International Journal of Geo-Information

MDPI

Article

Landslide Recognition Based on Machine Learning Considering Terrain Feature Fusion

Jincan Wang [1], Zhiheng Wang [2,*], Liyao Peng [3] and Chenzhihao Qian [4]

[1] School of Earth Sciences, China University of Geosciences (Wuhan), Wuhan 430079, China; wangjincan@cug.edu.cn
[2] School of Geology and Geomatics, Tianjin Chengjian University, Tianjin 300384, China
[3] School of Resources and Environmental Engineering, Wuhan University of Technology, Wuhan 430079, China; peng_liyao@whut.edu.cn
[4] Laboratory Cultivation Base of Environment Process and Digital Simulation, Beijing Laboratory of Water Resources Security, Capital Normal University, Beijing 100048, China; 2230902110@cnu.edu.cn
* Correspondence: wangzhiheng@tcu.edu.cn

Abstract: Landslides are one of the major disasters that exist worldwide, posing a serious threat to human life and property safety. Rapid and accurate detection and mapping of landslides are crucial for risk assessment and humanitarian assistance in affected areas. To achieve this goal, this study proposes a landslide recognition method based on machine learning (ML) and terrain feature fusion. Taking the Dawan River Basin in Detuo Township and Tianwan Yi Ethnic Township as the research area, firstly, landslide-related data were compiled, including a landslide inventory based on field surveys, satellite images, historical data, high-resolution remote sensing images, and terrain data. Then, different training datasets for landslide recognition are constructed, including full feature datasets that fusion terrain features and remote sensing features and datasets that only contain remote sensing features. At the same time, different ratios of landslide to non-landslide (or positive/negative, P/N) samples are set in the training data. Subsequently, five ML algorithms, including Extreme Gradient Boost (XGBoost), Adaptive Boost (AdaBoost), Light Gradient Boost (LightGBM), Random Forest (RF), and Convolutional Neural Network (CNN), were used to train each training dataset, and landslide recognition was performed on the validation area. Finally, accuracy (A), precision (P), recall (R), F1 score (F1), and intersection over union (IOU) were selected to evaluate the landslide recognition ability of different models. The research results indicate that selecting ML models suitable for the study area and the ratio of the P/N samples can improve the A, R, F1, and IOU of landslide identification results, resulting in more accurate and reasonable landslide identification results; Fusion terrain features can make the model recognize landslides more comprehensively and align better with the actual conditions. The best-performing model in the study is LightGBM. When the input data includes all features and the P/N sample ratio is optimal, the A, P, R, F1, and IOU of landslide recognition results for this model are 97.47%, 85.40%, 76.95%, 80.95%, and 71.28%, respectively. Compared to the landslide recognition results using only remote sensing features, this model shows improvements of 4.51%, 35.66%, 5.41%, 22.27%, and 29.16% in A, P, R, F1, and IOU, respectively. This study serves as a valuable reference for the precise and comprehensive identification of landslide areas.

Keywords: landslide hazards; landslide identification; topographic features; machine learning; sample ratio

check for updates

Citation: Wang, J.; Wang, Z.; Peng, L.; Qian, C. Landslide Recognition Based on Machine Learning Considering Terrain Feature Fusion. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 306. https://doi.org/10.3390/ijgi13090306

Academic Editors: Wolfgang Kainz, Dimitrios D. Alexakis and Adrianos Retalis

Received: 27 July 2024
Revised: 15 August 2024
Accepted: 22 August 2024
Published: 28 August 2024

Copyright: © 2024 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

Landslides are one of the most destructive natural disasters, characterized by their wide distribution and frequent occurrence. They cause significant loss of life and property in many countries around the world each year. Additionally, landslides mainly occur in mountainous regions with complex terrain and difficult access, which adds to the

challenges of on-site investigation and rapid post-disaster rescue and reconstruction efforts [1–4]. Landslide identification is the process of determining the extent of a landslide area by studying its morphology and characteristics. The technical methods for landslide recognition using remote sensing images mainly include visual interpretation, pixel-based recognition methods, object-oriented recognition methods, and machine learning-based recognition methods [5–8]. By utilizing these technological approaches, the location of landslides can be identified promptly, saving significant manpower and material resources for rescue operations. This is essential for risk assessment in disaster-prone areas and facilitating humanitarian assistance [9–12].

The traditional methods for identifying and mapping landslides mainly rely on on-site investigations. However, this approach is challenging to implement and inefficient in mountainous regions. The survey results are often not comprehensive and accurate enough [13]. With the emergence of remote sensing technology, satellite remote sensing images are considered the primary data source for detecting landslides and updating inventory maps. Researchers can identify landslides through visual interpretation of high-resolution remote sensing images [14,15]. Although the accuracy of visually interpreting landslide areas is relatively high, it heavily relies on the experience of professional personnel. This reliance results in drawbacks such as being time-consuming, labor-intensive, highly subjective, and lacking automation and intelligence [7,16,17]. With the development of computer equipment and artificial intelligence, researchers have begun to use ML techniques for landslide identification. ML methods are an automated modeling approach used for data analysis. They can learn the fundamental relationships present in the data, build analysis models, and generate precise results through iterative learning processes [16,18–21]. Using ML for landslide identification not only compensates for the shortcomings of traditional methods, such as being time-consuming, labor-intensive, and highly subjective but also significantly outperforms traditional methods in terms of accuracy and predictive performance [22,23].

The current ML methods used for landslide recognition include Logistic Regression (LR) [24], Artificial Neural Networks (ANN) [6,25], K-Nearest Neighbors (KNN) [26,27], Support Vector Machines (SVM) [6,28,29], RF [24,30], CNN [31–33], etc. Due to differences in climate conditions and geographical locations, the models suitable for various research areas also vary. Comparing different methods helps to better evaluate the applicability of each method in the research area, thereby ensuring the reliability of landslide identification results [34]. Several studies have been published comparing the performance of various ML models in landslide identification. Wang, H. et al. [24] utilized five ML algorithms to detect landslides in Lantau Island, Hong Kong. The research results indicate that various methods perform well in landslide identification. In terms of accuracy, CNN ranks first, Boosting methods rank second, followed by RF, LR, and SVM; Singh, P. et al. [35] used SVM, Classification and Regression Trees (CART), Minimum Distance, RF, and other methods to identify landslides in the Rudraprayang area. They found that RF and SVM outperformed other ML methods. In landslide-related research utilizing ML, the ratio of P/N samples in the training data is typically 1:1 [24,36,37]. However, an inappropriate ratio of P/N samples may result in inadequate training of the ML model or data contamination, consequently diminishing the model's performance [38–40]. Therefore, finding suitable ML methods and appropriate P/N sample ratios for specific research areas is crucial for obtaining more accurate and reasonable landslide identification results.

Landslides often occur in mountainous areas and may be influenced by the terrain and topography of the study area, so terrain features can be utilized to aid in identifying landslides [9,13,41,42]. At present, some scholars use high-resolution remote sensing images as the main data source and terrain features as auxiliary information, such as fusing remote sensing images with digital elevation models and derived information to identify landslides [43–46]. Tavakkoli Piralilou, S. et al. [47] conducted landslide identification research in the Rasuwa area by combining three ML methods and multi-scale methods, using optical images and DEM as data sources. The overall accuracy of the identification results was close to 90%. Wang, H. et al. [18] fused multi-source data (including remote

sensing images and DEM) and ML technology to conduct co-seismic landslide detection in the Wenchuan earthquake and Jiuzhaigou Valley Scenic and Historic Interest Area earthquake zones. The results show that the proposed landslide identification method has superior performance and versatility. However, these studies did not quantify the effect of terrain features on landslide identification results, nor did they analyze in-depth the specific role of terrain feature fusion in landslide identification. There are also studies indicating that fusion terrain features may not improve the effectiveness of landslide identification [48]. Therefore, it is crucial to conduct ablation experiments to analyze the impact and role of terrain features on landslide recognition using ML technology for terrain feature fusion.

To address the aforementioned issues, this article focuses on the Dadu River Basin in Detuo Township and Tianwan Yi Ethnic Township as the research area. It proposes a landslide recognition method based on multiple ML techniques and terrain feature fusion. The main goals are to: (1) construct multiple ML models and consider the proportion of P/N samples in the model training data for landslide identification research; (2) compare the performance of different models using various evaluation indicators to determine the most suitable ML method and P/N sample ratio for landslide identification in the study area, aiming to achieve more accurate and reasonable landslide identification results; and (3) conduct ablation experiments that fusion terrain features in landslide identification and analyze the impact and role of terrain features on landslide identification in the study area.

2. Materials and Data

2.1. Overview of the Study Area

The study area is the Dadu River Basin in Detuo Township and Tianwan Yi Ethnic Township, Sichuan Province, with the geographic coordinates of 102°3′0″–102°18′04″ E and 29°24′20″–29°38′49″ N, and an elevation ranging from 924 to 3197 m. The average annual precipitation exceeds 600 mm. In terms of topography and geomorphology, the study area is situated in the hilly region at the western edge of the Sichuan Basin. It features complex and undulating topography, displaying a terrain pattern sloping from northwest to southeast. In terms of stratigraphic lithology, the main types of rock formations in the region include sandstone, shale, mudstone, limestone, etc. These types of rocks are characterized by easy erosion, numerous joints, and fissures, making them susceptible to damage by external forces. The complex geological conditions, frequent precipitation, strong natural erosion, and frequent earthquakes in the study area have led to frequent landslides and geological hazards in the region. These events pose a serious threat to the lives and properties of the local people.

In analyzing the landslide data in Detuo Township and Tianwan Yi Township, we observed that the landslides were primarily clustered within a specific area along the Dadu River within the region. The closer to the Dadu River, the higher the frequency of landslides, especially within 1200 m around the Dadu River. Landslides within this 1200 m range account for more than 80% of all landslides. For example, large landslides such as the Mogangling landslide (E 102°09′41″, N 29°37′30″) and the Lantianwan landslide (E 102°10′46″, N 29°35′27″) occur within this range [49], as shown in Figure 1. Among them, the Moganling landslide has a length of about 650 m, a width of about 700 m, an average thickness of 120 m, and a volume of about 2400×10^4 m^3. The terrain of the landslide area presents a gentle upward slope and a steep downward slope. The rock mass is significantly affected by the Moxi Fault and Daduhe Fault, showing a loose structure and dense cracks. The Lantianwan landslide has a longitudinal length of about 1000 m, a width of about 600 m, an average thickness of 50 m, and a volume of about 3.0×10^7 m^3. The bedrock where the landslide is located is Chengjiang granite, and the rock mass on the back wall of the landslide is in a strong to moderately weathered state. Therefore, in this study, a width of 1200 m on each side of the Dadu River within Detuo Township and Tianwan Yi Ethnic Township was selected as the study area, as shown in Figure 2.

Figure 1. (**a**) Mogangling landslide; and (**b**) Lantianwan landslide.

