et al. 1999). Hence, utilizing these influencing factors for landslide susceptibility prediction is critical to minimizing the impact of landslide disasters, making LSM crucial for visualizing and analysing landslide susceptibility issues.

Landslide susceptibility refers to identifying areas prone to landslides based on the spatial distribution of conditioning factors associated with the instability process, without considering temporal influences (Dobrovolny 1971). This approach is valuable in regions with limited historical landslide event records, sparse rainfall data, or unknown magnitude/intensity of triggering earthquakes (Chacón et al. 2006). LSM assesses the contribution of various landslide conditioning factors (LCFs) to the spatial possibility of landslide events and the probability of slope failure under specific natural and artificial conditions. Common methods for landslide susceptibility prediction include physically based models (Montgomery and Dietrich 1994; Nery and Vieira 2015; Dikshit et al. 2019), knowledge-driven models (Mao et al. 2022), statistical models (Ramesh and Anbazhagan 2015; Zhang G et al. 2016; Chen et al. 2019), and machine learning (ML) models (Bui et al. 2020; Huan et al. 2023). Physically based models are challenging to apply to large-scale analysis, and knowledge-driven models are difficult to quantify or objectively evaluate. While statistical models and ML models can yield favourable results, the latter tends to be more accurate and efficient. Recent advancements have introduced various ML approaches into geosciences (Ghorbanzadeh et al. 2019; Zhang B et al. 2021; 2023). Integration landslide susceptibility prediction models that combines shallow ML techniques, e.g. support vector machines (SVMs) (Zhao S and Zhao 2021), artificial neural networks (ANNs) (Mariano et al. 2020; Zhao P et al. 2022), and decision trees (DTs) (Hong et al. 2018; Pham QB et al. 2021), and gradient boosting decision tree (GBDT) (Huan et al. 2023), have been introduced into LSM. However, shallow neural networks, learning from data described by manually pre-defined characteristics, may struggle to fully capture the linear or non-linear relationship between LCFs and landslide occurrences. Deep learning (DL) methods have gained popularity, often outperforming conventional shallow neural networks by extracting informative characteristics automatically from raw data with minimal pre-processing due to their complex architecture (LeCun et al. 2015). As a result, some studies have applied DL to LSM and achieved significant progress (Huang et al. 2020; Thi Ngo et al. 2021).

Convolutional neural networks (CNNs) are a crucial component of DL methods, as they can automatically extract valuable features from multidimensional data, taking into account the impact of the internal structure, spatial correlation, and neighborhood characteristics on prediction results (Lecun Y et al. 1998). CNNs not only capture auto-correlations within a massive imagery but also extract cross-correlations across multidimensional features, allowing them to leverage spatial auto-correlations and

characteristic cross-correlations for improved prediction models (Zuo and Xiong 2017; Li et al. 2020; Zhang B, Xu, et al. 2023). However, applying CNNs to LSM faces a critical challenge, i.e. the inconsistent data representation of 1D or 2D. While CNNs commonly use one or more two-dimensional images as input, LCFs in LSM are often represented as points or irregular polygons (Yi et al. 2020). To address the data representation issue and enhance the application of CNNs to LSM, Wang Y et al. (2019) constructed three CNN architectures with different data representation dimensions (1D, 2D, and 3D) in LSM. The 1D-CNN leverages local correction and progressively learns intricate representations from LCF vectors. The 2D-CNN converts each 1D LCF vector into a 2D matrix, effectively extracting hidden valuable features. In contrast, the 3D-CNN not only learns LCF representations but also captures local spatial information, partially overcoming the limitations of 1D- and 2D-CNN architectures (Yi et al. 2020; Azarafza et al. 2021). However, 3D-CNN-based LSMs still face challenges in extracting of spatial auto-correlation and deep features efficiently, optimizing large amounts of parameters, and incorporating the complete landslide inventory.

One of main constraints of the precision and rationality of CNN-based LSM is the focus of present studies on known landslides, neglecting the great threat posed by the potential landslides. Identifying landslide areas involves recognizing regions showing clear signs of deformation and having the potential to cause disasters in the near future. Common methods of landslide inventory identification include field survey and remote sensing recognition. While field surveys are highly accurate and controllable, they are inefficient and costly. Optical remote sensing imagery processing and recognition methods, including statistical methods, ML methods (Ghorbanzadeh et al. 2019), and DL methods (Ji et al. 2020; Wang H et al. 2021), provide alternative approaches for landslide identification, mapping, and monitoring. However, in practical research, the small probability event of landslides poses significant class imbalance problems, and parameter selection and optimization constrain model efficiency, affecting the effectiveness of landslide inventory identification based on semantic segmentation of remote sensing imagery.

This study aims to propose a detailed landslide inventory recognition method using remote sensing imagery and incorporate the identified results into a 3D-CNN-based framework for landslide susceptibility prediction and mapping in Zhangjiajie, Hunan Province, China. For the potential landslide inventory automatic identification, the attention U-Net backbone networks are utilized, and Bayesian optimization and focal loss are introduced to address the category imbalance problem in landslide inventory semantic segmentation. For the landslide susceptibility analysis, based on the easily measurable and widely available LCFs of topography, hydrology, meteorology, geology and human activities, along with the identified landslides from remote sensing imagery, a 3D-CNN method is proposed for LSM based on multi-scale sampling strategy.

2. Study area and dataset

2.1. Study area

Hunan Province has a total of 18,567 potential geological disaster sites, accounting for 6.5% of the national total, ranking fourth in China. Among these sites, landslides

are the main type, with a cumulative count of 11,405 disaster sites. Zhangjiajie City is one of the key areas for geological disaster prevention and control in Hunan Province. The geographical coordinates of Zhangjiajie City, located in the northwest of Hunan Province, range between $28^{\circ}52' N$ and $29^{\circ}48' N$ and $109^{\circ}40' E$ and $111^{\circ}20' E$, with an area of 9516 km^2 . Zhangjiajie City experiences a mid-tropical monsoon climate characterized by abundant and concentrated precipitation from May to July, totaling an average annual rainfall of 1200–1500 mm. Rainfall plays a crucial role in the occurrence and development of landslides, especially in shallow deposits within the study area. The primary land use types in the study area are forest land and crop-land. The permanent population of Zhangjiajie City has exceeded 1.51 million.

The altitudes in Zhangjiajie City range between 67 and 1840 m above sea level (asl). There are historical records of 192 medium landslides, twelve large landslides, and an extra-large landslide in Zhangjiajie City, as shown in Figure 1. The study area is situated between the Wuling Mountains and the northern Xiangxi Mountains in the northeast of the Yunnan-Guizhou Plateau, resulting in a distinct alluvial terrace and karst landscape. Several rivers traverse Zhangjiajie City, with the Lishui River and Loushui River being the main river systems.

Zhangjiajie City is situated in the tectonic background of the South China Plate. In the regional tectonic system, it falls within the third uplift zone of Neocathaysian tectonic system. As shown in Table 1 and Figure 2(a), the lithology of the study area primarily consists of sandstone, siltstone and mudstone. Historical landslides occurred predominantly in the limestone, dolomite, and shale of the Middle Cambrian Series

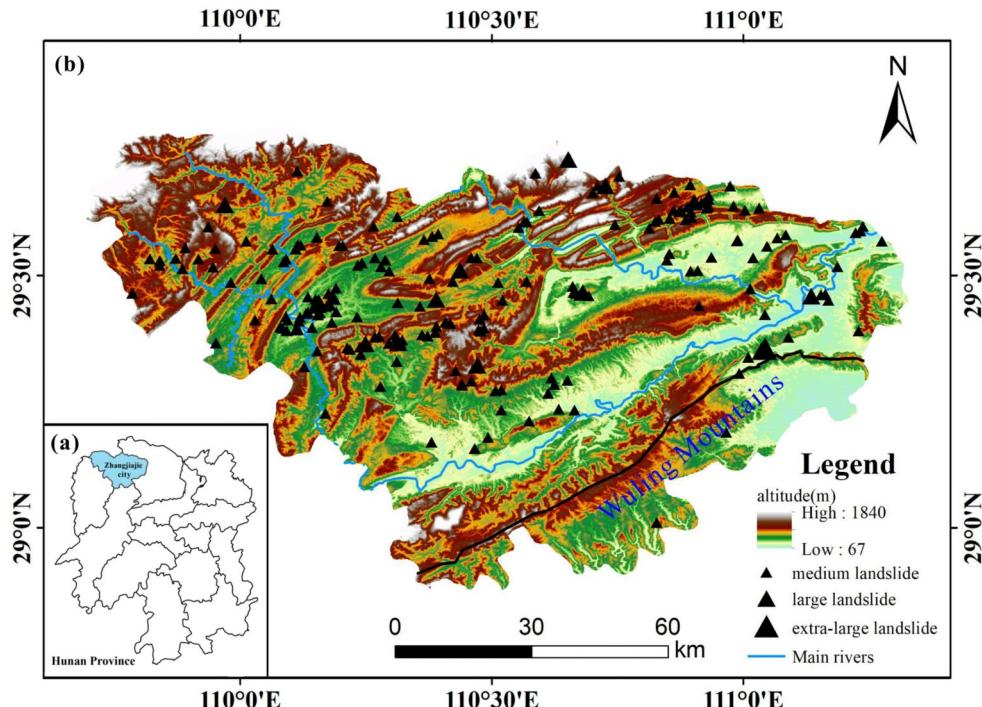
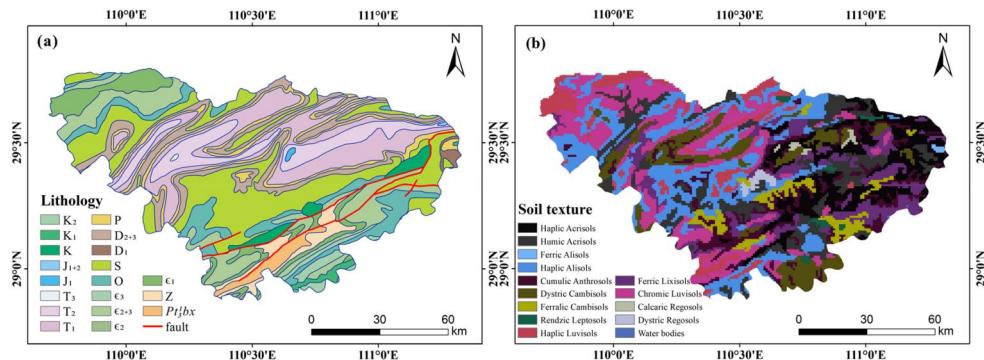


Figure 1. (a) Location of the study area in Hunan Province, China, and (b) location of historical landslides.

Table 1. Lithological units and their description in the study area.

| Lithological unit | Description | Geological Era |
|-----------------------|---|-------------------------|
| 1 K ₂ | Red sandstone, calcium mudstone, and glutenite. | Cretaceous |
| 2 K ₁ | Red siltstone, mudstone, glutenite, and basalt. | Cretaceous |
| 3 K | Merging layer. | Cretaceous |
| 4 J ₁₊₂ | Merging layer of the Upper and Middle Series. | Jurassic |
| 5 J ₁ | Sandstone, feldspar quartz sandstone, shale, conglomerate, and basalt. | Jurassic |
| 6 T ₃ | Sandstone, mudstone, shale, and quartz conglomerate. | Triassic |
| 7 T ₂ | Variegated calcareous mudstone, siltstone, and dolomite. | Triassic |
| 8 T ₁ | Limestone, dolomite, and argillaceous limestone. | Triassic |
| 9 P | Upper Series: siliceous rock, shale, limestone, and sandstone. Lower Series: chert limestone and magnesian marl. | Permian |
| 10 D ₂₊₃ | Merging layer of the Middle and Lower Series. | Devonian |
| 11 D ₁ | Purple sandstone, shale, and glutenite. | Devonian |
| 12 S | Siltstone, shale, sandy limestone, and shell limestone. | Silurian |
| 13 O | Merging layer. | Ordovician |
| 14 ε ₃ | Dolomite and limestone. | |
| 15 ε ₂₊₃ | Merging layer of the Upper and Middle Series. | Cambrian |
| 16 ε ₂ | Limestone, dolomite, and shale. | Cambrian |
| 17 ε ₁ | Limestone, marl, shale, and carbonaceous shale. | Cambrian |
| 18 Z | Syncline. | Sinian |
| 19 Pt ₃ bx | Banxi Group: purple slate, metamorphic sandstone, limestone, pebbled sandstone, and mafic rock. | Lower Upper Proterozoic |

**Figure 2.** (a) Geological map and (b) soil texture of the study area.

(ε₂), the dolomite and limestone of the Upper Cambrian Series (ε₃), the Ordovician strata (O), the purple sandstone and shale of the Lower Devonian Series (D₁), the shale, limestone, sandstone, chert limestone, and magnesian marl of the Permian (P), and the red sandstone and calcium mudstone of the Upper Cretaceous (K₂), exhibiting fragmentation and extensive weathering. There is a development of mud cementation, along with well-developed joints and fissures. The soil structure in the study area is loose, as shown in Figure 2(b), exhibiting high water dispersion and permeability, while cohesion and bonding are poor. Due to the slopes' unstable nature and the influence of surface water, ground water, and human activities, the shear strength of soil body diminishes, leading to an increased risk of sediment mass avalanches on

the slopes. This exacerbates the weathering process and reduces slope stability, ultimately resulting in slope collapse.