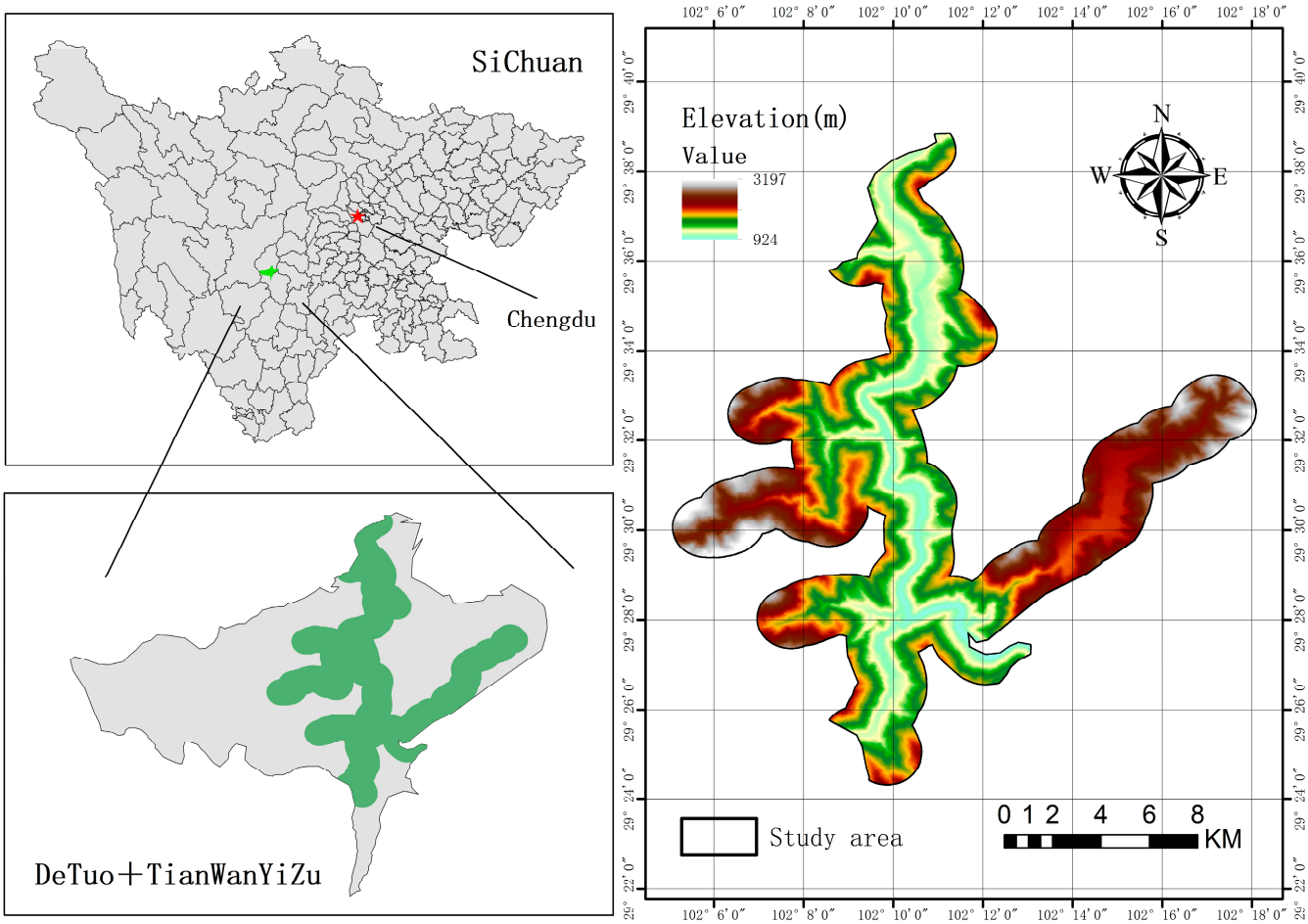

Figure 2. The study area's location.

2.2. Data

2.2.1. Basic Data

The remote sensing image was captured by the GF-2 satellite on 10 September 2022, with a spatial resolution of 2 M. Four spectral bands—blue (440–510 nm), green (520–590 nm), red (630–685 nm), and near-infrared (760–850 nm)—were primarily utilized in the study. The true-color optical image of the study area is depicted in Figure 3a. DEM data downloaded from NASA (https://search.asf.alaska.edu, accessed on 2 January 2024) were provided by Japan's Advanced Land Observing Satellite (ALOS). ALOS is mainly responsible for high-precision terrain mapping, natural disaster monitoring, and management, with a resolution of 12.5 M, as depicted in Figure 3b. The pixel-level landslide inventory utilized for training

and testing was obtained through manual interpretation, illustrated in Figure 3c,d. The training area in the Dadu River Basin within Tianwan Yi Township covers 80.84 km^2, while the validation area in the Dadu River Basin within Detuo Township spans 73.87 km^2.

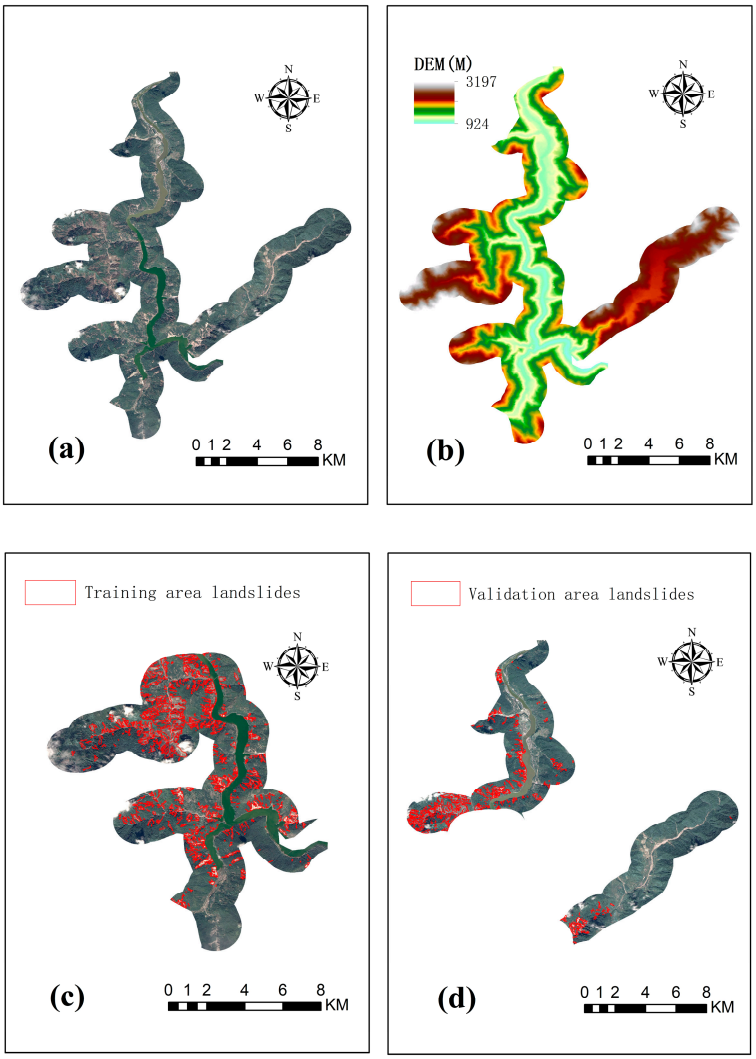

Figure 3. Basic data for landslide identification includes: (**a**) optical image on 10 September 2022; (**b**) DEM; (**c**) landslide area in Dadu River Basin, Tianwan Yi Ethnic Township; and (**d**) landslide area in Dadu River Basin, Detuo Township.

2.2.2. Feature Set Construction

Topographic features are essential in landslide development, providing the necessary spatial conditions for landslides to occur, including sufficient air frontage, i.e., the slope face of the slope. This is related to elevation characteristics, gradient, slope direction, and other factors [50,51]. This study comprehensively considers both remote sensing features (spectral, texture, and thematic indices of remote sensing images collectively referred to as remote sensing features) and terrain features to construct a multidimensional and comprehensive initial feature set. This approach makes full use of all valuable initial feature information for the rapid extraction of landslides. It also lays a foundation for subsequent comparative analysis of the features and the final improvement of landslide extraction accuracy. Table 1 lists all the features constructed based on high-resolution remote sensing images and digital elevation models, with a spatial resolution of 2 M for remote sensing features and 12.5 M for terrain features.

Table 1. Multidimensional features list.

Type	Feature Name	Characterization
Remote sensing features	Red	Red band
	Green	Green band
	Blue	Blue band
	Nir	Near-infrared band
	Gray level co-occurrence matrix (GLCM) contrast	This reflects the total amount of localized changes in the image.
	GLCM correlation	Reflects the length of extension of a specific gray value in a particular direction.
	GLCM dissimilarity	Similar to contrast, it is a measure of the amount of local variation in an image.
	GLCM entropy	This reflects the complexity of the image.
	GLCM homogeneity	The image reflects the local gray uniformity.
	GLCM mean	This reflects the uniformity of the gray tone in the image.
	GLCM second moment	This reflects the localized homogeneity of the image.
	GLCM variance	This reflects the roughness or fineness of the image.
	NDVI	Normalized Difference Vegetation Index, which has the formula (Nir − Red)/(Nir + Red).
	NDWI	Normalized Difference Water Index, which has the formula (Nir − Green)/(Nir + Green)
Topographic features	DEM	Elevation of a surface or feature in relation to a reference point.
	Slope	The degree of tilt of the Earth's surface.
	Aspect	The direction in which the Earth's surface is tilted.
	Relief	The degree of undulation on the Earth's surface.
	Planform curvature	The curvature of the terrain surface in the horizontal direction reflects the degree of curvature of contour lines.
	Profile curvature	The curvature of the terrain surface perpendicular to the contour line reflects the degree of unevenness of the slope surface.

3. Workflow and Methods

3.1. Workflow

This study utilized remote sensing imagery and DEM data to extract remote sensing and terrain features. Landslide identification was performed at the pixel level using XGBoost, AdaBoost, LightGBM, RF, and CNN methods, which have shown superior performance in landslide research [18,24,52]. The main workflow was as follows:

(1) Remote sensing feature data, such as spectra, textures, and thematic indices from remote sensing images in the study area, were extracted. Terrain feature data such as elevation, slope, and aspect of the study area were extracted based on DEM. Consequently, a landslide identification feature dataset was constructed.
(2) Based on the results of feature set construction, two different training datasets were designed: (a) containing only remote sensing features and (b) containing both remote sensing features and terrain features. Additionally, various proportions were set for the number of landslide and non-landslide samples in the training dataset, including 1:1, 1:2, 1:3, 1:4, and 1:5, respectively.
(3) XGBoost, AdaBoost, LightGBM, RF, and CNN were used for landslide recognition using various training datasets. The landslide identification performance of different models was evaluated using A, P, R, F1, and IOU.
(4) By evaluating the performance of different models, suitable ML methods and P/N sample ratios for the study area were identified, and accurate and reasonable landslide identification results were obtained for the test area.
(5) Ablation experiments were conducted to evaluate the influence of terrain feature indicators on landslide identification outcomes and offer additional explanations and

discussions. Simultaneously, the effects of ML models and P/N sample ratio on the results were assessed.

3.2. Methods

3.2.1. XGBoost

XGBoost is a ML tool based on the Gradient Boosting Decision Tree (GBDT) algorithm, proposed by Chen, T. et al. [53] in 2016. XGBoost uses a CART as the basic learner and optimizes GBDT to achieve integrated learning of multiple CARTs. Specifically, the method integrates several fundamental models, including decision trees for classification and regression, as well as linear models, to create a robust model for addressing ML problems like classification and regression. The principle of constructing an XGBoost model is illustrated in Figure 4. Initially, an initial tree (weak classifier) is built for training the model, and the residuals between the model's predicted and actual values are calculated. Subsequently, multiple rounds of iterations are conducted, with each iteration generating a new tree to capture the residuals from the previous prediction of the model. Finally, the outcomes of all the trees are weighted and combined to derive the final prediction. The predicted value Y_i is determined as follows:

$$Y_i = \sum_{K=1}^{K} f_k(x_i) \tag{1}$$

where K denotes all CART trees built, x_i denotes the features of the ith sample, and denotes the predicted value of the kth tree.

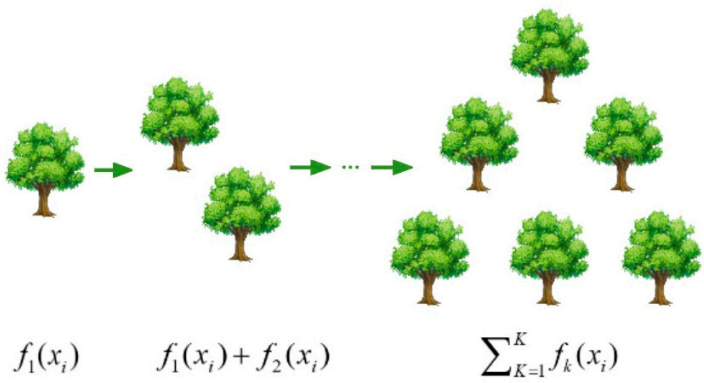

Figure 4. XGBoost conceptual model.

3.2.2. AdaBoost

AdaBoost was proposed by Freund, Y. et al. [54] in 1995. The AdaBoost model is also an ensemble learning algorithm based on GBDT. However, unlike the XGBoost model, the AdaBoost model iteratively trains multiple weak classifiers, adjusting the weight of the samples each time. This allows the samples that were misclassified by the previous weak classifier to receive more attention in the next round of training, ultimately combining these weak classifiers into a more powerful classifier. The principle of AdaBoost model construction is illustrated in Figure 5. Initially, the sample weights are initialized, with equal weights assigned to each sample. Iterative training is then performed. For each weak classifier, the error rate on the training set is calculated, and the weight coefficients are updated. Subsequently, all weak classifiers are combined in a weighted manner to obtain the final strong classifier. Finally, new samples are classified using the integrated strong classifiers.

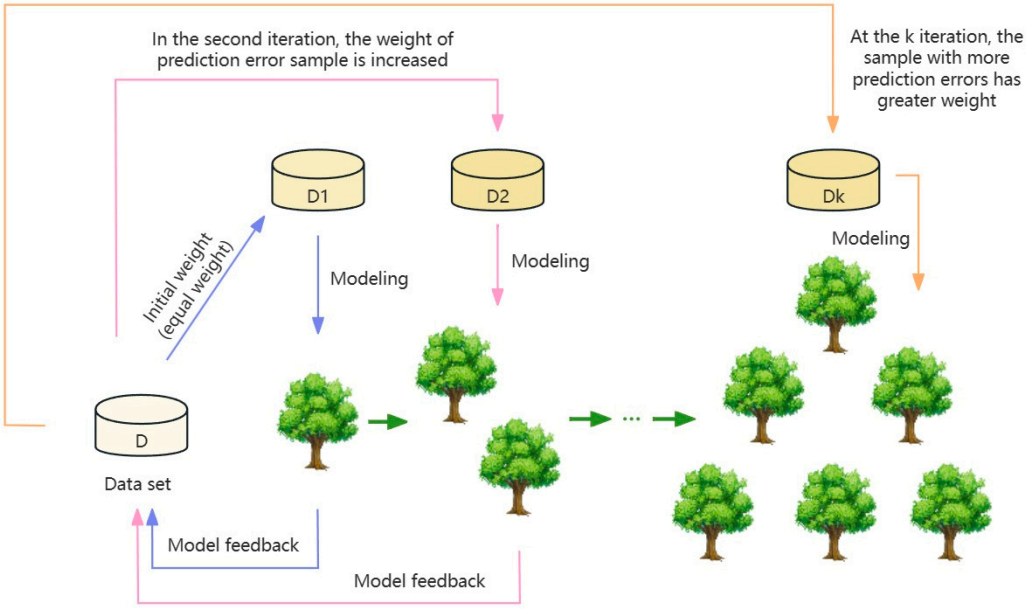

Figure 5. AdaBoost conceptual model.

3.2.3. LightGBM

LightGBM was proposed by Ke, G. et al. [55] at the end of 2016. The implementation principle of LightGBM is similar to GBDT, which uses CART to iteratively train the data in order to obtain the optimal model. The model construction principle is illustrated in Figure 6. A series of CARTs are trained sequentially, generating a weak classifier in each iteration. The next round of weak classifiers will continue to train on the residuals of the previous round of weak classifiers. Finally, the prediction results of all regression trees (additive model) will be aggregated to obtain a strong classifier. LightGBM uses a histogram algorithm to discretize the data, optimize the optimal CART segmentation points, and adopt a leaf-wise growth strategy. This approach differs from traditional level-wise growth, as it selects the leaf node that maximizes the objective function gain after each split for splitting. This strategy can reduce the loss function value more rapidly and enhance the speed of model training. Therefore, LightGBM has the characteristics of efficiency, accuracy, scalability, and reduced overfitting, making it suitable for sorting, classification, regression, and various other ML tasks.

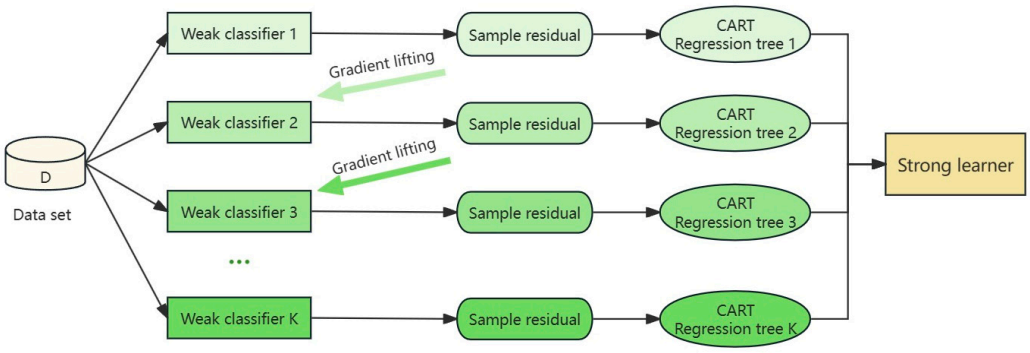

Figure 6. LightGBM conceptual model.

3.2.4. Random Forest

RF was proposed by Breiman L based on the Bagging ensemble learning theory [56]. It is essentially a classification method based on ensemble learning, which generates a single consensus prediction by combining a large number of decision trees [24]. The key is to create independent training sets, generate various classification models using different

training sample sets, and then combine these models into a random forest model to enhance the diversity among models. Finally, collective decisions are made through voting, and the mode of selection is based on the voting results as the output. The random forest model constructed in this paper is illustrated in Figure 7.

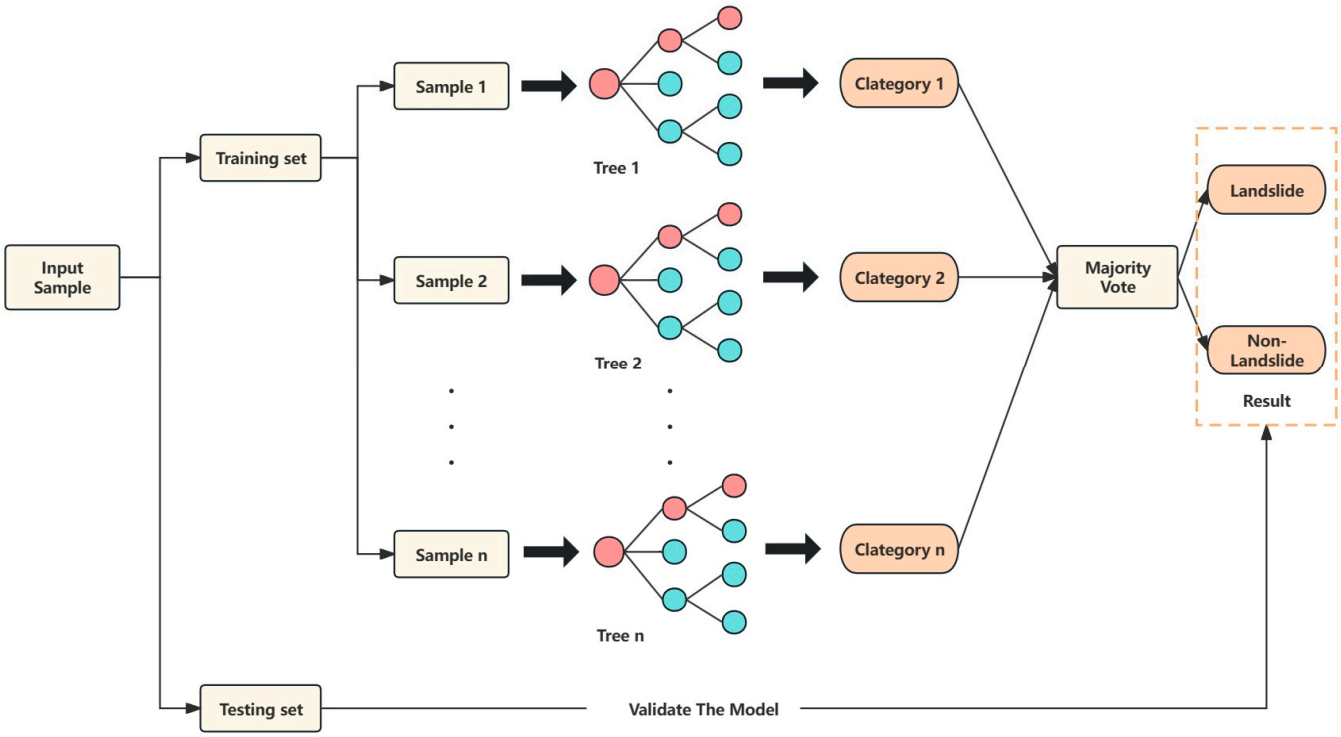

Figure 7. Flow chart of the Random Forest model.

3.2.5. CNN

CNN, as one of the most popular deep learning algorithms, has attracted much attention due to its outstanding contributions to computer vision [57]. At the same time, the classification of high-dimensional data is also a strength of CNNs. In recent years, an increasing number of scholars have applied it to research related to landslide recognition. CNN consists of four key components: convolutional layer, activation layer, pooling layer, and fully connected layer. Based on these layers, different scholars have proposed many carefully designed CNN structures to solve various practical problems [24]. The CNN structure constructed in this article is shown in Figure 8. (The main reference for model construction is [18]).

3.2.6. Model Evaluation

In this study, A, P, R, F1, and IOU are selected as the accuracy evaluation indices for landslide identification results. First of all, we need to clarify the concepts of the four types of landslide categorization prediction samples, i.e., TP: the predicted category is landslide and the prediction is consistent with the actual category; FP: the predicted category is landslide, but the prediction is inconsistent with the actual category; TN: the predicted category is non-landslide and the prediction is consistent with the actual category; and FN: the predicted category is non-landslide, but the prediction is inconsistent with the actual category [58,59].