2.1. Dataset

2.2.1. Remote sensing imagery

The Sentinel 2 L1C remote sensing imagery from July 2019, with minimal cloud coverage, was employed to identify potential landslide inventory. Most of the landslide-prone areas in Zhangjiajie City are located in mountainous regions covered by evergreen needleleaf forest. The imagery remains suitable for landslide inventory mapping due to minimal changes in vegetation cover throughout the year. Landslides result in exposed soil that can be clearly distinguished from vegetation in remote sensing spectrum. Damaged or altered vegetation in landslide areas can be detected using near-infrared and red bands, while significant topographic changes caused by landslides are identifiable through different reflectance characteristics in multispectral data. In this study, a composition of four pre-processed L1C bands (bands 2, 3, 4, and 8) was utilized for identifying potential landslide inventory through DL semantic segmentation. Bands 2 (blue, 458–523 nm), 3 (green, 543–578 nm), 4 (red, 650–680 nm), and 8 (near-infrared, 785–900 nm) from Sentinel 2 imagery provide valuable information for distinguishing various land cover types and identifying landslide-specific spectral characteristics. Band 2 aids in recognizing features of water bodies and vegetation, Band 3 is used for vegetation differentiation and health monitoring, Band 4 effectively distinguishes between soil and vegetation, and Band 8 is highly sensitive to vegetation, making it useful for vegetation monitoring. Each band has a specific use and advantage, and bands 2, 3, 4, and 8 have unique roles in capturing and analyzing landslide features. Before segmentation, the imagery also underwent imagery cropping, multi-band imagery enhancement, and additional steps for DL, e.g. delineation of training area and labelling of samples. The training area selected for this study is Wulingyuan District (Figure 3a) of Zhangjiajie City, which comprises a total of 275 investigated landslide prone sites, compiled until 2021 through field surveys conducted by the Hunan Institute of Geological Disaster Investigation and Monitoring (<http://hnsdzy.hunan.gov.cn/dzs/>, 30 June, 2022) (Figure 3b). Among these landslide prone sites are 242 small (with slip mass less than 100,000 m³) landslides, 32 medium (ranging from 100,000 to 1,000,000 m³ slip mass) landslides, and a large (ranging from 1,000,000 to 10,000,000 m³ slip mass) landslide, classified based on the landslide mass volume criteria as outlined in DZ/T 0261-2014 (Ministry of Land and Resources of the People's Republic of China 2014). These landslide prone sites have resulted in 14 fatalities, causing a total of 37.3 million China Yuan in direct economic losses and 43.08 million China Yuan in potential economic losses due to the disasters. However, investigating all landslide prone sites in whole Zhangjiajie City would entail significant workloads and costs. Therefore, we extracted all landslide prone sites of Zhangjiajie City based on the Sentinel 2 remote sensing imagery of the city and the investigated sites in Wuling District. Based on these investigated landslide prone sites, their areas were manually interpreted and delineated according to the land cover texture and topographic index on the Sentinel 2 remote sensing imagery, as shown in

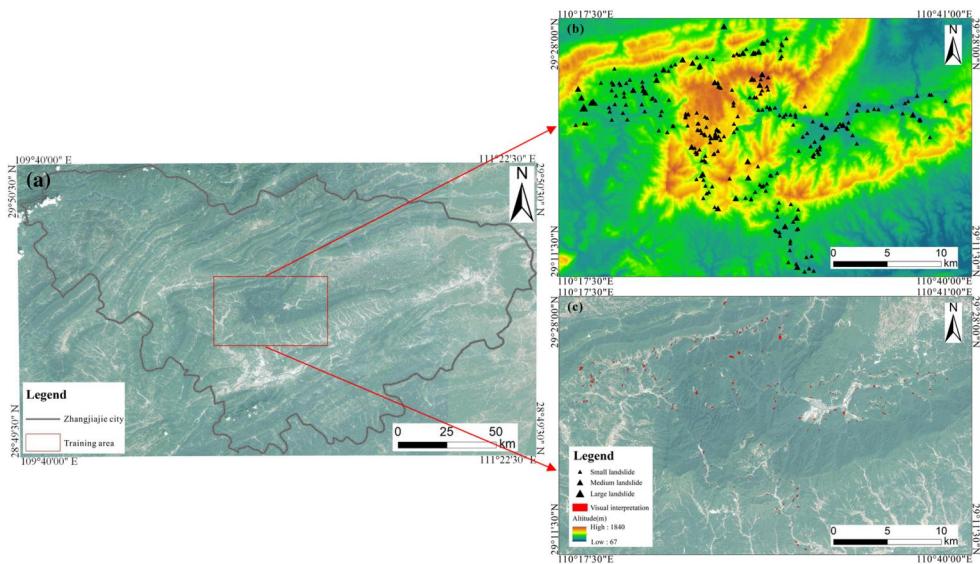


Figure 3. Distributions of training area (Wulingyuan District) and investigated landslide sites: (a) extent of training area on the Sentinel 2 imagery; (b) investigated landslide sites on the DEM; and (c) interpreted boundaries of investigated landslides on the Sentinel 2 imagery.

Figure 3(c). A total of 2428 sliced images labelled as landslide or non-landslide, with a size of 128×128 pixels, and each pixel with a spatial resolution of 10 m, were used to recognize potential landslide inventory. The aforementioned four bands are each used as one channel of the input data for the attention U-Net. This combination of the four channels allows for the integrated input of the band information. Then, 80% and 20% of them were randomly chosen as training dataset and testing dataset, respectively.

2.2.2. Landslide conditioning factors

The occurrence of landslide disaster is influenced by various inducing conditions. Therefore, selecting appropriate landslide conditioning factors (LCFs) as model input is of great significance for landslide susceptibility prediction. Based on easily measurable and widely available data, fifteen LCFs of topography, hydrology, meteorology, geology, and human activities, along with the identified landslides, were utilized for the landslide susceptibility prediction and mapping. As shown in Table 2, these LCFs include altitude, slope, aspect, plane curvature, profile curvature, relief, roughness, rainfall, topographic wetness index (TWI), normalized difference vegetative index (NDVI), distance to roads, distance to rivers, land use/land cover (LULC), soil texture, and lithology. Traditionally, rainfall is often regarded as a triggering factor for landslides rather than a static environmental LCF (Lin et al. 2021). However, contrasting to monthly or seasonal rainfalls, the annual rainfall characteristics of various regions in Zhangjiajie City exhibit substantial variation. Therefore, annual rainfall is utilized as a static environmental LCF to depict the spatial variability in rainfall characteristics across the study area. The assumption is that higher annual rainfall corresponds to increased landslide susceptibility when other LCFs remain similar. All the

Table 2. Landslide conditioning factors used in this study.

| No. | LCF | Categories | Data source | Resolution (Scale) | Description |
|-----|------------------------|--|---|--------------------|---|
| 1 | Altitude (m) | 67–317, 317–527, 527–750, 750–1037, and 1037–1840. | Digital elevation model (DEM) | 90 × 90 m | Affecting rainfall and related surface runoff (Botzen et al. 2013) |
| 2 | Slope (°) | 0–8.382, 8.382–15.680, 15.680–23.664, 23.664–33.440, and 33.440–71.275. | DEM | 90 × 90 m | Influencing slope stability as well as surface runoff, infiltration, and drainage density (Nefeslioglu et al. 2008) |
| 3 | Aspect | Flat (−1), north (0–22.5°), northeast (22.5–67.5°), east (67.5–112.5°), southeast (112.5–157.5°), south (157.5–202.5°), southwest (202.5–247.5°), west (247.5–292.5°), northwest (292.5–337.5°), and north (337.5–360°). | DEM | 90 × 90 m | The orientation of the slope. |
| 4 | Plane curvature | −6.555 to −0.460, −0.460 to 0.120, −0.120 to 0.137, 0.137 to 0.478, and 0.478 to 4.314. | DEM | 90 × 90 m | Reflecting the divergence and convergence of water on the surface (Huang et al. 2020) |
| 5 | Profile curvature | −5.902 to −0.569, −0.569 to 0.178, −0.178 to 0.116, 0.116 to 0.507, and 0.507 to 6.575. | DEM | 90 × 90 m | Affecting the flow velocity variation of slope (Pham et al. 2017) |
| 6 | Relief (m) | 0–45, 45–80, 80–120, 120–176, and 176–756. | DEM | 90 × 90 m | Difference of the maximum and minimum elevations within a certain area. |
| 7 | Roughness | 1–1.054, 1.054–1.137, 1.137–1.261, 1.261–1.509, and 1.509–3.321. | DEM | 90 × 90 m | Undulation changes and erosion degree of the ground surface. |
| 8 | Rainfall (mm/yr) | 0–1000, 1000–1200, 1200–1400, 1400–1600, and 1600–2000. | Hunan Provincial Institute of Land Resources Planning (www.hngtghy.com , Changsha, China) | 1000 × 1000 m | An important external condition. |
| 9 | TWI | −1.025–3.277, 3.277–6.456, 6456–9.542, 9.542–12.067, and 12.067–22.822. | DEM | 90 × 90 m | Predicting areas susceptible to the surface of saturated soil. |
| 10 | NDVI | 0.244–0.500, 0.500–0.632, 0.632–0.736, 0.736–0.816, and 0.816–0.900. | Resource and Environment Science Data Center (www.resdc.cn , Beijing, China) | 30 × 30 m | Measuring the degree of vegetation coverage and the status of vegetation growth (Tien Bui et al. 2019) |
| 11 | Distance to rivers (m) | 0–603.738, 603.738–1297.998, 1297.998–2012.461, 2012.461–2747.58, 2747.581– | OpenStreetMap (www.openstreetmap.org) | | openhistoricalmap.org , Cambridge, UK |

(continued)

Table 2. Continued.

| No. | LCF | Categories | Data source | Resolution (Scale) | Description |
|-----|-----------------------|---|--|--------------------|--|
| 12 | Distance to roads (m) | 90 × 90 m Erosion of river. 3520.369, 3520.369– 4374.037, 4374.037– 5331.313, 5331.313– 6468.114, and 6468.114–9198.070. | OpenStreetMap | 90 × 90 m | Reflecting the intensity of human influence. |
| 13 | LULC | Cropland, forests, grassland, wetland, water bodies, and artificial surfaces. | GLOBELAND30 (www.globallandcover.com , Beijing, China) | 30 × 30 m | The soil properties of cohesive force, friction angle, soil bulk density, and pore pressure will change under different land covers (Nguyen et al. 2019) |
| 14 | Soil texture | Haplic acrisols, humic acrisols, ferric alisols, haplic alisols, cumulic anthroposols, dystric cambisols, ferralic cambisols, rendzic leptosols, haplic luvisols, ferric lixisols, chromic luvisols, calcareous regosols, dystric regosols, and water bodies. | Harmonized World Soil Database (HWSD) built by the Food and Agriculture Organization of the United Nations (UNFAO, www.fao.org , Rome, Italy) and the International Institute for Applied Systems Analysis (IIASA, iiasa.ac.at , Laxenburg, Austria) | 1:5,000,000 | Representing the texture of the soil materials that influence landslide occurrence. |
| 15 | Lithology | As shown in Table 1. | Geological map database of China | 1:500,000 | Influencing the occurrence of erosion, ground stability, and landslide (Arulbalaji et al. 2019) |

LCFs were converted into raster format with a spatial resolution of 90 m, as shown in **Figure 4**. However, due to the terrain slope and roughness factors were derived from the relief factor, the relief factor was removed, remaining fourteen LCFs as input variables for LSM.