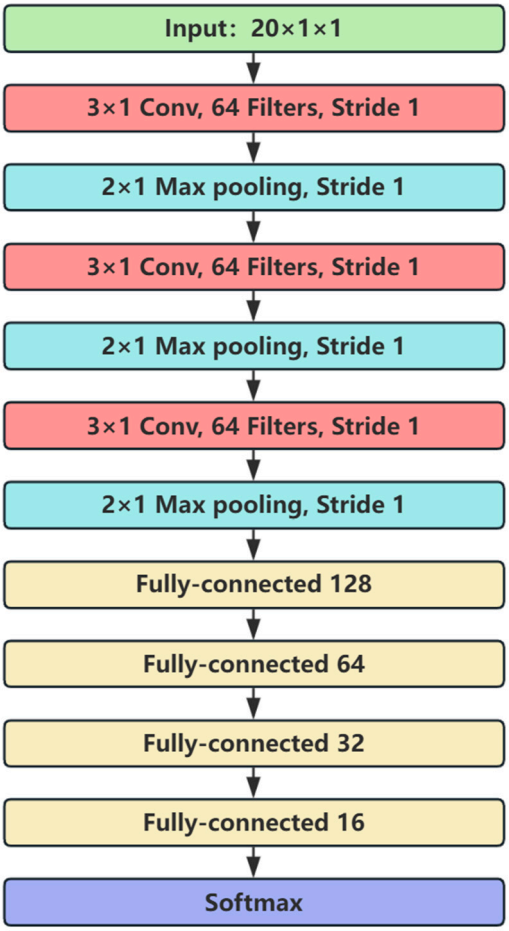

Figure 8. CNN structure diagram.

Definitions and formulas for the A, P, R, F1, and IOU indicators are as follows:

A quantified the percentage of correctly predicted samples in landslide identification:

$$A = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{2}$$

P quantified the ratio of correctly predicted landslide samples to all predicted landslide samples:

$$P = \frac{TP}{(TP + FP)} \tag{3}$$

R quantified the ratio of predicted landslide samples among all actual landslide samples:

$$R = \frac{TP}{(TP + FN)} \tag{4}$$

F1 is the harmonic mean of P and R, used to balance accuracy and recall in order to provide comprehensive performance metrics:

$$F1 = \frac{2 \times P \times R}{(P + R)} \tag{5}$$

In computer vision, especially in object detection and semantic segmentation tasks, IOU is a commonly used evaluation metric to measure the degree of overlap between model prediction results and real labels [60]. Generally speaking, any method of generating

boundary polygons can be validated by using IOU based on an accurate inventory dataset of target polygons [48]. The formula is as follows:

$$\mathrm{IOU} = \mathrm{O}/\mathrm{U} \tag{6}$$

where O denotes the Area of Overlap between actual and predicted landslide samples, and U denotes the Area of Union between actual and predicted landslide samples.

4. Results

4.1. Landslide Recognition Results Based on ML Models and Ablation Experiments

This study used a 5 m resolution grid unit as the unit for landslide identification. It first extracted data from the prepared data layer and the landslide inventory data of the Dadu River Basin. Statistically, the training area had a total of 3,260,161 pixels, of which 11.5% were interpreted as landslide pixels. The validation area comprised a total of 2,948,849 pixels, with 6.5% being identified as landslide pixels. Therefore, the total number of landslide samples in the training data was 375,882, and the total number of non-landslide samples was 2,884,279. The total number of landslide samples in the validation data was 193,124, and the total number of non-landslide samples was 2,755,725. The positive and negative samples were combined and normalized separately. Corresponding labels were added to the training data to create the final training and validation sample dataset. Further, five different ratios of the P/N samples (i.e., 1:1, 1:2, 1:3, 1:4, 1:5) were used for model training. Two training data layers were set up for each sample ratio: one with remote sensing features and terrain features and the other with only remote sensing features. By using these sample ratios, each algorithm generated 10 models, resulting in a total of 50 models being trained.

In this study, ML and deep learning algorithms were developed and trained in the Python environment. Grid search was employed to identify the optimal parameters for the algorithms. The summary of the optimal parameters for each algorithm is presented in Table 2.

Table 2. Summary of the optimal parameters of each algorithm.

Algorithm Name	Optimal Parameter
XGBoost	n_estimators: 400 Learning_rate: 0.1 max_depth: 4
AdaBoost	n_estimators: 2000 Learning_rate: 1
LightGBM	num_boost_round: 700 num_leaves: 50 learning_rate: 0.1
RF	Number of decision trees: 120
CNN	Optimizer: Adam epochs: 10 Learning_rate: 0.001 batch_size: 6000

For the landslide identification results, each model is named in the format of A_B_C, where A represents the abbreviation of the algorithm, B indicates the number of data layers used for training (14 for remote sensing features and 6 for terrain features), and C denotes the reciprocal of the ratio of positive and negative samples in the training data. For example, when training data with remote sensing features and terrain features using the CNN algorithm, with a ratio of the P/N samples of 1:5, the resulting model is referred to as CNN_20_5. In addition, selecting all landslide samples in the training area as positive samples and randomly choosing negative samples can impact the model's performance.

Therefore, this study conducted ten independent retention verifications. Under the same training settings, the final performance of this model is determined by averaging the results from ten verifications [18]. The results of landslide identification in the validation area using XGBoost, AdaBoost, LightGBM, RF, CNN, and optimal parameters are shown in Figures 9–13.

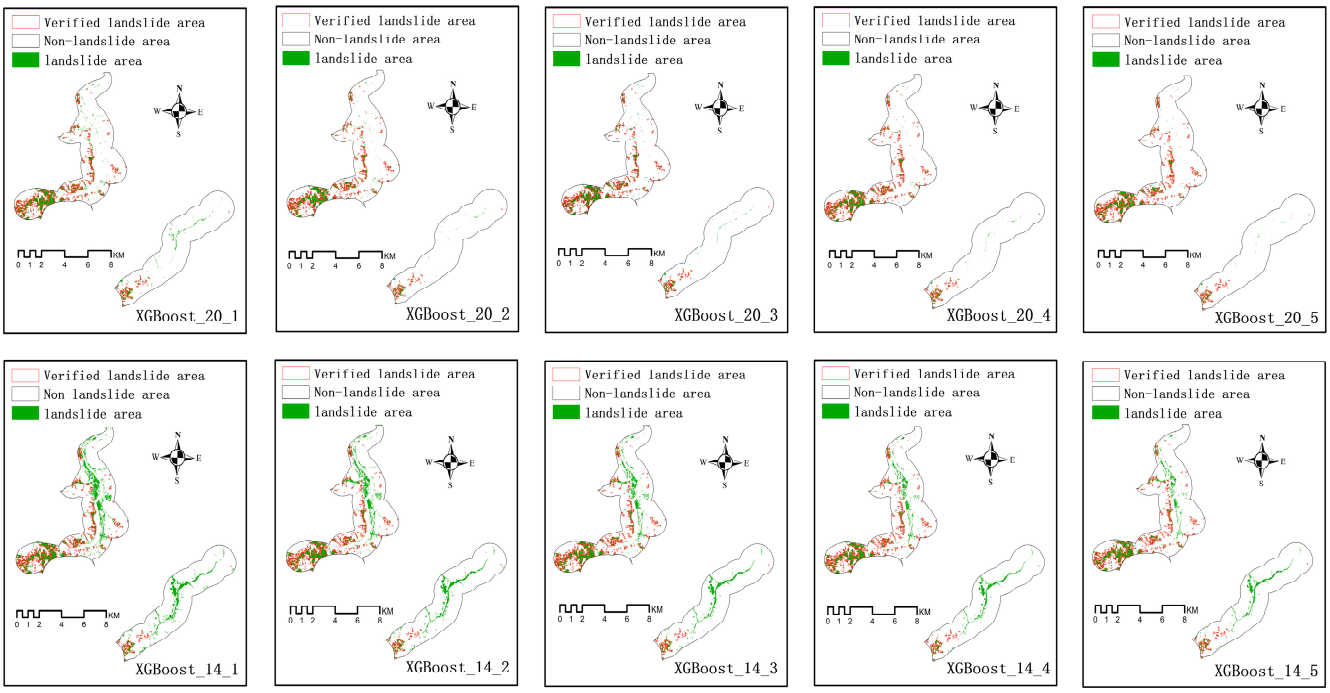

Figure 9. Landslide identification results based on XGBoost.

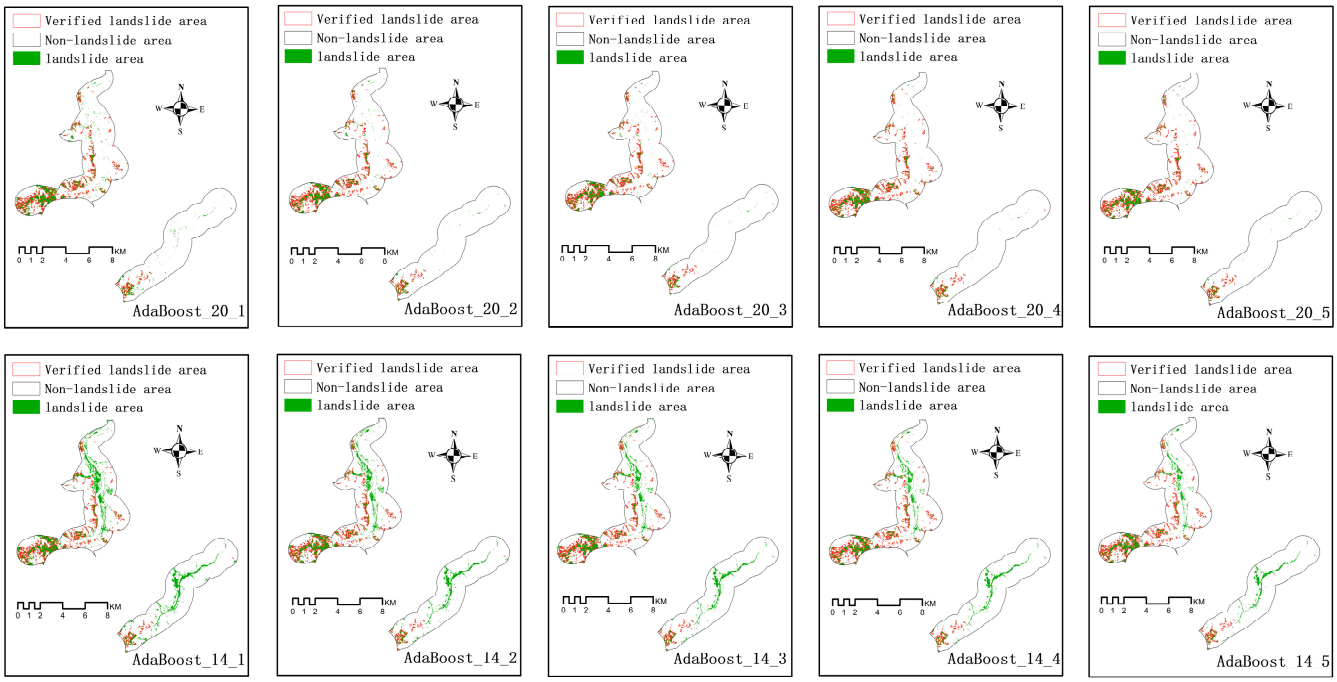

Figure 10. Landslide identification results based on AdaBoost.