Cited references in **Table 2** (Nefeslioglu et al. 2008; Botzen et al. 2013; Pham BT et al. 2017; Arulbalaji et al. 2019; Nguyen et al. 2019; Tien Bui et al. 2019; Huang et al. 2020).

3. Methods

Landslide identification is conducted using Sentinel 2 remote sensing imagery, with the results is subsequently utilized for susceptibility analysis. The identification

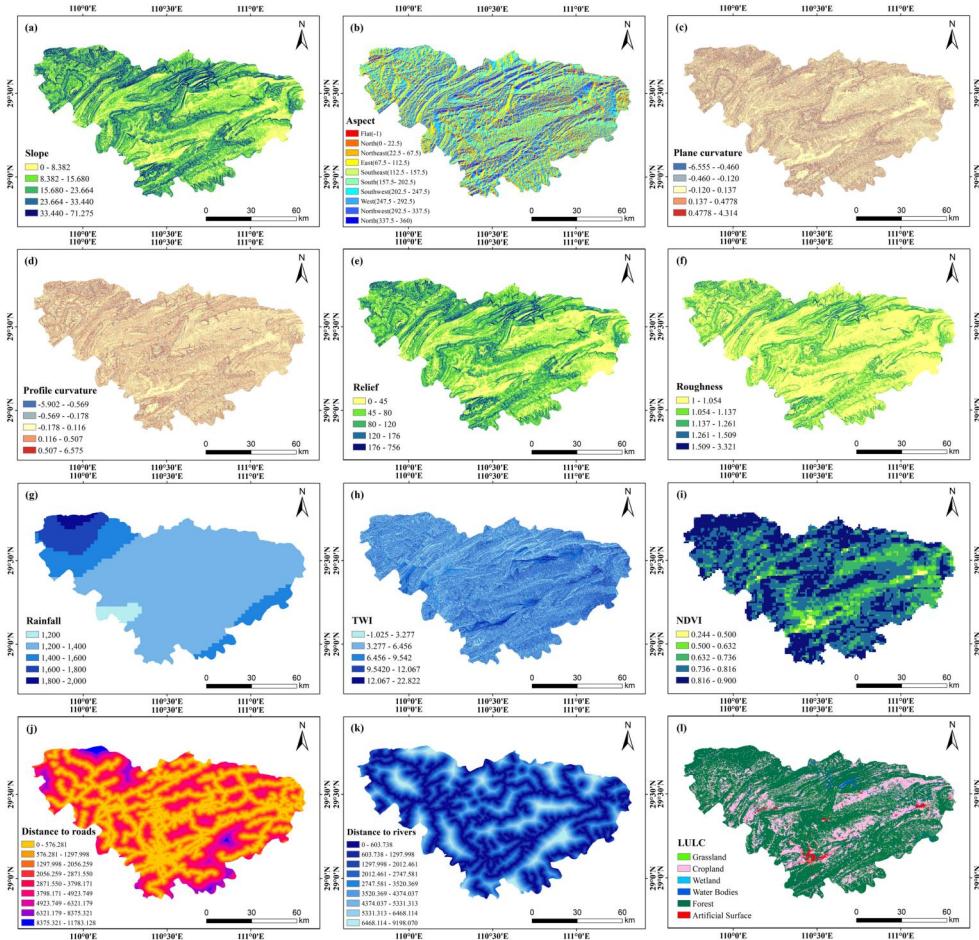


Figure 4. Landslide conditioning factors for LSM: (a) slope, (b) aspect, (c) plane curvature, (d) profile curvature, (e) relief, (f) roughness, (g) rainfall, (h) TWI, (i) NDVI, (j) distance to roads, (k) distance to Rivers, and (l) LULC.

employs an attention U-Net model, while the susceptibility analysis leverages a multi-scale three-dimensional convolutional neural network (3D-CNN) model.

3.1. Landslide inventory mapping based on attention U-Net

3.1.1. Backbone networks of attention U-Net

U-Net is a fully convolutional network (FCN) model suitable for processing small-sized targets, requiring less training data and providing higher segmentation accuracy compared to other CNNs (Ronneberger et al. 2015). In the encoding part of the U-Net, the input image undergoes progressively filtering and down-sampling to produce multi-scale features. To further improve the segmentation performance, attention U-Net is introduced into this study. The attention mechanism can enhance the U-Net model's focus on important regions and improve the accuracy of segmentation results (Oktay et al. 2018). Information extracted from coarse scale is employed in

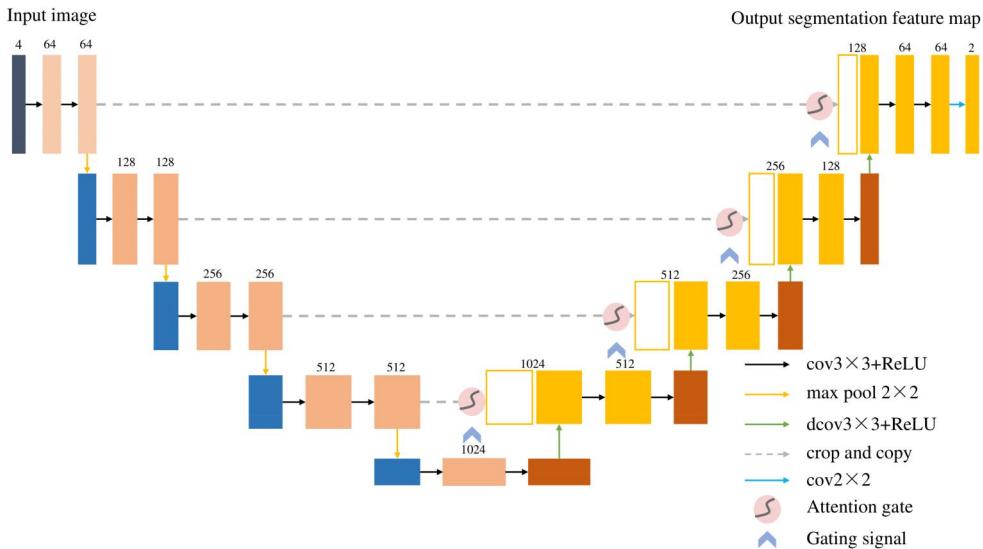


Figure 5. Backbone networks of the attention U-Net.

attention gating to disambiguate irrelevant and noisy responses in skip connections. Therefore, attention U-Net is chosen as the backbone networks in this study.

The detailed networks of attention U-Net are shown in Figure 5, consisting of an encoder and a decoder in an axisymmetric shape. The encoder acquires the image features by down-sampling them through a convolutional layer. The decoder performs successive up-sampling of the feature image to gradually recover image detail features, with a jump connection set up between the decoder and its corresponding encoder to preserve detailed location information effectively. Specifically, the encoder part down-samples four times and mainly consists of 3×3 convolutional layers, 2×2 maximum pooling layers, and the rectified linear unit (ReLU) function for extracting image feature information. The decoder part performs four corresponding up-sampling operations to recover image information, with the main structure comprising the attention module and successive 2×2 deconvolution layers. By concatenating the corresponding feature patterns of the encoder and decoder, the shallow networks' detailed location information is effectively preserved to assist in semantic segmentation of imagery.

3.1.2. Focal loss function

Landslide hazards, as small probability events, account for a relatively small proportion of pixels in remote sensing imagery, while the background has a large proportion of pixels, leading to a serious category imbalance issue. Designing a suitable loss function does not lose image information while optimizing the network structure, making it the preferred method to address the category imbalance problem in imagery semantic segmentation.

Focal loss is a loss function specifically designed to handle the category imbalance problem. Its purpose is to assign weights α_t to the losses $FL()$ of samples based on their ease probability of differentiation p_t , with smaller weights being added to

categories that are easier to differentiate and larger weights being added to categories that are more difficult to discriminate. The easily differentiated categories are those with classification confidence levels approaching 0 or 1, while the remaining categories are considered hard-to-discriminate.

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (1)$$

where $(1 - p_t)^\gamma$ is a category weight modifier, and γ is an adjustable focus parameter. The larger its value, the more attention the networks are trained to direct towards hard-to-distinguish categories.

3.2. Three-dimensional CNNs for LSM

3.2.1. Three-dimensional data representation

Generally, 1D-CNNs treat the input data a vector (Figure 6a), where each row represents a sample of landslide sites, and each column represents an LCF attribute of that sample. Thus, each grid cell in the input data can be represented by a vector, the length of which is defined by the number of LCFs. The number of LCFs is compared with the number of attribute values for each LCF, and the larger of the two numbers is chosen as the size of the corresponding 2D matrix, which serves as the input of 2D-CNNs (Figure 6b). However, 1D-CNNs and 2D-CNNs only make use of local landslide information and ignore environmental information, and the 2D data contains many zero values, resulting in sparse data. To address these limitations, this study utilizes a 3D-CNN strategy, treating the entire input data as a multi-channel image. The LCFs are considered individual channels of the study area image and are fed into the networks simultaneously (Figure 6c). The input data of the study area is represented as a 3D matrix of size $c \times n \times n$, where n denotes the rows and columns of the imagery, and c denotes the number of LCFs.

The simplest 3D-CNNs consist of a kernel convolutional layer of size $m \times m \times m$, a maximum pooling layer, and a fully connected layer. When the size of the input data is $c \times n \times n$, the output feature map size of the convolution layer is $(c-m+1) \times (n-m+1) \times (n-m+1)$. Each grid cell is connected to an $m \times m \times m$ neighborhood

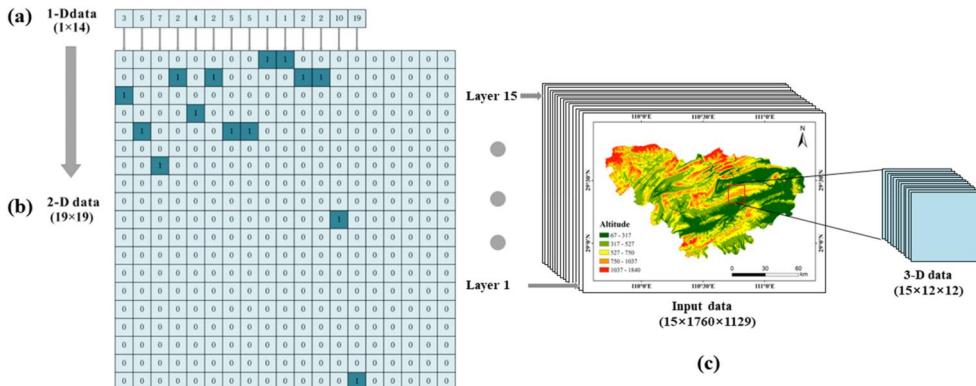


Figure 6. Data presentations for (a) 1D-CNNs, (b) 2D-CNNs, and (c) 3D-CNNs.

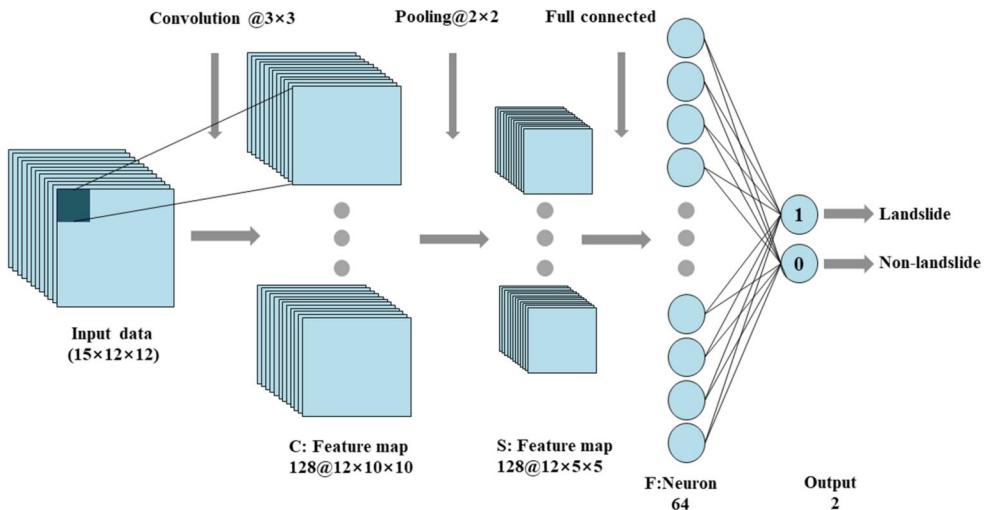


Figure 7. Schematic diagram of 3D-CNNs architecture.

in the input data. Next, a maximum pooling layer of size $a \times a$ outputs N feature maps of size $(\frac{c-m+1}{a}) \times (\frac{n-m+1}{a}) \times (\frac{m-1}{a})$. The fully connected layer has k neurons, and the output layer has 2 neurons, representing landslide and non-landslide, respectively. **Figure 7** illustrates the 3D-CNNs architecture for this study (where $c=14$, $n=22$, $N=128$, $m=3$, $a=2$, and $k=64$), depicting the combined structure of a convolutional layer, a pooling layer, and a fully-connected layer.

3.2.2. Multi-scale sampling strategy

According to the spatial autocorrelation, the occurrence of a landslide is closely related to its surrounding environment. However, LCFs originate from varied sources, each with differing resolutions and scales. It necessitates both up-sampling and down-sampling procedures to homogenize LCFs into a unified spatial resolution. Therefore, a multi-scale sampling strategy is introduced in this study to preserve the original-scale characteristics of LCFs from diverse sources, minimizing uncertainties that may arise during these necessary procedures, as shown in **Figure 8**. Firstly, all LCF images are stacked as a factor dataset. Secondly, a marching window is introduced, centred on landslide or non-landslide within a certain range, and it expands outward to extract spatial information within larger range, generating sub-datasets of different scales. Thirdly, all sub-datasets are resampled to a unified small scale, which contains multi-scale spatial information. Finally, sub-datasets with different scales are combined to create multi-scale datasets, serving as input data for the 3D-CNN model. In this study, small, medium, and large size sample scales are used, and they were set as 12, 17, and 22 after extensive experimental verification. Finally, all LCF images were resampled to the smallest scale size of 12 during muti-scale data fusion.

3.2.3. Model architecture

To avoid model overfitting and extract more detailed features, a moderate network depth and small convolutional energy is adopted base on the visual geometry group

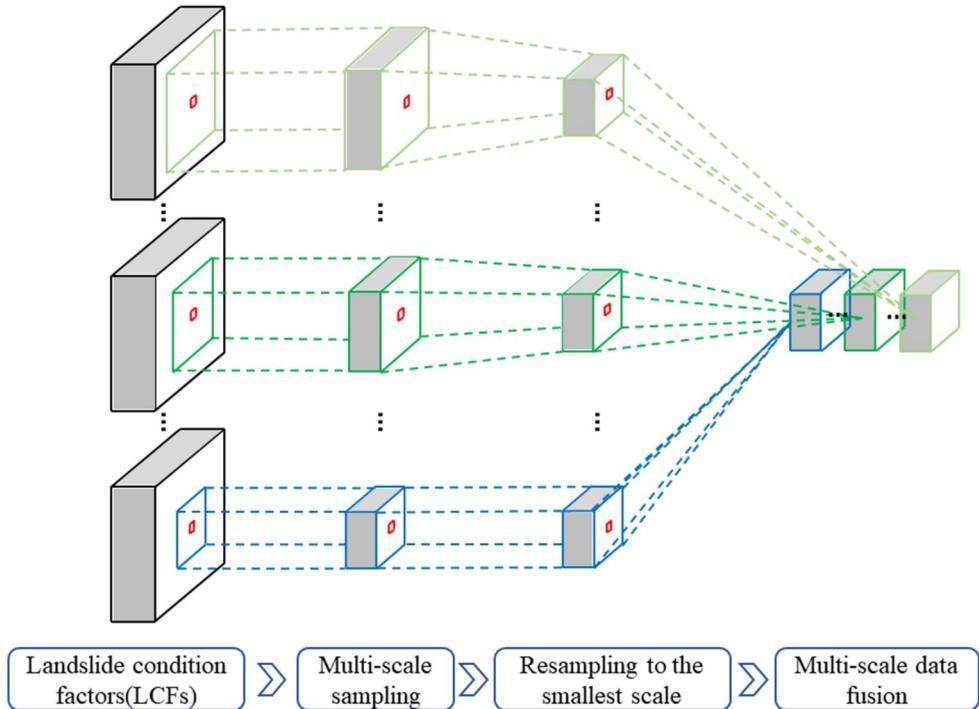


Figure 8. Construction of a multi-scale LCF dataset. Red square indicated centred pixel labelled as landslide or non-landslide.

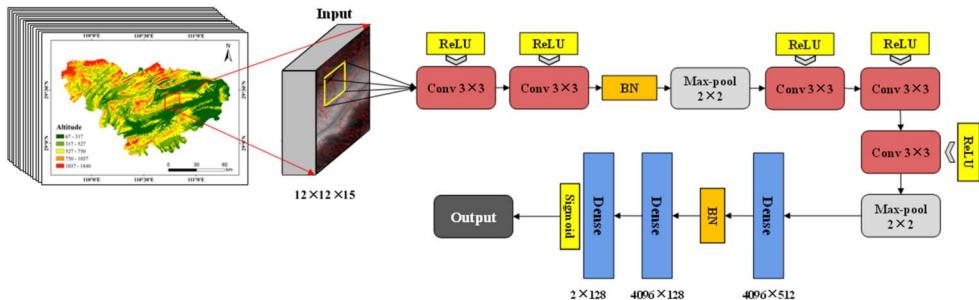


Figure 9. Architecture of the 3D-CNN model for LSM base on VGG.

network (VGG). VGG has a moderate network structure, small convolution kernels, and pooled sampling domains, which can capture more detailed features of 2D remote-sensing imagery (Liu and Deng 2015). Based on the original VGG networks, as shown in Figure 9, the following improvements are made according the experimental needs: (1) To improve computational efficiency, a smaller convolution window with a size of $12 \times 12 \times 15$ and a stride of 1 is used for data processing. (2) The network structure is optimized by employing two 3×3 sized convolutional kernels superimposed with three 3×3 sized convolutional kernels. This approach provides a similar perceptual field to using one large convolutional kernel, while reading deeper information with a smaller memory footprint. (3) A sample weight function is



introduced into the networks, wherein the algorithm multiplies the training loss of each sample by its weight. This is a good solution to the problem of low training accuracy caused by sample imbalance. (4) The data is normalized to a decimal in the (0,1) interval, and the dimension expression is changed to a dimensionless expression, thus speeding up the gradient descent process to find the optimal solution. (5) To overcome vanishing gradient, the batch normalization method is employed to bring the eigenvalue distribution of the layer back to the standard normal distribution, avoiding the vanishing gradient and accelerating convergence.

4. Results and discussions

4.1. Landslide inventory identification and mapping

4.1.1. Hyperparameter optimization

Hyperparameter optimization is a crucial but challenging task in the DL training process. Bayesian optimization is an iterative optimization method based on probabilistic models, employed to find the optimal solution within a limited number of evaluations of the objective function (Victoria and Maragatham 2021). At the core of Bayesian optimization lie Gaussian process regression and the acquisition function (AC). Gaussian process predicts a new distribution of an input variable x given known observations of input variable and objective function of model accuracy (x^* , y^*), while the AC guides the selection of the next candidate point x based on Gaussian process sampling. In this study, expected improvement (EI) serves as the AC and is defined by the formula:

$$EI(x) = \begin{cases} (\mu(x) - f(x^+))\Phi(z) + \sigma(x)\Phi'(z), & \sigma(x) > 0 \\ 0, & \sigma(x) = 0 \end{cases} \quad (2)$$

$$z = \frac{\mu(x) - f(x^+)}{\sigma(x)} \quad (3)$$

where $\Phi(\cdot)$ represents the normal cumulative distribution function, $\phi(\cdot)$ denotes the normal probability density function, and $f(x^+)$ represents the current maximum value of the objective function. For each x , $\mu(x)$ and $\sigma(x)$ are the predicted mean and standard deviation of the Gaussian process, respectively, and EI is computed to identify the x corresponding to the maximum EI value.

In this study, hyperparametric optimization was conducted for the learner, the optimal learning rate and the weight coefficients α and γ of the loss function. The search targets of the learner were adaptive moment estimation (Adam) (Kingma and Ba 2015) and stochastic gradient descent (SGD) (Ruder 2016), the learning rate was searched in the values of 1e-2, 1e-3 and 1e-4, the α value of the loss function was searched in the range of 0 to 1, the γ value was searched in the range of 0 to 5, and the number of hyperparametric optimization iterations was set to 30.

The specific hyperparameter optimization results from 30 iterations are shown in Table 3. According to the Bayesian hyperparameter optimization results, the model accuracy reached over 95% in 11 out of 30 iterations of the experiment, with an

Table 3. Bayesian hyperparameter optimization results for the U-Net model.

| Iteration | Accuracy | Learner | Learning | α | γ |
|-----------|----------|---------|----------|----------|----------|
| 1 | 0.2999 | Adam | 0.001 | 0.573 | 4.549 |
| 2 | 0.3261 | Adam | 0.01 | 0.877 | 1.885 |
| 3 | 0.2356 | Adam | 0.01 | 0.892 | 3.941 |
| 4 | 0.5572 | SGD | 0.001 | 0.969 | 4.944 |
| 5 | 0.9884 | Adam | 0.001 | 0.639 | 3.856 |
| 6 | 0.5657 | SGD | 0.001 | 0.282 | 4.436 |
| 7 | 0.9879 | SGD | 0.001 | 0.298 | 4.519 |
| 8 | 0.7253 | Adam | 0.01 | 0.818 | 1.655 |
| 9 | 0.4531 | Adam | 0.001 | 0.999 | 3.363 |
| 10 | 0.6094 | SGD | 0.01 | 0.17 | 3.366 |
| 11 | 0.9882 | SGD | 0.01 | 0.165 | 0.727 |
| 12 | 0.7606 | SGD | 0.001 | 0.046 | 4.731 |
| 13 | 0.9885 | Adam | 0.01 | 0.784 | 0.371 |
| 14 | 0.5195 | Adam | 0.001 | 0.866 | 4.877 |
| 15 | 0.8878 | Adam | 0.001 | 0.105 | 3.226 |
| 16 | 0.9888 | Adam | 0.01 | 0.338 | 3.809 |
| 17 | 0.6622 | SGD | 0.01 | 0.949 | 2.831 |
| 18 | 0.9882 | SGD | 0.001 | 0.121 | 0.316 |
| 19 | 0.8302 | Adam | 0.01 | 0.514 | 1.502 |
| 20 | 0.9877 | Adam | 0.001 | 0.587 | 0.891 |
| 21 | 0.9877 | Adam | 0.01 | 0.591 | 3.737 |
| 22 | 0.5728 | SGD | 0.01 | 0.736 | 0.92 |
| 23 | 0.6053 | SGD | 0.01 | 0.283 | 2.219 |
| 24 | 0.9541 | Adam | 0.01 | 0.748 | 2.511 |
| 25 | 0.9871 | SGD | 0.01 | 0.55 | 0.307 |
| 26 | 0.5283 | SGD | 0.01 | 0.763 | 2.976 |
| 27 | 0.9886 | SGD | 0.001 | 0.551 | 0.665 |
| 28 | 0.4843 | Adam | 0.01 | 0.293 | 1.563 |
| 29 | 0.4416 | Adam | 0.001 | 0.151 | 2.28 |
| 30 | 0.3985 | Adam | 0.01 | 0.977 | 0.164 |

average accuracy of 71%. This indicates that the Bayesian hyperparameter optimization introduced in this study can efficiently search for a better combination of hyperparameters. The accuracy on the testing dataset reached 98.88% when the number of optimization iterations reached 16. The hyperparameters obtained in this iteration are considered as the best values obtained in these 30 iterations of hyperparameter search. Specifically, the learner is Adam, the learning rate is 0.01, and the values of α and γ coefficients in the focal loss function are 0.338 and 3.809, respectively. The subsequent U-Net model can be constructed using this set of hyperparameters.