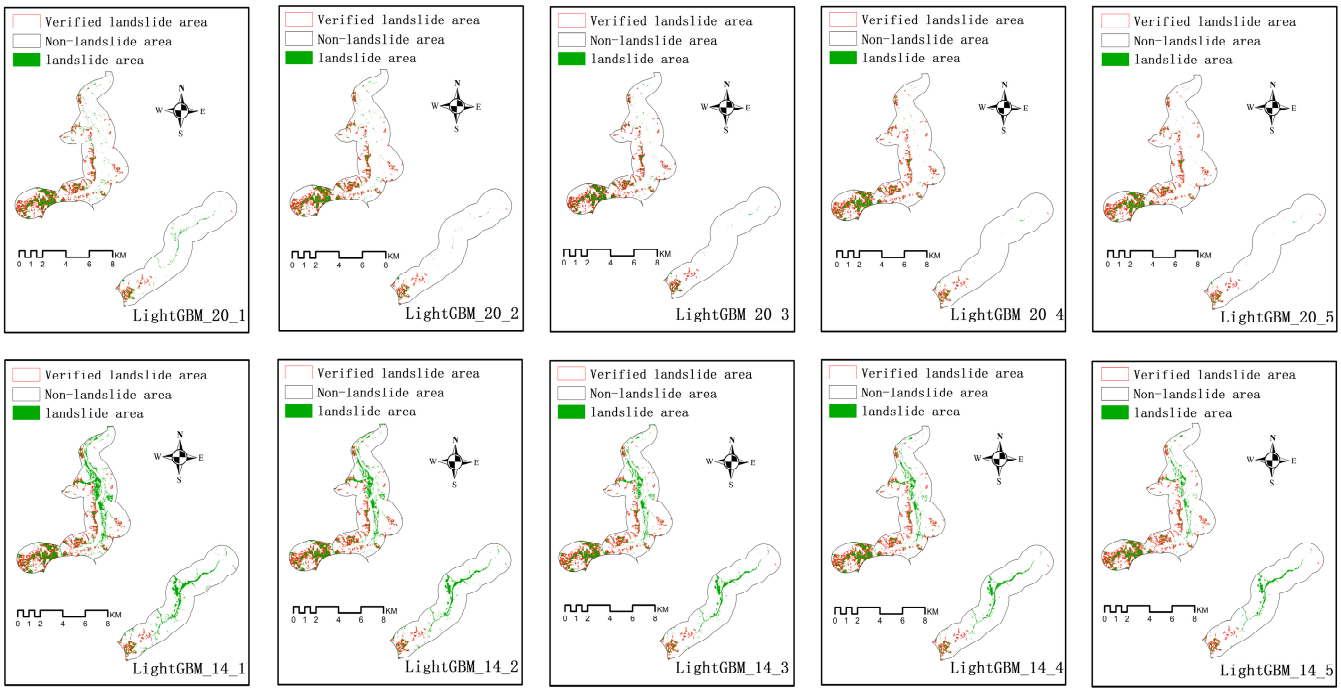

Figure 11. Landslide identification results based on LightGBM.

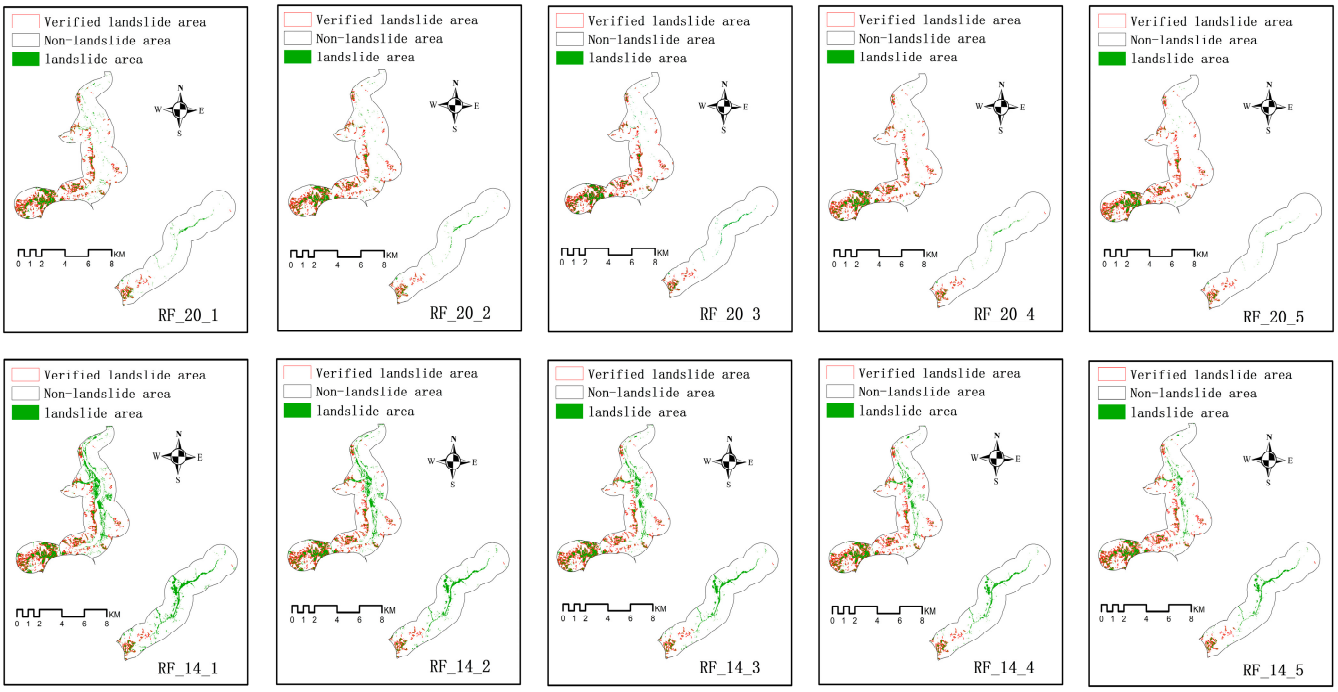

Figure 12. Landslide identification results based on RF.

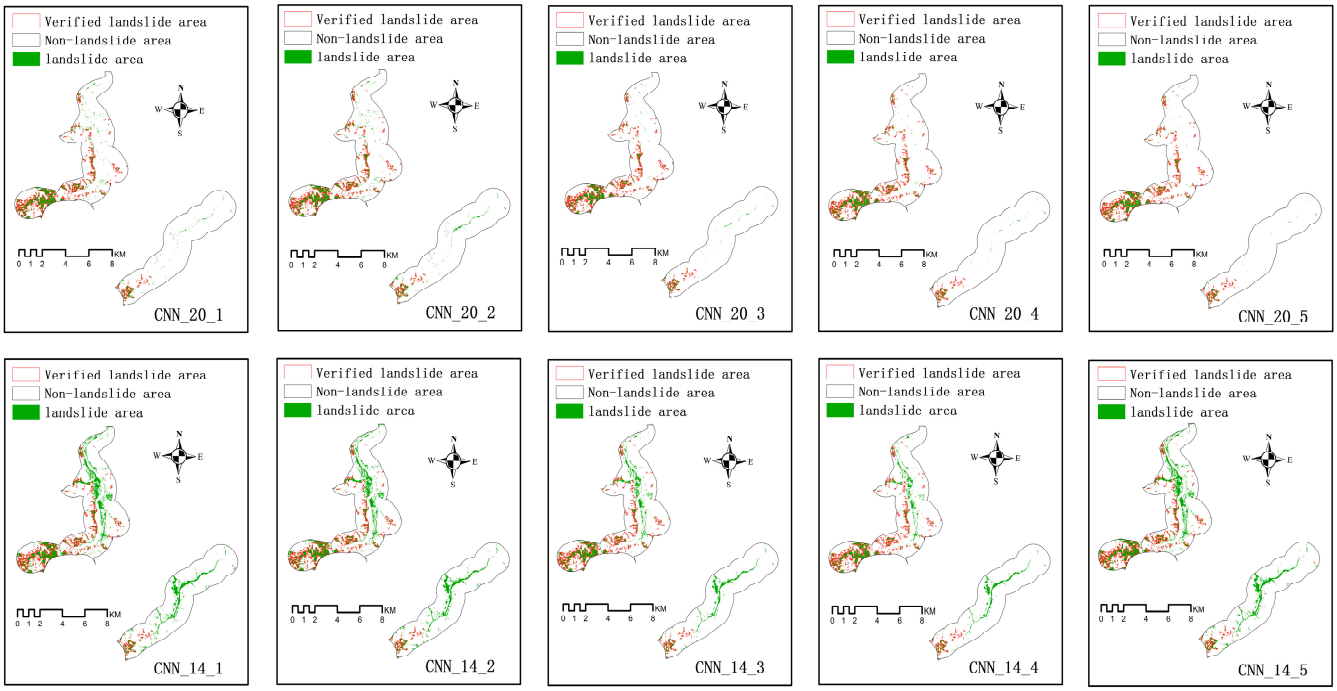

Figure 13. Landslide identification results based on CNN.

4.2. Model Performance Evaluation

The performance evaluation indicators of XGBoost, AdaBoost, LightGBM, RF, and CNN using different ratios of the P/N samples and various data layers for landslide identification in the test area are presented in Tables 3–7. A comparison line chart illustrating various performance evaluation indicators of the models is depicted in Figure 14. These performance evaluation indicators encompass A, R, P, F1, and IOU. The larger the value, the better the model performance.

Overall, when evaluating the performance of the model in terms of landslide identification results A, P, R, F1, and IOU in the validation area, the LightGBM model usually performs better, followed by the XGBoost, CNN, AdaBoost, and RF models. Among the models constructed by fully considering multiple ML methods and different P/N sample ratios while maintaining consistent input features, the best-performing model is LightGBM_20_3, with five evaluation indicators of 97.47%, 85.40%, 76.95%, 80.95%, and 71.28%, respectively. The worst-performing model is RF_20_1, with five evaluation indicators of 96.65%, 71.97%, 85.14%, 78.01%, and 64.84%, respectively. It can be understood that through comparative research, using appropriate ML methods and P/N sample ratios can increase the A, P, F1, and IOU values of landslide recognition results by 1.18%, 13.43%, 2.94%, and 6.44% compared to poorer ML methods and P/N sample ratios. Despite a decrease of 8.19% in the R metric, the overall performance of the model has still improved. When terrain feature data and remote sensing feature data are combined and input into model training, the performance of all algorithms significantly improves compared to using only remote sensing feature data. Taking the superior LightGBM_20_3 model as an example, compared to the LightGBM_14_3 model, incorporating terrain features improved the A, P, R, F1, and IOU of landslide recognition results by 4.51%, 35.66%, 5.41%, 22.27%, and 29.16%, respectively.