4.1.2. Model training

After data pre-processing, the entire dataset (2428 sliced remote sensing images with their corresponding label images) was randomly divided into a training dataset (80%) and a testing dataset (20%). A total of 1942 sliced remote sensing images of size 128×128 pixels, with a spatial resolution of $100 \text{ m}^2/\text{pixel}$, along with their corresponding label images, were used for model training. Additionally, 486 sliced remote sensing images and their corresponding labelled images were used for model validation. In this study, the early stopping function was introduced during the training process, designed to halt the training process when the loss rate no longer decreases after n rounds of epochs, with $n=10$ set in this study. Four DL semantic segmentation models based on the attention U-Net, with Adam optimizer and batch size of 4, were designed for comparison, including: (1) Model-1, training 50 epochs using the

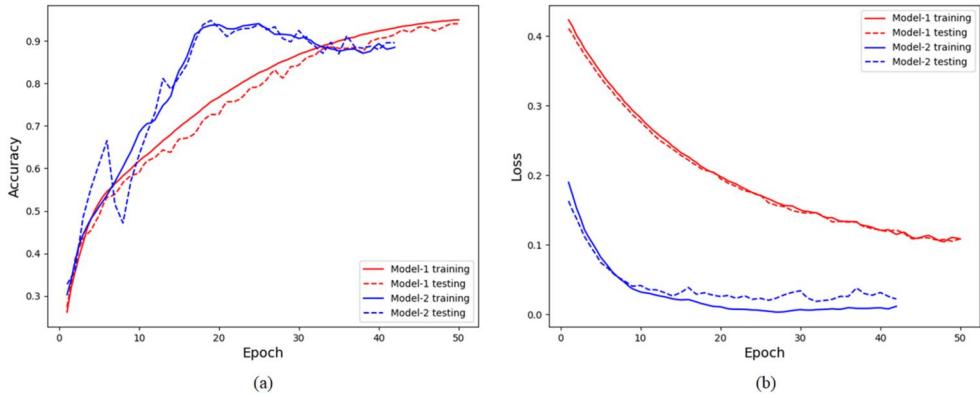


Figure 10. (a) Accuracy and (b) loss rate curves of model-1 and model-2.

focal loss function, (2) Model-2, training 50 epochs using the bivariate cross-entropy loss function, (3) Model-3, training 100 epochs using the focal loss function, and (4) Model-4, training 100 epochs using the bivariate cross-entropy loss function. All four models achieved good convergence and stopped training after 50 epochs and 75 epochs, respectively, and Figure 10 shows the dynamic accuracy and loss function plots for Model-1 and Model-2.

4.1.3. Model evaluation

In this study, the performances of four attention U-Net models, trained with different combinations of loss function and number of epochs, were evaluated using the evaluation metrics of accuracy, precision, recall, F1-score, and mIoU. Higher values of these evaluation metrics, ranging from 0 to 1, indicate better overall performance and prediction ability of the model. Accuracy is computed as the ratio of correctly predicted instances to the total number of instances:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

where TP , TN , FP , and FN represent the numbers of true positive, true negative, false positive, and false negative sampling images in the training or testing datasets, respectively.

Precision signifies the ratio of true positive (TP) to all predicted positive ($TP + FP$), while recall represents the ratio of true positive (TP) to all determined positive ($TP + FN$). The F α -score is the harmonic mean of precision and recall with a weight coefficient α , and for F1-score, $\alpha = 1$.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

Table 4. Results of performance evaluation of four models.

| Model | Parameters | Accuracy | Precision | Recall | F1-score | mIoU |
|---------|---------------------------------------|----------|-----------|--------|----------|-------|
| Model-1 | Focal loss, 50 epochs | 0.935 | 0.995 | 0.754 | 0.820 | 0.735 |
| Model-2 | Binary cross entropy loss, 50 epochs | 0.498 | 0.981 | 0.499 | 0.498 | 0.492 |
| Model-3 | Focal loss, 100 epochs | 0.875 | 0.993 | 0.696 | 0.758 | 0.671 |
| Model-4 | Binary cross entropy loss, 100 epochs | 0.501 | 0.961 | 0.501 | 0.500 | 0.485 |

$$F\alpha - score = \frac{(\alpha^2 + 1)Precision * Recall}{\alpha^2(Precision + Recall)} \quad (7)$$

The mIoU measures the overlap between the predicted result and an actual dataset, calculated as the average of the intersection over union for each class.

$$mIoU = \frac{1}{2} \left(\frac{TP}{FN + FP + TP} + \frac{TN}{FP + FN + TN} \right) \quad (8)$$

Table 4 shows the differences in model performance after adding different module. Focal loss significantly addressed the data category imbalance problem, improving the model's performance while maintaining a stable recall rate. However, when the number of training epochs was set to 100, the models converged at around 75 epochs and automatically stops training, resulting in a less pronounced improvement in the models' performance in all aspects.

Furthermore, the capabilities of Model-1 and Model-3 were evaluated in the actual landslide remote sensing recognition. Firstly, the remote sensing imagery of a rectangular area within the cross-validation area was interpreted through human-computer interactions ([Figure 11a](#)). Human-computer interpretation outlines landslide boundaries based on differences in image texture, mainly expressed in terms of both vegetation cover and topographic characteristic, between the landslide area and surrounding background areas. This method ensures a large accuracy rate, but is not suitable for landslide identification in large areas as it relies too much on the empirical professional knowledge and has high time costs. The human-computer interpreted results were compared with the landslide identification results of Model-1 and Model-3, and the comparison reveals that Model-1 ([Figure 11b](#)) is able to extract more landslide areas, while Model-3 ([Figure 11c](#)) extracts more fragmented areas and misses many small landslide areas. Secondly, [Figure 12](#) displays six actual landslides recognized by Model-1 and Model-3. The results demonstrate that Model-1, with the focal loss and 50 epochs, has enhanced the performance of landslide recognition, indicating its feasibility.

4.1.4. Landslide inventory Mapping

Model-1 was employed to recognize landslides throughout the study area. After using the optimal Model-1 proposed in this study for initial landslide recognition, some areas, e.g. roads, residential areas, water bodies, and other features that were incorrectly identified as landslides were effectively removed by overlaying them with corresponding map layers. In total, 61,238 landslide inventory areas of various sizes were

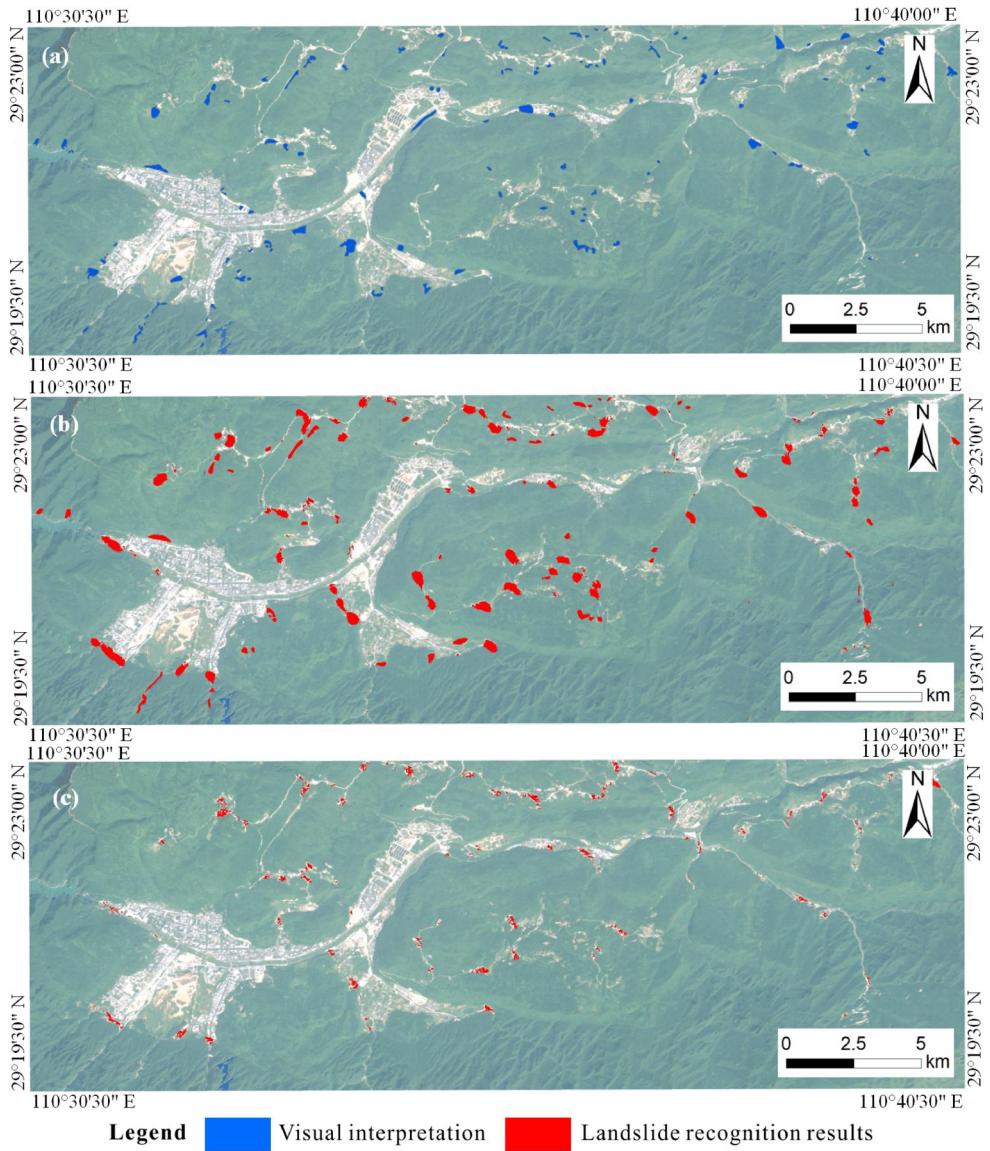


Figure 11. Potential landslide inventory on the sentinel 2 imagery: (a) human-computer interpretation result, (b) landslide recognition results by model-1, and (c) landslide recognition results by model-3.

retained, as shown in Figure 13. The recognized landslide areas are primarily situated within the strata of the Middle Cambrian Series (ϵ_2), the Upper Cambrian Series (ϵ_3), the Ordovician (O), the Lower Devonian Series (D_1), the Permian (P), and the Upper Cretaceous (K_2), aligning with historical landslide occurrences. These landslides are predominantly observed as thin and dark yellow strips in remote sensing imagery, located in mountainous areas, beside roads, next to buildings, or on the inlet side of rivers and lakes.

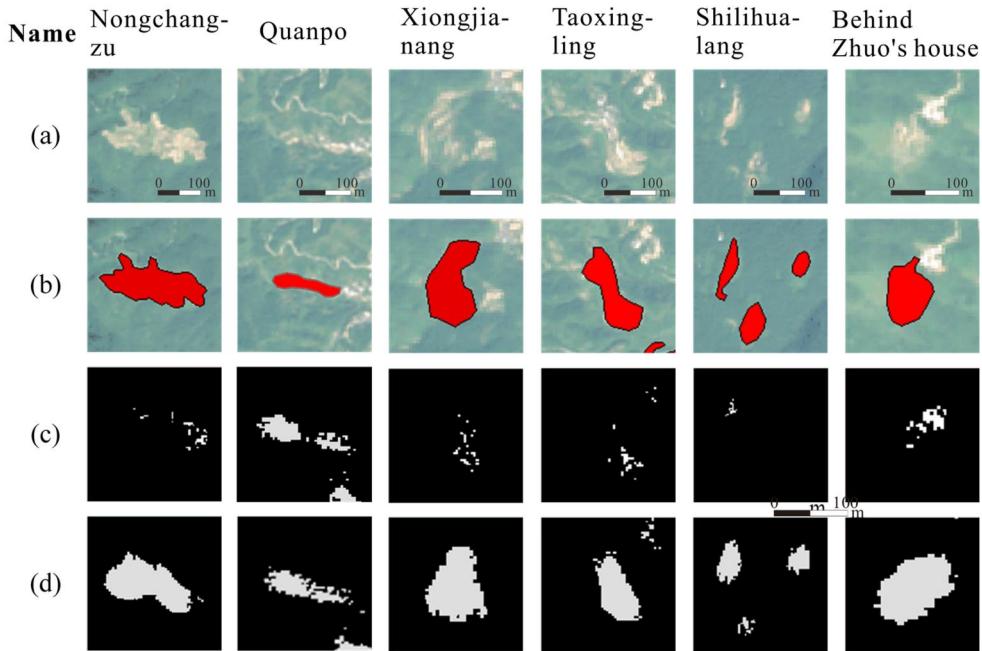


Figure 12. Six landslides comparison: (a) original sentinel 2 image, (b) human-computer interpretation on the sentinel 2 imagery, (c) recognized by model-1, and (d) recognized by model-3.

4.2. Landslide susceptibility evaluation and mapping

4.2.1. Model training

All data from the fourteen LCFs that were pre-processed, along with identified landslide inventory from remote sensing imagery, were fed into the above models, implemented in Python 3.7 using the Keras package, an open source DL framework based on TensorFlow (<http://www.tensorflow.org>, 30 June 2023), and the Scikit-learn package (<https://scikit-learn.org/stable>, 20 June 2023). Eight CPUs of Intel CoffeLake i7-9700F with the main frequency of 4.7 GHz and a NVIDIA Quadro P2000 GPU with a 5 Gb memory were used for the experiments for performance benchmarking, and the hyperparameters used in these models are listed in [Table 5](#). Four distinct 3D-CNN models were compared, utilizing different marching window sizes: the first three models employed marching window sizes of small-scale (12×12), medium-scale (17×17), and large-scale (22×22), respectively, while the fourth model integrated a multi-scale sampling process. Notably, all models demonstrated convergence within 25 epochs, as shown in [Figure 14](#).

4.2.2. Model evaluation

Huan et al. (2023) utilized a gradient boosting decision tree and logistic regression stacking ensemble model (GBDT + LR) to map the land susceptibility in Zhangjiajie City using a similar LCF dataset, however, historical landslides were used as sample labels, differing from the recognized landslide areas from remote sensing imagery in this study. [Figure 15](#) illustrates the receiver operating characteristic (ROC) curves of four 3D-CNN and the GBDT + LR models, evaluated by the area under the curve

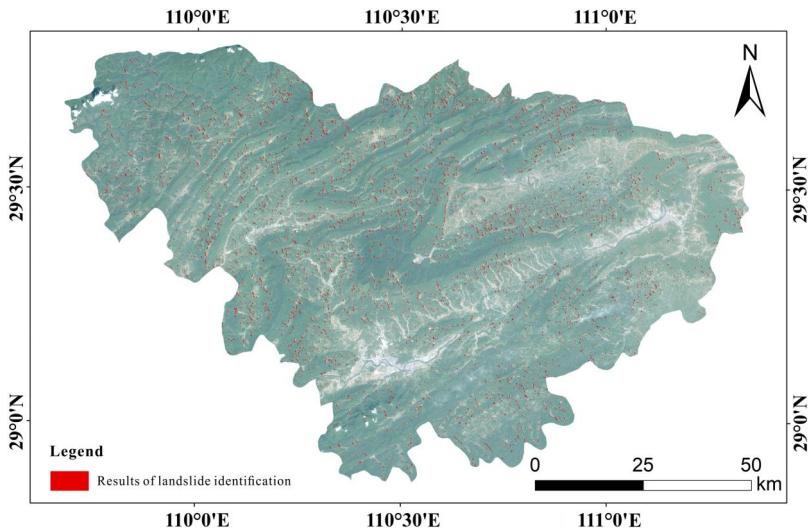


Figure 13. Results of landslide identification on the Sentinel 2 imagery.

Table 5. Hyperparameters of the models.

| No. | Parameters | Values |
|-----|---------------------------|----------------------|
| 1 | Convolutional kernel size | (3,3) |
| 2 | Pooling size | (2,2) |
| 3 | Loss function | Binary cross-entropy |
| 4 | Optimizer | Adam |
| 5 | Epoch | 25 |
| 6 | Batch size | 64 |
| 7 | Learning rate | 0.01 |
| 8 | Activation function | ReLU |

(AUC). **Table 6** lists the accuracy, precision, recall, F1-score, and AUC values of the four CNN methods and the GBDT + LR model on the testing dataset. In comparison to an AUC value of 0.8168 of the GBDT + LR model, the AUC values of four CNN models exceed 92%, and the accuracies are equal to or greater than 97%. The CNN model with the multi-scale sampling strategy achieves the highest accuracy and AUC value, while its recall and F1-score are significantly higher than those of other three models, with the F1-score being more than 10% higher than those of other three models, indicating its better prediction ability for positive examples.

Evaluating the statistical significance of performance differences among the models often necessitates non-parametric tests, and McNemar's test and Wilcoxon signed-rank test have been widely applied for this purpose (de Leeuw et al. 2006; Pham BT et al. 2019). To discern statistical differences among the models, McNemar's test was conducted using the testing dataset, as shown in **Table 7**. Among the models with different sampling scales, the multi-scale model exhibits statistically superior results, underscoring the efficacy of the multiscale sampling strategy.

4.2.3. Landslide susceptibility mapping

The proposed four 3D-CNN models were employed for LSM in the study area, and five levels of landslide susceptibility were ranked using natural breakpoint

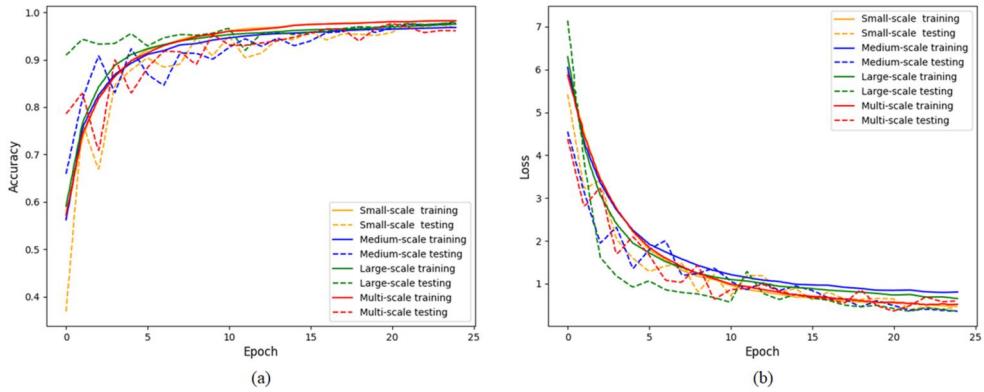


Figure 14. (a) Accuracy and (b) loss curves of the four CNN models for LSM.

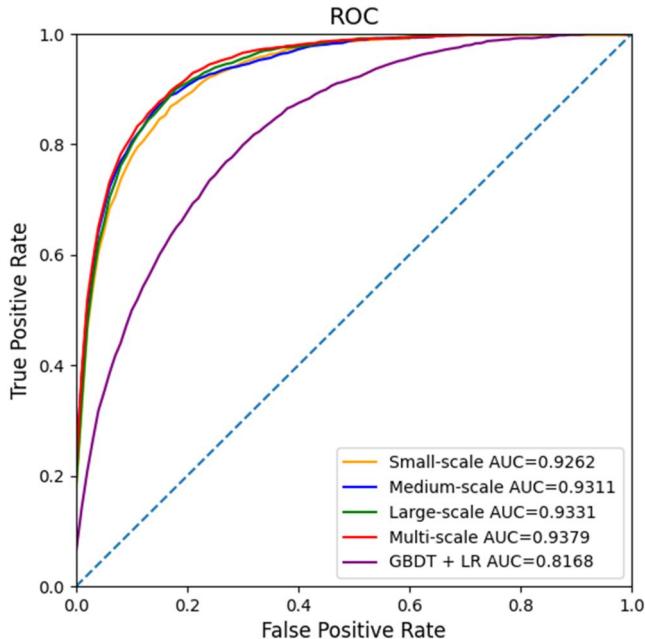


Figure 15. Receiver operating characteristic curves of the four 3D-CNN and GBDT + LR models.

classification. As shown in Figures 16(a,b), the predictions of the two models of small- and medium-scale differ significantly from the actual situations, with most of the landslides being ranked as very low or very high levels landslide susceptibility, and the eastern part of the study area being ranked as very high susceptibility, without sufficient predictive capacity for the northern, western and southeastern parts. The predictive results of the models of large- and multi-scale are more consistent with the actual situations, as shown in Figure 16(c,d), and the western, eastern, and southeastern parts of the study have been predicted as the highest landslide susceptibility, whose spatial distribution pattern is similar to that of the mountainous areas. Moreover, the zones of very high landslide susceptibility predicted by the multi-scale

**Table 6.** Evaluation of the four 3D CNN and GBDT + LR models.

| NO. | Models | Accuracy | Precision | Recall | F1-score | AUC |
|-----|--------------|----------|-----------|--------|----------|--------|
| 1 | Small-scale | 0.97 | 0.90 | 0.51 | 0.51 | 0.9262 |
| 2 | Medium-scale | 0.97 | 0.99 | 0.51 | 0.50 | 0.9311 |
| 3 | Large-scale | 0.98 | 0.99 | 0.50 | 0.50 | 0.9331 |
| 4 | Multi-scale | 0.98 | 0.83 | 0.57 | 0.61 | 0.9379 |
| 5 | GBDT + LR | 0.81 | 0.80 | 0.57 | 0.67 | 0.8168 |

Table 7. McNemar's test results using the testing dataset.

| Model pairs | McNemar's (χ^2) | p-value | Significant at | |
|-----------------------------|------------------------|---------------|----------------|-----|
| | | | 90% | 95% |
| Multi-scale vs Small-scale | 60.220 | $p < 0.00001$ | Yes | Yes |
| Multi-scale vs Medium-scale | 63.856 | $p < 0.00001$ | Yes | Yes |
| Multi-scale vs Large-scale | 11.811 | $p = 0.00058$ | Yes | Yes |
| Multi-scale vs GBDT + LR | 12.397 | $p = 0.00042$ | Yes | Yes |

model predominantly coincide with strata housing historical landslide occurrences and recognized landslide areas.

The CNN model with multi-scale sampling strategy has two advantages in landslide susceptibility analysis. On the one hand, landslides are controlled by complex influencing factors. Multi-scale sampling strategy can capture the information of landslide location and the environmental information near the landslide point through marching window. On the other hand, different scales increase the receptive field for 3D-CNN and enhance the acquisition ability of deep information, which is conducive to exploring the complex linear or nonlinear relationship between LCFs and landslide occurrences.