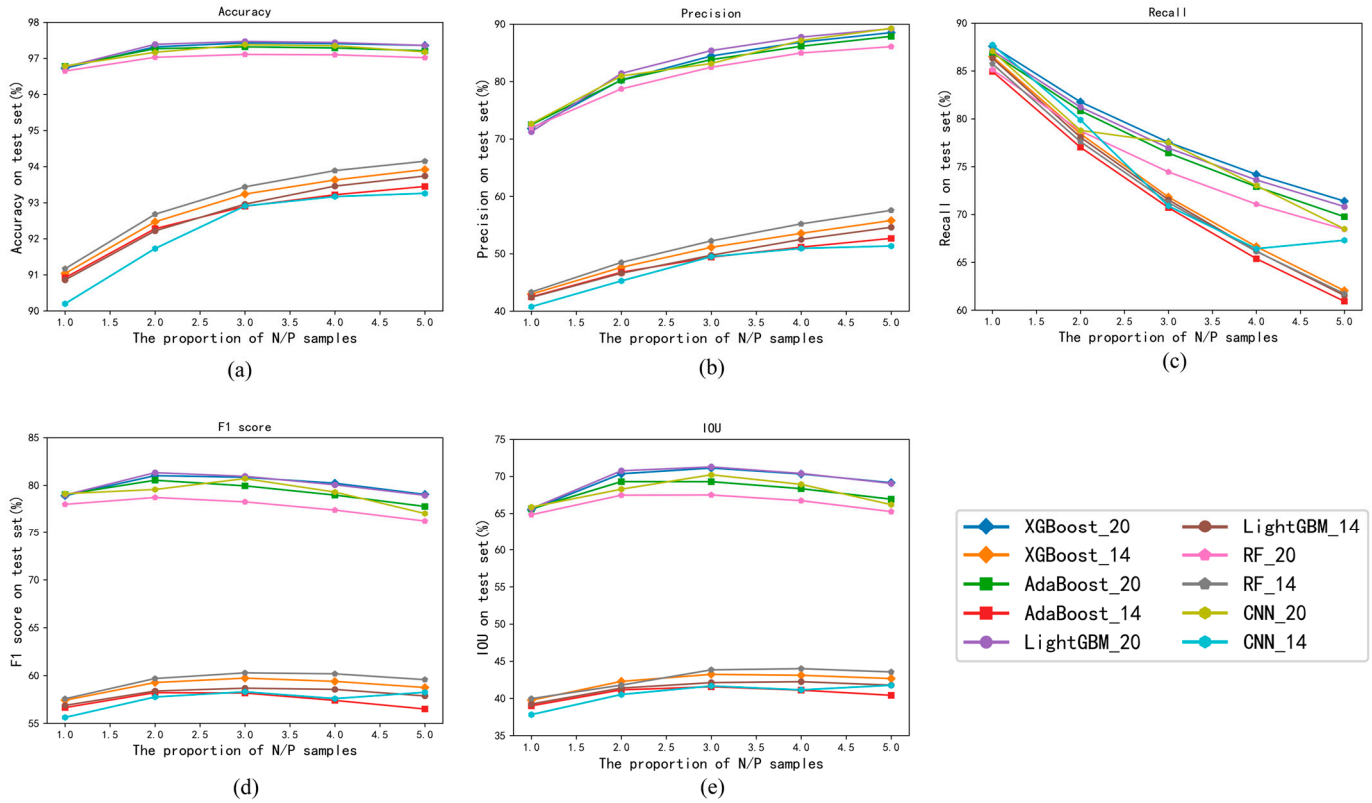

Figure 14. Performance comparison of different models: (**a**) accuracy on the test set; (**b**) precision on the test set; (**c**) recall on the test set; (**d**) F1 score on the test set; and (**e**) IOU on the test set.

Table 3. Performance evaluation indicators of XGBoost.

Model Name	A (%)	P (%)	R (%)	F1 (%)	IOU (%)
XGBoost_14_1	91.04	42.97	86.51	57.42	39.76
XGBoost_14_2	92.47	47.63	78.43	59.27	42.29
XGBoost_14_3	93.24	51.12	71.83	59.73	43.24
XGBoost_14_4	93.63	53.56	66.64	59.39	43.11
XGBoost_14_5	93.92	55.79	62.04	58.75	42.67
XGBoost_20_1	96.73	71.82	87.59	78.92	65.52
XGBoost_20_2	97.32	80.29	81.77	81.02	70.36
XGBoost_20_3	97.43	84.45	77.54	80.85	71.13
XGBoost_20_4	97.41	86.87	74.19	80.03	70.33
XGBoost_20_5	97.36	88.53	71.40	79.05	69.15

Table 4. Performance evaluation indicators of AdaBoost.

Model Name	A (%)	P (%)	R (%)	F1 (%)	IOU (%)
AdaBoost_14_1	90.92	42.49	84.95	56.65	39.01
AdaBoost_14_2	92.28	46.78	77.03	58.21	41.15
AdaBoost_14_3	92.90	49.40	70.72	58.17	41.57
AdaBoost_14_4	93.22	51.16	65.38	57.40	41.11
AdaBoost_14_5	93.45	52.68	60.93	56.50	40.42
AdaBoost_20_1	96.78	72.49	86.89	79.04	65.62
AdaBoost_20_2	97.27	80.25	80.84	80.54	69.30
AdaBoost_20_3	97.32	83.82	76.42	79.95	69.29
AdaBoost_20_4	97.29	86.15	72.91	78.98	68.35
AdaBoost_20_5	97.21	87.88	69.77	77.79	66.93

Table 5. Performance evaluation indicators of LightGBM.

Model Name	A (%)	P (%)	R (%)	F1 (%)	IOU (%)
LightGBM_14_1	90.86	42.41	86.33	56.88	39.25
LightGBM_14_2	92.22	46.61	78.08	58.38	41.39
LightGBM_14_3	92.96	49.74	71.54	58.68	42.12
LightGBM_14_4	93.46	52.50	66.17	58.55	42.25
LightGBM_14_5	93.74	54.61	61.54	57.87	41.78
LightGBM_20_1	96.76	72.25	87.13	78.99	65.66
LightGBM_20_2	97.39	81.42	81.23	81.33	70.76
LightGBM_20_3	97.47	85.40	76.95	80.95	71.28
LightGBM_20_4	97.44	87.75	73.61	80.06	70.40
LightGBM_20_5	97.36	89.21	70.83	78.96	69.06

Table 6. Performance evaluation indicators of RF.

Model Name	A (%)	P (%)	R (%)	F1 (%)	IOU (%)
RF_14_1	91.17	43.33	85.74	57.57	39.97
RF_14_2	92.68	48.49	77.61	59.69	41.79
RF_14_3	93.44	52.24	71.25	60.28	43.84
RF_14_4	93.89	55.21	66.14	60.18	44.00
RF_14_5	94.15	57.56	61.72	59.57	43.56
RF_20_1	96.65	71.97	85.14	78.01	64.84
RF_20_2	97.03	78.72	78.74	78.73	67.47
RF_20_3	97.11	82.50	74.45	78.27	67.50
RF_20_4	97.10	84.98	71.08	77.41	66.73
RF_20_5	97.02	86.08	68.45	76.26	65.26

Table 7. Performance evaluation indicators of CNN.

Model Name	A (%)	P (%)	R (%)	F1 (%)	IOU (%)
CNN_14_1	90.20	40.78	87.73	55.62	37.82
CNN_14_2	91.83	45.27	79.89	57.76	40.53
CNN_14_3	92.91	49.50	70.98	58.30	41.67
CNN_14_4	93.17	50.94	66.43	57.59	41.15
CNN_14_5	93.26	51.35	67.31	58.25	41.77
CNN_20_1	96.79	72.65	87.07	79.13	65.86
CNN_20_2	97.17	80.97	78.80	79.57	68.28
CNN_20_3	97.38	83.15	77.53	80.71	70.22
CNN_20_4	97.35	87.15	73.02	79.30	68.93
CNN_20_5	97.18	89.28	68.50	77.05	66.20

5. Discussion

5.1. The Effectiveness of Comparing Different ML Methods

Compared with traditional landslide identification methods, the utilization of ML algorithms significantly enhances the accuracy of the results [22,23]. Nevertheless, there is still debate among researchers regarding the selection and application of these algorithms in various study areas [61]. This study employed five commonly used and high-performance ML methods for landslide identification research. The landslide identification results of the validation area showed that all five models could obtain accurate and comprehensive results. Among them, LightGBM was identified as the most suitable method for the study area. Although the RF method has demonstrated superior performance in some studies [35,62], it performed poorly in this study. Compared to the worst-performing RF method, the LightGBM method improved the A, P, R, F1, and IOU of landslide recognition results by 0.36%, 2.9%, 2.5%, 2.68%, and 3.78%, respectively, while maintaining consistent training data and P/N sample ratios. Therefore, when conducting landslide identification, utilizing multiple ML methods for comparative research can help determine the applicabil-

ity of different methods in the study area, leading to more accurate and reliable landslide identification results.

5.2. The Influence of Different P/N Sample Ratios on Landslide Identification Results

In the study, different P/N sample ratios have a significant impact on landslide recognition. As the proportion of negative samples in the training data increases, the accuracy of landslide recognition results obtained by all models will increase, while the recall rate will decrease. The other three indicators, namely A, F1, and IOU, show a trend of first increasing and then decreasing. When the ratio of P/N samples in the training data is 1:3, the overall recognition performance of the model is usually optimal. In the lightGBM model with the best performance, the landslide recognition results of its optimal (1:3) and poor (1:1) P/N sample ratios are compared. The results show that although the optimal P/N sample ratio reduces the R value of the landslide recognition results by 10.18%, it improves the A, P, F1, and IOU indicators by 0.71%, 13.15%, 1.96%, and 5.62%, respectively. Overall, the identified landslides are more accurate and reasonable. It is worth noting that in some regions, landslide identification results with a P/N sample ratio of 1:4 and 1:5 can effectively avoid the "speckled image" (see Figure 15).

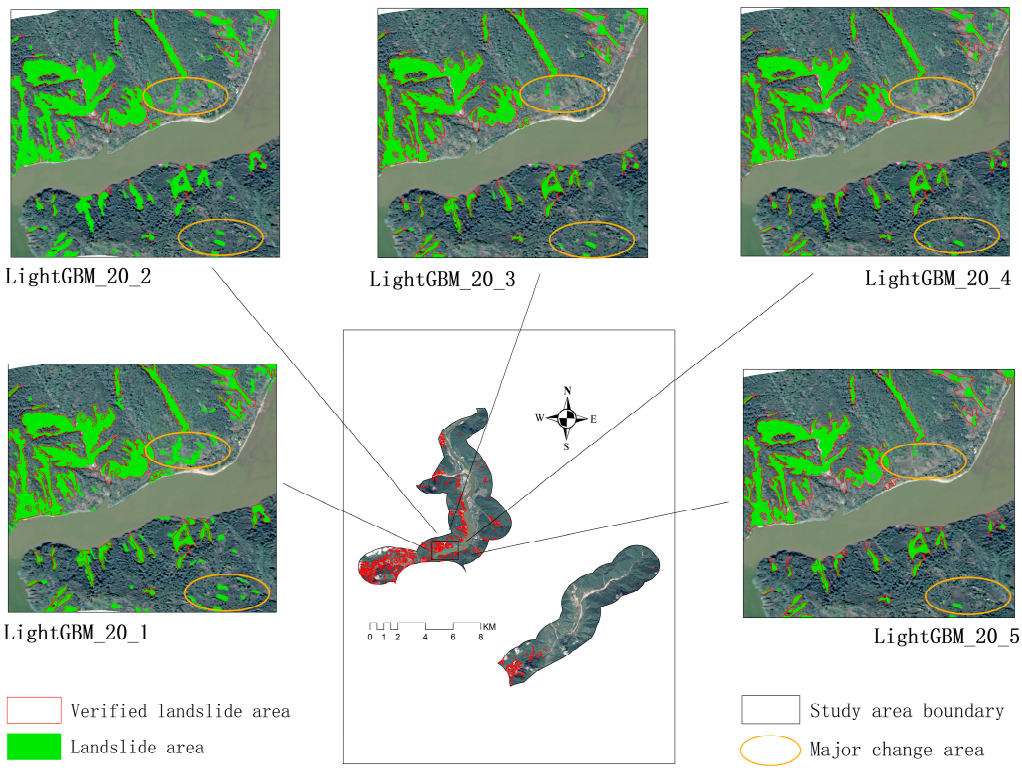

Figure 15. Landslide identification effect with different ratios of P/N samples.