4.3. Limitations

In this study, the multispectral characteristics of four bands (bands 2, 3, 4, and 8) from Sentinel 2 imagery are primarily combined with DL semantic segmentation to identify landslide inventory. The texture and index characteristics of remote sensing imagery, mainly expressed in terms of both land cover texture and topographic index, can also serve as important basis for landslide recognition. Currently, the input to the attention U-Net consists of the raw band information. Currently, band fusion information is not considered as input features. For instance, the NDVI, an important vegetation index calculated from the red and near-infrared bands, effectively distinguishes vegetation from bare soil and can be further evaluated and utilized in future research to improve landslide identification accuracy (Shahabi et al. 2021). Moreover, landslide recognition can be further enhanced by incorporating physical models, e.g. hydrological models, geological models, and terrain models. While our proposed model achieved accurate landslide recognition results, it is important to note that not all landslides were successfully predicted, and the result exhibited more or less bias when integrating into LSM process. Xu et al. (2023) proposed a landslide image identification method, which can handle multimodal data in landslides and add knowledge as features to the landslide image identification model. Currently, knowledge graphs have been increasingly applied in the field of geosciences; therefore,

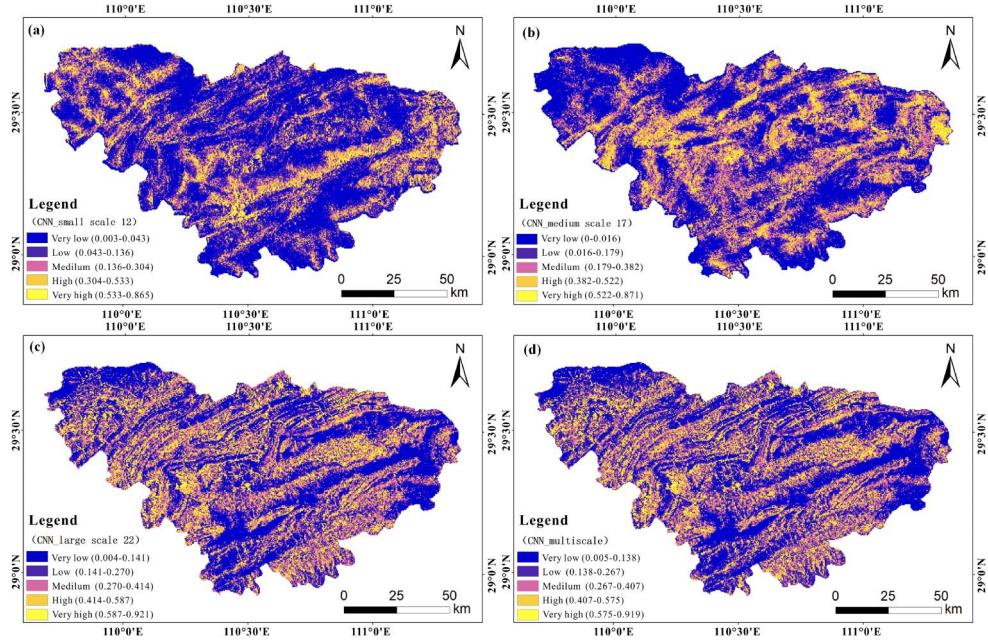


Figure 16. Landslide susceptibility maps using 3D-CNN models of (a) small-scale, (b) medium-scale, (c) large-scale, and (d) multi-scale.

combining knowledge mapping with remote sensing imagery will be a focus of our future research work.

The process of LSM often involves merging LCFs from diverse sources, necessitating up-sampling and down-sampling, especially for raster-based thematic LCFs. Additionally, the manual area delineation of investigated landslide in the Wulingyuan District were crucial for training the attention U-Net classification model using the Sentinel imagery. However, these steps introduce significant uncertainties due to varying resolutions and scales across data sources. Consequently, these uncertainties impact the overall uncertainty of LSM outcomes. Luo et al. (2019) discussed the classification uncertainty of landslide susceptibility maps using the rank probability score (RPS) series in Qingchuan county, China. Future research will prioritize quantitatively evaluating uncertainties associated with LCFs' up-sampling and down-sampling and landslide area delineation, shedding light on their impact on LSM outcomes, and offering a scientific foundation for improved landslide management and prevention strategies.

In LSM, some landslide sites being classified into the very low or very high landslide susceptibility categories. This discrepancy may be attributed to the rare occurrence of landslides, leading to a significant data imbalance between positive and negative samples for DL training, thereby affecting the generalization of the proposed model. To enhance the stability of the networks training and mitigate the risk of local optimum, He et al. (2019) have employed structural similarity (SSIM) as a loss function for tasks like road extraction. Similarly, Wan et al. (2021) introduced a hybrid loss function to CNNs to address category imbalance. Meanwhile, solving the



category imbalance at the data level can also improve the generalization ability of LSM models. Gupta and Shukla (2023) used two undersampling techniques, i.e. easy ensemble (EE) and balance cascade (BC), for reducing the effect of imbalance in data. Consequently, these aspects may play a significant role in LSM in the future.

5. Conclusions

The main work of this study is to construct a 3D-CNN model using multi-scale fused LCFs (e.g. topography, hydrology, meteorology, geology, and human activities) and the recognized landslide inventory from Sentinel 2 remote sensing imagery, to produce LSMs for Zhangjiajie City, Hunan Province, China. It completes two parts of research work, including landslide inventory recognition from remote sensing imagery and landslide susceptibility analysis by a multi-scale sampling 3D-CNN method. The specific research conclusions are as follows:

(1) An attention U-Net model of remote sensing imagery semantic segmentation for landslide inventory recognition is proposed, combining focal loss and Bayesian optimization. The experimental results demonstrate that the accuracy of the proposed model in this study reaches 93.5%, F1-score reaches 82%, and mIOU reaches 73.5%, largely solving the problem of category imbalance caused by landslides as small probability events. The model successfully extracted a more complete landslide boundary in the large-scale identification of landslide inventory in the study area, laying the foundation for the subsequent evaluation of landslide susceptibility.

(2) In the landslide susceptibility analysis, to extract and analyse complex spatial relationships between landslides and LCFs, this study constructs a 3D-CNN model and introduces a multi-scale sampling strategy to mine the multi-scale neighborhood information and deep information of landslides. The results show that the 3D-CNN model based on the multi-scale sampling strategy achieves the highest AUC, accuracy, precision, recall, and F1_score. The proposed 3D-CNNs with multi-scale sampling capture the spatial correlation between pixels and neighbouring features by setting different sampling scales and convolution operations. The hybrid multiscale 3D-CNN model performs well in predictive mapping to assist decision makers in large-scale land-use planning and geohazard prevention.

Acknowledgments

The authors would like to express their gratitude to the MapGIS Laboratory Co-Constructed by the National Engineering Research Centre for Geographic Information System of China and Central South University, for providing MapGIS® software (Wuhan Zondy Cyber-Tech Co. Ltd., Wuhan, China). We also thank Mr. Dongliang Huang (Senior Engineer at the Hunan Provincial Planning Institute of Land and Resources) for providing and processing the dataset.

Code availability statement

The source code (CNN4LSM) for this study is available in Python at Github (<https://github.com/csugeo3d-ai/CNN4LSM>).

Disclosure statement

The authors declare no conflicts of interest.

Funding

This study was supported by grants from the Hunan Provincial Natural Science Foundation (Grant Nos. 2022JJ30708 and 2023JJ60188), the Hunan Provincial Natural Resource Science and Technology Planning Program (Grant No. 20230123XX), the Changsha Municipal Natural Science Foundation (Grant No. kq2208054), and the National Natural Science Foundation of China (Grant No. 42072326).

Data availability statement

The dataset used in this study is not publicly available due to a data privacy agreement with the Hunan Institute of Geological Disaster Investigation and Monitoring. However, it can be obtained from the corresponding author upon reasonable request.

References

- Arulbalaji P, Padmalal D, Sreelash K. 2019. GIS and AHP techniques based delineation of groundwater potential zones: a case study from Southern Western Ghats, India. *Sci Rep.* 9(1):2082. doi: [10.1038/s41598-019-38567-x](https://doi.org/10.1038/s41598-019-38567-x).
- Azrafza M, Azrafza M, Akgün H, Atkinson PM, Derakhshani R. 2021. Deep learning-based landslide susceptibility mapping. *Sci Rep.* 11(1):24112. doi: [10.1038/s41598-021-03585-1](https://doi.org/10.1038/s41598-021-03585-1).
- Botzen WJW, Aerts JCJH, van den Bergh JCJM. 2013. Individual preferences for reducing flood risk to near zero through elevation [Article]. *Mitig Adapt Strateg Glob Change.* 18(2): 229–244. doi: [10.1007/s11027-012-9359-5](https://doi.org/10.1007/s11027-012-9359-5).
- Bui DT, Tsangaratos P, Nguyen V-T, Liem NV, Trinh PT. 2020. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *CATENA.* 188:104426. doi: [10.1016/j.catena.2019.104426](https://doi.org/10.1016/j.catena.2019.104426).
- Chacón J, Irigaray C, Fernández T, El Hamdouni R. 2006. Engineering geology maps: landslides and geographical information systems [Review]. *Bull Eng Geol Environ.* 65(4):341–411. doi: [10.1007/s10064-006-0064-z](https://doi.org/10.1007/s10064-006-0064-z).
- Chen W, Shahabi H, Shirzadi A, Hong H, Akgun A, Tian Y, Liu J, Zhu AX, Li S. 2019. Novel hybrid artificial intelligence approach of bivariate statistical-methods-based kernel logistic regression classifier for landslide susceptibility modeling [Article]. *Bull Eng Geol Environ.* 78(6):4397–4419. doi: [10.1007/s10064-018-1401-8](https://doi.org/10.1007/s10064-018-1401-8).
- Chung C-JF, Fabbri AG. 1999. Probabilistic prediction models for landslide hazard mapping. *Photogramm Eng Remote Sens.* 65(12):1389–1399.
- de Leeuw J, Jia H, Yang L, Liu X, Schmidt K, Skidmore AK. 2006. Comparing accuracy assessments to infer superiority of image classification methods. *Int J Remote Sens.* 27(1):223–232. doi: [10.1080/01431160500275762](https://doi.org/10.1080/01431160500275762).
- Dikshit A, Satyam N, Pradhan B. 2019. Estimation of rainfall-induced landslides using the TRIGRS model. *Earth Syst Environ.* 3(3):575–584. doi: [10.1007/s41748-019-00125-w](https://doi.org/10.1007/s41748-019-00125-w).
- Dobrovolsky E. 1971. Landslide susceptibility in and near anchorage as interpreted from topographic and geologic maps, in the great Alaska earthquake of 1964-Geology volume. Publication 1603. US Geological survey open file report. p. 86–329.
- Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena SR, Tiede D, Aryal J. 2019. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing.* 11(2):196. doi: [10.3390/rs11020196](https://doi.org/10.3390/rs11020196).