Therefore, we believe that appropriately increasing the proportion of negative samples in the selection of P/N sample ratios for landslide identification can effectively improve the effectiveness of landslide identification. However, exceeding a certain threshold can lead to more landslide samples being incorrectly predicted as non-landslides, resulting in a decrease in the overall recognition accuracy of the model. In practical research, different P/N sample ratios should be used for model training and prediction based on the specific requirements and drawing needs of the study area in order to obtain more suitable P/N sample ratios and better landslide identification results, rather than following a 1:1 or other predefined fixed ratios.

5.3. The Influence of Terrain Features on Landslide Identification Results

To verify whether the performance of the model improves after fusing terrain features, we created two different training datasets for landslide identification studies: one dataset contains solely remotely sensed features, while the other dataset includes both remotely sensed features and terrain features.

From the evaluation indicators of landslide recognition results, while maintaining the consistency of machine learning methods and P/N sample ratios, the A, P, R, F1, and IOU of landslide recognition results obtained by fusing terrain features have significantly improved. The average improvement of the five indicators (the average difference between XX_20_X and XX_14_X) is 4.53%, 32.77%, 4.29%, 21.66%, and 26.63%, respectively. From the landslide recognition result map (please refer to the example of using the LightGBM model in Figure 16), both models can effectively extract landslide areas. However, due to the similarity between remote sensing features of buildings and roads and landslide areas, models relying solely on remote sensing features may mistakenly classify buildings and roads as landslide areas to a certain extent, and there is also a serious noise phenomenon. The model that incorporates fusion terrain features has a strong recognition ability for non-landslide areas. It can effectively distinguish whether buildings and roads belong to non-landslide areas.

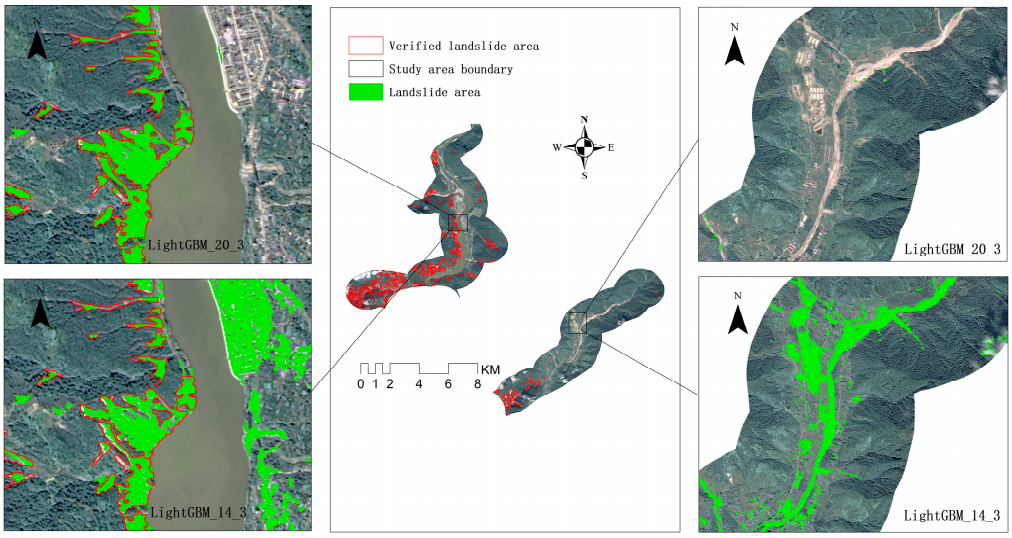

Figure 16. Landslide recognition effect of different training data.

Therefore, we can conclude that in this study, terrain features can significantly enhance the effectiveness of landslide identification, improve the model's ability to distinguish non-landslide areas like buildings and roads, and consequently, more accurately and comprehensively identify landslide areas.

5.4. Comparative Analysis of Similar Studies

To further validate the effectiveness of the methods used in the study for landslide identification, we compared them with similar literature on landslide identification. We selected the A, P, and IOU indicators from the optimal identification results for comparison, as shown in Table 8.

The comparison of the results of four similar studies shows that this study ranks first in terms of A (1/3, indicating first place out of three), second in terms of P (2/4), and third in terms of IOU (3/3). This study outperformed others in terms of A and P and lagged behind in IOU by only 0.42% compared to the second place. It is important to note that the identification of landslides may still be influenced by the climate conditions and geographical location of the study area [34]. This will introduce certain uncertainties to the comparison of similar studies. However, through a comprehensive evaluation of

the accuracy of the results of this study, the mapping effect, and comparison with similar studies, our proposed landslide identification method is effective and can ensure the accuracy and comprehensiveness of the identification results.

Table 8. Comparison of landslide identification results in similar studies.

Author	Study Area	A (%)	P (%)	IOU (%)
The author of this study	Detuo Township, Sichuan Province, China	97.47 [1]	85.40 [2]	71.28 [3]
Ghorbanzadeh, O. et al. (2019) [48]	Rasuwa Region, Nepal	/	83.31 [3]	78.26 [1]
Wang, H. et al. (2021) [24]	Lantau Island, Hong Kong, China	89.32 [2]	92.58 [1]	/
Cai, J. et al. (2022) [63]	Hongkou Township, Sichuan Province, China	/	71.66 [4]	71.70 [2]
Asadi, A. et al. (2024) [19]	Kyushu Island, Japan	83.6 [3]	/	/

Note: The numerical superscript represents the ranking in this indicator, "/" indicates that the evaluation indicator was not used in the study.

6. Conclusions

This study utilizes five ML methods, XGBoost, AdaBoost, LightGBM, RF, and CNN, to detect landslides at the pixel level. Through comparative analysis, the most suitable ML method and P/N sample ratio for the study area are determined, resulting in more precise and comprehensive landslide recognition outcomes. Additionally, the study investigates the impact of terrain features on landslide recognition through ablation experiments. Subsequently, a comparative analysis of similar studies was conducted to validate the effectiveness of the proposed method. The research results led to the following conclusions:

(1) By considering the fusion of various ML methods, different P/N sample ratios, and terrain features, we have obtained the most suitable landslide recognition model for the study area. The A, P, R, F1, and IOU of the landslide recognition results obtained by this model reached 97.47%, 85.40%, 76.95%, 80.95%, and 71.28%, respectively. Compared with similar studies, it is concluded that the landslide recognition method proposed in this study is effective and can ensure the accuracy and comprehensiveness of the recognition results. Simultaneously, it can provide a reference basis for future scholars in landslide recognition research regarding methodology and feature construction.

(2) Choosing appropriate ML methods and P/N sample ratios can improve the overall performance of landslide recognition. Compared to the RF method with the worst performance, the LightGBM method with the best performance in the study increased the A, P, R, F1, and IOU of landslide identification results by 0.36%, 2.9%, 2.5%, 2.68%, and 3.78%, respectively. While maintaining the consistency of the ML method and the input data, the appropriate P/N sample ratio also enhances the accuracy and validity of landslide identification results. Taking LightGBM as an example, compared to the poor P/N sample ratio, the optimal P/N sample ratio increases by 0.71%, 13.15%, 1.96%, and 5.62% in A, P, F1, and IOU indicators, respectively.

(3) Compared with models trained using only remote sensing feature data, models trained using both remote sensing feature data and terrain feature data can significantly improve the ability to distinguish between landslide and non-landslide areas. Taking the superior LightGBM model as an example, incorporating terrain features increases the A, P, R, F1, and IOU of landslide recognition results by 4.51%, 35.66%, 5.41%, 22.27%, and 29.16%, respectively. This enhancement leads to higher accuracy and more comprehensive landslide recognition results.

Author Contributions: Jincan Wang and Zhiheng Wang designed the research; Chenzhihao Qian and Liyao Peng processed the data; Jincan Wang analyzed the data; and Jincan Wang and Zhiheng Wang wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 41971310, 42271103.

Data Availability Statement: The authors have not obtained permission to publish the data. Therefore, the data can be obtained from the corresponding author upon reasonable request.

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

References

1. Wang, Y.; Wu, X.; Chen, Z.; Ren, F.; Feng, L.; Du, Q. Optimizing the Predictive Ability of Machine Learning Methods for Landslide Susceptibility Mapping Using SMOTE for Lishui City in Zhejiang Province, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 368. [CrossRef]
2. Zhou, X.; Wen, H.; Zhang, Y.; Xu, J.; Zhang, W. Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. *Geosci. Front.* **2021**, *12*, 101211. [CrossRef]
3. Wang, Z.; Wang, D.; Guo, Q.; Wang, D. Regional landslide hazard assessment through integrating susceptibility index and rainfall process. *Nat. Hazards* **2020**, *104*, 2153–2173. [CrossRef]
4. Darrow, M.; Nelson, V.; Grilliot, M.; Wartman, J.; Jacobs, A.; Baichtal, J.; Buxton, C. Geomorphology and initiation mechanisms of the 2020 Haines, Alaska landslide. *Landslides* **2022**, *19*, 2177–2188. [CrossRef]
5. Koarai, M.; Sato, H.P.; Une, H.; Kamiya, I. Interpretation of high-resolution satellite images to detect the landform changes and disaster damages: Case study of the northern Pakistan earthquake. In Proceedings of the SPIE Asia-Pacific Remote Sensing, Goa, India, 12 December 2006. [CrossRef]
6. Danneels, G.; Pirard, E.; Havenith, H. Automatic landslide detection from remote sensing images using supervised classification methods. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 3014–3017.
7. Chen, T.; Trinder, J.C.; Niu, R. Object-Oriented Landslide Mapping Using ZY-3 Satellite Imagery, Random Forest and Mathematical Morphology, for the Three-Gorges Reservoir, China. *Remote Sens.* **2017**, *9*, 333. [CrossRef]
8. Song, Y.; Hao, L.; Yan, L.; Wang, Y.; Chang, H. Application of Support Vector Machine in Landslide Identification, China. *J. Lanzhou Univ. (Nat. Sci. Ed.)* **2022**, *58*, 727–734. [CrossRef]
9. SS, V.C.; Shaji, E. Landslide identification using machine learning techniques: Review, motivation, and future prospects. *Earth Sci. Inform.* **2022**, *15*, 2063–2090. [CrossRef]
10. Dias, H.C.; Hölbling, D.; Grohmann, C.H. Rainfall-Induced Shallow Landslide Recognition and Transferability Using Object-Based Image Analysis in Brazil. *Remote Sens.* **2023**, *15*, 5137. [CrossRef]
11. Naidu, S.; Sajinkumar, K.S.; Oommen, T.; Anuja, V.J.; Samuel, R.A.; Muraleedharan, C. Early warning system for shallow landslides using rainfall threshold and slope stability analysis. *Geosci. Front.* **2018**, *9*, 1871–1882. [CrossRef]
12. Yin, W.; Niu, C.; Bai, Y.; Zhang, L.; Ma, D.; Zhang, S.; Zhou, X.; Xue, Y. An Adaptive Identification Method for Potential Landslide Hazards Based on Multisource Data. *Remote Sens.* **2023**, *15*, 1865. [CrossRef]
13. Guzzetti, F.; Mondini, A.; Cardinali, M.; Fiorucci, F.; Santangelo, M.; Chang, K. Landslide inventory maps: New tools for an old problem. *Earth-Sci. Rev.* **2012**, *112*, 42–66. [CrossRef]
14. Ghorbanzadeh, O.; Shahabi, H.; Crivellari, A.; Homayouni, S.; Blaschke, T.; Ghamisi, P. Landslide detection using deep learning and object-based image analysis. *Landslides* **2022**, *19*, 929–939. [CrossRef]
15. Xu, C. Preparation of earthquake-triggered landslide inventory maps using remote sensing and GIS technologies: Principles and case studies. *Geosci. Front.* **2015**, *6*, 825–836. [CrossRef]
16. Lu, Z.; Peng, Y.; Li, W.; Yu, J.; Ge, D.; Han, L.; Xiang, W. An Iterative Classification and Semantic Segmentation Network for Old Landslide Detection Using High-Resolution Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4408813. [CrossRef]
17. Galli, M.; Ardizzone, F.; Cardinali, M.; Guzzetti, F.; Reichenbach, P. Comparing landslide inventory maps. *Geomorphology* **2008**, *94*, 268–289. [CrossRef]
18. Wang, H.; Zhang, L.; Wang, L.; Fan, R.; Zhou, S.; Qiang, Y.; Peng, M. Machine learning powered high-resolution co-seismic landslide detection. *Gondwana Res.* **2023**, *123*, 217–237. [CrossRef]
19. Asadi, A.; Baise, L.G.; Koch, M.; Moaveni, B.; Chatterjee, S.; Aimaiti, Y. Pixel-based classification method for earthquake-induced landslide mapping using remotely sensed imagery, geospatial data and temporal change information. *Nat. Hazards* **2024**, *120*, 5163–5200. [CrossRef]
20. Kavzoglu, T.; Colkesen, I.; Sahin, E.K. Machine Learning Techniques in Landslide Susceptibility Mapping: A Survey and a Case Study. In *Landslides: Theory, Practice and Modelling*; Pradhan, S., Vishal, V., Singh, T., Eds.; Advances in Natural and Technological Hazards Research; Springer: Cham, Switzerland, 2019; Volume 50. [CrossRef]
21. Alcántara-Ayala, I.; Sassa, K. Contribution of the International Consortium on Landslides to the implementation of the Sendai Framework for Disaster Risk Reduction: Engraining to the Science and Technology Roadmap. *Landslides* **2021**, *18*, 21–29. [CrossRef]
22. Merghadi, A.; Yunus, A.P.; Dou, J.; Whiteley, J.; ThaiPham, B.; Bui, D.T.; Avtar, R.; Abderrahmane, B. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth Sci. Rev.* **2020**, *7*, 10325. [CrossRef]