- Gupta SK, Shukla DP. 2023. Handling data imbalance in machine learning based landslide susceptibility mapping: a case study of Mandakini River Basin, North-Western Himalayas. *Landslides*. 20(5):933–949. doi: [10.1007/s10346-022-01998-1](https://doi.org/10.1007/s10346-022-01998-1).
- Guzzetti F, Carrara A, Cardinali M, Reichenbach P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*. 31(1-4):181–216. doi: [10.1016/S0169-555X\(99\)00078-1](https://doi.org/10.1016/S0169-555X(99)00078-1).
- Guzzetti F, Reichenbach P, Cardinali M, Galli M, Ardizzone F. 2005. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*. 72(1-4):272–299. doi: [10.1016/j.geomorph.2005.06.002](https://doi.org/10.1016/j.geomorph.2005.06.002).
- He H, Yang D, Wang S, Wang S, Li Y. 2019. Road extraction by using atrous spatial pyramid pooling integrated encoder-decoder network and structural similarity loss. *Remote Sensing*. 11(9):1015. doi: [10.3390/rs11091015](https://doi.org/10.3390/rs11091015).
- Hong H, Liu J, Bui DT, Pradhan B, Acharya TD, Pham BT, Zhu AX, Chen W, Ahmad BB. 2018. Landslide susceptibility mapping using J48 decision tree with adaboost, bagging and rotation forest ensembles in the Guangchang area (China). *CATENA*. 163:399–413. doi: [10.1016/j.catena.2018.01.005](https://doi.org/10.1016/j.catena.2018.01.005).
- Huan YK, Song L, Khan U, Zhang BY. 2023. Stacking ensemble of machine learning methods for landslide susceptibility mapping in Zhangjiajie City, Hunan Province, China. *Environ Earth Sci*. 82(1):35. doi: [10.1007/s12665-022-10723-z](https://doi.org/10.1007/s12665-022-10723-z).
- Huang F, Chen J, Du Z, Yao C, Huang J, Jiang Q, Chang Z, Li S. 2020. Landslide susceptibility prediction considering regional soil erosion based on machine-learning models. *IJGI*. 9(6): 377. doi: [10.3390/ijgi9060377](https://doi.org/10.3390/ijgi9060377).
- Huang F, Zhang J, Zhou C, Wang Y, Huang J, Zhu L. 2020. A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction [Article]. *Landslides*. 17(1):217–229. doi: [10.1007/s10346-019-01274-9](https://doi.org/10.1007/s10346-019-01274-9).
- Ji S, Yu D, Shen C, Li W, Xu Q. 2020. Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks [Article]. *Landslides*. 17(6):1337–1352. doi: [10.1007/s10346-020-01353-2](https://doi.org/10.1007/s10346-020-01353-2).
- Kingma DP, Ba JL. 2015. Adam: a method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015, May 7, 2015 - May 9, 2015; San Diego, CA, United states.
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning [Review]. *Nature*. 521(7553):436–444. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- Lecun Y, Bottou L, Bengio Y, Haffner P. 1998. Gradient-based learning applied to document recognition. *Proc IEEE*. 86(11):2278–2324. doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791).
- Li H, Li X, Yuan F, Jowitt SM, Zhang M, Zhou J, Zhou T, Li X, Ge C, Wu B. 2020. Convolutional neural network and transfer learning based mineral prospectivity modeling for geochemical exploration of Au mineralization within the Guandian-Zhangbaling area, Anhui Province, China. *Appl Geochem*. 122(1):104747. doi: [10.1016/j.apgeochem.2020.104747](https://doi.org/10.1016/j.apgeochem.2020.104747).
- Lin Q, Lima P, Steger S, Glade T, Jiang T, Zhang J, Liu T, Wang Y. 2021. National-scale data-driven rainfall induced landslide susceptibility mapping for China by accounting for incomplete landslide data. *Geosci Front*. 12(6):101248. doi: [10.1016/j.gsf.2021.101248](https://doi.org/10.1016/j.gsf.2021.101248).
- Liu S, Deng W. 2015. Very deep convolutional neural network based image classification using small training sample size. 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015, November 3, 2016 - November 6, 2016; Kuala Lumpur, Malaysia.
- Luo X, Lin F, Chen Y, Zhu S, Xu Z, Huo Z, Yu M, Peng J. 2019. Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. *Sci Rep*. 9(1):15369. doi: [10.1038/s41598-019-51941-z](https://doi.org/10.1038/s41598-019-51941-z).
- Mao Z, Shi S, Li H, Zhong J, Sun J. 2022. Landslide susceptibility assessment using triangular fuzzy number-analytic hierarchy processing (TFN-AHP), contributing weight (CW) and random forest weighted frequency ratio (RF weighted FR) at the Pengyang county, Northwest China. *Environ Earth Sci*. 81(3):86. doi: [10.1007/s12665-022-10193-3](https://doi.org/10.1007/s12665-022-10193-3).

- Mariano DN, Carotenuto F, Cevasco A, Confuorto P, Di Martire D, Firpo M, Pepe G, Raso E, Calcaterra D. **2020**. Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. *Landslides*. 17(8):1897–1914. doi: [10.1007/s10346-020-01392-9](https://doi.org/10.1007/s10346-020-01392-9).
- Ministry of Land and Resources of the People's Republic of China. **2014**. Specification of comprehensive survey for landslide, collapse, and debris flow (1:50,000). Beijing: Ministry of Land and Resources of the People's Republic of China.
- Montgomery DR, Dietrich WE. **1994**. A physically based model for the topographic control on shallow landsliding. *Water Resour Res*. 30(4):1153–1171. doi: [10.1029/93WR02979](https://doi.org/10.1029/93WR02979).
- Nefeslioglu HA, Duman TY, Durmaz S. **2008**. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). *Geomorphology*. 94(3-4):401–418. doi: [10.1016/j.geomorph.2006.10.036](https://doi.org/10.1016/j.geomorph.2006.10.036).
- Nery TD, Vieira BC. **2015**. Susceptibility to shallow landslides in a drainage basin in the Serra do Mar, Sao Paulo, Brazil, predicted using the SINMAP mathematical model [Article]. *Bull Eng Geol Environ*. 74(2):369–378. doi: [10.1007/s10064-014-0622-8](https://doi.org/10.1007/s10064-014-0622-8).
- Nguyen V-T, Tran TH, Ha NA, Ngo VL, Nadhir A-A, Tran VP, Duy Nguyen H, M. A M, Amini A, Prakash I, et al. **2019**. GIS based novel hybrid computational intelligence models for mapping landslide susceptibility: a case study at Da Lat City, Vietnam. *Sustainability*. 11(24):7118. doi: [10.3390/su11247118](https://doi.org/10.3390/su11247118).
- Oktay O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S,Y, Hammerla N, Kainz B, et al. **2018**. Attention U-Net: learning where to look for the pancreas. [accessed April 01, 2018]:[arXiv:1804.03999 p.]. <https://ui.adsabs.harvard.edu/abs/2018arXiv180403999O>. doi: [10.48550/arXiv.1804.03999](https://doi.org/10.48550/arXiv.1804.03999).
- Pham QB, Achour Y, Ali SA, Parvin F, Vojtek M, Vojteková J, Al-Ansari N, Achu AL, Costache R, Khedher KM, et al. **2021**. A comparison among fuzzy multi-criteria decision making, bivariate, multivariate and machine learning models in landslide susceptibility mapping. *Geomatics Nat Hazards Risk*. 12(1):1741–1777. doi: [10.1080/19475705.2021.1944330](https://doi.org/10.1080/19475705.2021.1944330).
- Pham BT, Jaafari A, Prakash I, Bui DT. **2019**. A novel hybrid intelligent model of support vector machines and the multiboost ensemble for landslide susceptibility modeling. *Bull Eng Geol Environ*. 78(4):2865–2886. doi: [10.1007/s10064-018-1281-y](https://doi.org/10.1007/s10064-018-1281-y).
- Pham BT, Tien Bui D, Prakash I, Dholakia MB. **2017**. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *CATENA*. 149:52–63. doi: [10.1016/j.catena.2016.09.007](https://doi.org/10.1016/j.catena.2016.09.007).
- Ramesh V, Anbazhagan S. **2015**. Landslide susceptibility mapping along Kolli hills Ghat road section (India) using frequency ratio, relative effect and fuzzy logic models [Article]. *Environ Earth Sci*. 73(12):8009–8021. doi: [10.1007/s12665-014-3954-6](https://doi.org/10.1007/s12665-014-3954-6).
- Ronneberger O, Fischer P, Brox T. **2015**. U-Net: convolutional networks for biomedical image segmentation. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015, October 5, 2015 - October 9, Munich, Germany.
- Ruder S. **2016**. An overview of gradient descent optimization algorithms. [accessed September 01, 2016]:[arXiv:1609.04747 p.]. <https://ui.adsabs.harvard.edu/abs/2016arXiv160904747R>. doi: [10.48550/arXiv.1609.04747](https://doi.org/10.48550/arXiv.1609.04747).
- Shahabi H, Rahimzad M, Tavakkoli Piralilou S, Ghorbanzadeh O, Homayouni S, Blaschke T, Lim S, Ghamisi P. **2021**. Unsupervised deep learning for landslide detection from multispectral sentinel-2 imagery. *Remote Sensing*. 13(22):4698. doi: [10.3390/rs13224698](https://doi.org/10.3390/rs13224698).
- Thi Ngo PT, Panahi M, Khosravi K, Ghorbanzadeh O, Kariminejad N, Cerdá A, Lee S. **2021**. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geosci Front*. 12(2):505–519. doi: [10.1016/j.gsf.2020.06.013](https://doi.org/10.1016/j.gsf.2020.06.013).
- Tien Bui D, Shahabi H, Omidvar E, Shirzadi A, Geertsema M, Clague JJ, Khosravi K, Pradhan B, Pham BT, Chapi K, et al. **2019**. Shallow landslide prediction using a novel hybrid functional machine learning algorithm. *Remote Sensing*. 11(8):931. doi: [10.3390/rs11080931](https://doi.org/10.3390/rs11080931).
- Victoria AH, Maragatham G. **2021**. Automatic tuning of hyperparameters using Bayesian optimization [Article.]. *Evolving Systems*. 12(1):217–223. doi: [10.1007/s12530-020-09345-2](https://doi.org/10.1007/s12530-020-09345-2).

- Wan J, Xie Z, Xu Y, Chen S, Qiu Q. 2021. DA-RoadNet: a dual-attention network for road extraction from high resolution satellite imagery [Article]. *IEEE J Sel Top Appl Earth Observations Remote Sensing*. 14:6302–6315. doi: [10.1109/JSTARS.2021.3083055](https://doi.org/10.1109/JSTARS.2021.3083055).
- Wang Y, Fang Z, Hong H. 2019. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China [Article]. *Sci Total Environ*. 666:975–993. doi: [10.1016/j.scitotenv.2019.02.263](https://doi.org/10.1016/j.scitotenv.2019.02.263).
- Wang H, Zhang L, Yin K, Luo H, Li J. 2021. Landslide identification using machine learning [Article]. *Geoscience Frontiers*. 12(1):351–364. doi: [10.1016/j.gsf.2020.02.012](https://doi.org/10.1016/j.gsf.2020.02.012).
- Xu B, Zhang C, Liu W, Huang J, Su Y, Yang Y, Jiang W, Sun W. 2023. Landslide Identification Method Based on the FKGRNet Model for remote sensing images. *Remote Sensing*. 15(13):3407. doi: [10.3390/rs15133407](https://doi.org/10.3390/rs15133407).
- Yi Y, Zhang Z, Zhang W, Jia H, Zhang J. 2020. Landslide susceptibility mapping using multi-scale sampling strategy and convolutional neural network: a case study in Jiuzhaigou region. *CATENA*. 195:104851. doi: [10.1016/j.catena.2020.104851](https://doi.org/10.1016/j.catena.2020.104851).
- Zhang G, Cai Y, Zheng Z, Zhen J, Liu Y, Huang K. 2016. Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in Huizhou, China. *CATENA*. 142:233–244. doi: [10.1016/j.catena.2016.03.028](https://doi.org/10.1016/j.catena.2016.03.028).
- Zhang B, Li M, Huan Y, Khan U, Wang L, Wang F. 2023. Bedrock mapping based on terrain weighted directed graph convolutional network using stream sediment geochemical samplings. *Trans Nonferrous Met Soc China*. 33(9):2299–2814. doi: [10.1016/S1003-6326\(23\)66299-5](https://doi.org/10.1016/S1003-6326(23)66299-5).
- Zhang B, Li M, Li W, Jiang Z, Khan U, Wang L, Wang F. 2021. Machine learning strategies for lithostratigraphic classification based on geochemical sampling data: A case study in area of Chahanwusu River, Qinghai Province, China. *J Cent South Univ*. 28(5):1422–1447. doi: [10.1007/s11771-021-4707-9](https://doi.org/10.1007/s11771-021-4707-9).
- Zhang B, Xu K, Khan U, Li X, Du L, Xu Z. 2023. A lightweight convolutional neural network with end-to-end learning for three-dimensional mineral prospectivity modeling: a case study of the Sanhetun Area, Heilongjiang Province, Northeastern China. *Ore Geol Rev*. 163: 105788. doi: [10.1016/j.oregeorev.2023.105788](https://doi.org/10.1016/j.oregeorev.2023.105788).
- Zhao P, Masoumi Z, Kalantari M, Aflaki M, Mansourian A. 2022. A GIS-based landslide susceptibility mapping and variable importance analysis using artificial intelligent training-based methods. *Remote Sensing*. 14(1):211. doi: [10.3390/rs14010211](https://doi.org/10.3390/rs14010211).
- Zhao S, Zhao Z. 2021. A comparative study of landslide susceptibility mapping using SVM and PSO-SVM models based on grid and slope Units. *Math Prob Eng*. 2021:1–15. doi: [10.1155/2021/8854606](https://doi.org/10.1155/2021/8854606).
- Zuo R, Xiong Y. 2017. Big data analytics of identifying geochemical anomalies supported by machine learning methods. *Nat Resour Res*. 27(1):5–13. doi: [10.1007/s11053-017-9357-0](https://doi.org/10.1007/s11053-017-9357-0).