23. Ju, Y.; Xu, Q.; Jin, S.; Li, W.; Su, Y.; Dong, X.; Guo, Q. Loess Landslide Detection Using Object Detection Algorithms in Northwest China. *Remote Sens.* **2022**, *14*, 1182. [CrossRef]
24. Wang, H.; Zhang, L.; Yin, K.; Luo, H.; Li, J. Landslide identification using machine learning. *Geosci. Front.* **2021**, *12*, 351–364. [CrossRef]
25. Gorsevski, P.; Brown, M.; Panter, K.; Onasch, C.M.; Simic, A.; Snyder, J. Landslide Detection and Susceptibility Mapping Using LiDAR and An Artificial Neural Network Approach: A Case Study in the Cuyahoga Valley National Park, Ohio. *Landslides* **2016**, *13*, 467–484. [CrossRef]
26. Ramos-Bernal, R.N.; Vázquez-Jiménez, R.; Cantú-Ramírez, C.A.; Alarcón-Paredes, A.; Alonso-Silverio, G.A.; Bruzón, A.G.; Arrogante-Funes, F.; Martín-González, F.; Novillo, C.J.; Arrogante-Funes, P. Evaluation of conditioning factors of slope instability and continuous change maps in the generation of landslide inventory maps using machine learning (ml) algorithms. *Remote Sens.* **2021**, *13*, 4515. [CrossRef]
27. Mezaal, M.R.; Pradhan, B.; Rizeei, H.M. Improving landslide detection from airborne laser scanning data using optimized dempster–shafer. *Remote Sens.* **2018**, *10*, 1029. [CrossRef]
28. Chen, S.; Zhou, R. Landslide detection based on color feature model and svm in remote sensing imagery. *Spacecr. Recovery Remote Sens.* **2020**, *40*, 89–98.
29. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Alizadeh, M.; Chen, W.; Mohammadi, A.; Ahmad, B.B.; Panahi, M.; Hong, H.; et al. Landslide detection and susceptibility mapping by airsar data using support vector machine and index of entropy models in cameron highlands, malaysia. *Remote Sens.* **2018**, *10*, 1527. [CrossRef]
30. Li, X.; Cheng, X.; Chen, W.; Chen, G.; Liu, S. Identification of forested landslides using lidar data, object-based image analysis, and machine learning algorithms. *Remote Sens.* **2015**, *7*, 9705–9726. [CrossRef]
31. Chen, X.; Liu, M.; Li, D.; Jia, J.; Yang, A.; Zheng, W.; Yin, L. Conv-trans dual network for landslide detection of multi-channel optical remote sensing images. *Front. Earth Sci.* **2023**, *11*, 1182145. [CrossRef]
32. Sameen, M.I.; Pradhan, B. Landslide detection using residual networks and the fusion of spectral and topographic information. *IEEE Access* **2019**, *7*, 114363–114373. [CrossRef]
33. Ji, S.; Yu, D.; Shen, C.; Li, W.; Xu, Q. Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks. *Landslides* **2020**, *17*, 1337–1352. [CrossRef]
34. Sun, D.; Gu, Q.; Wen, H.; Xu, J.; Zhang, Y.; Shi, S.; Xue, M.; Zhou, X. Assessment of landslide susceptibility along mountain highways based on different machine learning algorithms and mapping units by hybrid factors screening and sample optimization. *Gondwana Res.* **2023**, *123*, 89–106. [CrossRef]
35. Singh, P.; Maurya, V.; Dwivedi, R. Pixel Based Landslide Identification Using Landsat 8 and GEE. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 8444–8447.
36. Hong, H. Assessing landslide susceptibility based on hybrid Best-first decision tree with ensemble learning model. *Ecol. Indic.* **2023**, *147*, 109968. [CrossRef]
37. Zeng, T.; Wu, L.; Peduto, D.; Glade, T.; Hayakawa, Y.S.; Yin, K. Ensemble learning framework for landslide susceptibility mapping: Different basic classifier and ensemble strategy. *Geosci. Front.* **2023**, *14*, 101645. [CrossRef]
38. Yang, C.; Liu, L.-L.; Huang, F.; Huang, L.; Wang, X.-M. Machine learning-based landslide susceptibility assessment with optimized ratio of landslide to non-landslide samples. *Gondwana Res.* **2023**, *123*, 198–216. [CrossRef]
39. Chang, Z.; Huang, J.; Huang, F.; Bhuyan, K.; Meena, S.R.; Catani, F. Uncertainty analysis of non-landslide sample selection in landslide susceptibility prediction using slope unit-based machine learning models. *Gondwana Res.* **2023**, *117*, 307–320. [CrossRef]
40. Liu, Q.; Tang, A.; Huang, D. Exploring the uncertainty of landslide susceptibility assessment caused by the number of non–landslides. *Catena* **2023**, *227*, 107109. [CrossRef]
41. Cruden, D.; Varnes, D. Landslides: Investigation and mitigation. Chapter 3: Landslide types and processes. *Transp. Res. Board Spec. Rep.* **1966**, *247*, 36–75.
42. Imanian, A.; Tangestani, M.H.; Asadi, A. Application of radar and optical satellite imagery data in landslide potential mapping of sheshpeer sub-catchment in iran. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 547–552. [CrossRef]
43. Li, Z.; Shi, W.; Lu, P.; Yan, L.; Wang, Q.; Miao, Z. Landslide mapping from aerial photographs using change detection-based Markov random field. *Remote Sens. Environ.* **2016**, *187*, 76–90. [CrossRef]
44. Li, S.; Li, Y.; An, Y. Automatic identification of landslide disasters based on change detection, China. *Remote Sens. Infor-Mation* **2010**, *1*, 27–31.
45. Han, Y.; Wang, P.; Zheng, Y.; Yasir, M.; Xu, C.; Nazir, S.; Hossain, M.S.; Ullah, S.; Khan, S. Extraction of Landslide Information Based on Object-Oriented Approach and Cause Analysis in Shuicheng, China. *Remote Sens.* **2022**, *14*, 502. [CrossRef]
46. Yang, S.; Wang, Y.; Wang, P.; Mu, J.; Jiao, S.; Zhao, X.; Wang, Z.; Wang, K.; Zhu, Y. Automatic Identification of Landslides Based on Deep Learning. *Appl. Sci.* **2022**, *12*, 8153. [CrossRef]
47. Piralilou, S.T.; Shahabi, H.; Jarihani, B.; Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.R.; Aryal, J. Landslide Detection Using Multi-Scale Image Segmentation and Different Machine Learning Models in the Higher Himalayas. *Remote Sens.* **2019**, *11*, 2575. [CrossRef]
48. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.R.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sens.* **2019**, *11*, 196. [CrossRef]

49. Deng, X. Research on Geological Hazard Risk Assessment of the Dadu River Detuo and Jiajun River Section. Master's Thesis, Chengdu University of Technology, Chengdu, China, 2011.
50. Zhuo, L.; Huang, Y.; Zheng, J.; Cao, J.; Guo, D. Landslide Susceptibility Mapping in Guangdong Province, China, Using Random Forest Model and Considering Sample Type and Balance. *Sustainability* **2023**, *15*, 9024. [CrossRef]
51. Moosavi, V.; Niazi, Y. Development of Hybrid Wavelet PacketStatistical Models (WP-SM) for Landslide Susceptibility Mapping. Lanslides. *Landslides* **2016**, *13*, 97–114. [CrossRef]
52. Tehrani, F.S.; Calvello, M.; Liu, Z.; Zhang, L.; Lacasse, S. Machine learning and landslide studies: Recent advances and applications. *Nat. Hazards* **2022**, *114*, 1197–1245. [CrossRef]
53. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.
54. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of online learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
55. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T. LightGBM: A Highly Efficient Gradient Boosting De-cision Tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3149–3157.
56. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
57. Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning. *Genet. Program. Evolvable Mach.* **2017**, *19*, 305–307.
58. Hong, H. Assessing landslide susceptibility based on hybrid multilayer perceptron with ensemble learning. *Bull. Eng. Geol. Environ.* **2023**, *82*, 382. [CrossRef]
59. An, X.; Mi, C.; Sun, D.; Wen, H. Comparative Study on Landslide Susceptibility in the Three Gorges Reservoir Area Based on Different Evaluation Units: Take Yunyang County in Chongqing as an Example. *J. Jilin Univ. (Earth Sci. Ed.)* **2023**, *53*, 1–11. [CrossRef]
60. Biswas, M.; Pramanik, R.; Sen, S.; Sinitca, A.; Kaplun, D.; Sarkar, R. Microstructural segmentation using a union of attention guided U-Net models with different color transformed images. *Sci. Rep.* **2023**, *13*, 5737. [CrossRef] [PubMed]
61. Yu, H.; Pei, W.; Zhang, J.; Chen, G. Landslide Susceptibility Mapping and Driving Mechanisms in a Vulnerable Region Based on Multiple Machine Learning Models. *Remote Sens.* **2023**, *15*, 1886. [CrossRef]
62. Ng, C.W.W.; Yang, B.; Liu, Z.Q.; Kwan, J.S.H.; Chen, L. Spatiotemporal modelling of rainfall-induced landslides using machine learning. *Landslides* **2021**, *18*, 2499–2514. [CrossRef]
63. Cai, H.; Han, H.; Zhang, Y.; Wang, L. Convolutional neural network landslide identification based on topographic feature fusion. *J. Earth Sci. Environ.* **2022**, *44*, 568–579.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.