


Article

Old Landslide Detection Using Optical Remote Sensing Images Based on Improved YOLOv8

Yunlong Li ¹, Mingtao Ding ^{1,2,*}, Qian Zhang ¹, Zhihui Luo ¹, Wubiao Huang ³ , Cancan Zhang ¹ and Hui Jiang ¹

¹ College of Geological Engineering and Geomatics, Chang'an University, Xi'an 710054, China; 2022126016@chd.edu.cn (Y.L.); 20221226022@chd.edu.cn (Q.Z.); 2022126059@chd.edu.cn (Z.L.); 2022126015@chd.edu.cn (C.Z.); 2022126019@chd.edu.cn (H.J.)

² Key Laboratory of Loess, Xi'an 710054, China

³ School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China; huangwubiao@whu.edu.cn

* Correspondence: mingtaoding@chd.edu.cn; Tel.: +86-151-0294-8168

Abstract: The reactivation of old landslides can be triggered by heavy destructive earthquakes, heavy rainfall, and ongoing human activities, thereby resulting in the occurrence of secondary landslides. However, most existing models are designed for detecting nascent landslides and there are few algorithms for old landslide detection. In this paper, we introduce a novel landslide detection model known as YOLOv8-CW, built upon the YOLOv8 (You Only Look Once) architecture, to tackle the formidable challenge of identifying old landslides. We replace the Complete-IOU loss function in the original model with the Wise-IOU loss function to mitigate the impact of low-quality samples on model training and improve detection recall rate. We integrate a CBAM (Convolutional Block Attention Module) attention mechanism into our model to enhance detection accuracy. By focusing on the southwest river basin of the Sichuan–Tibet area, we collect 558 optical remote sensing images of old landslides in three channels from Google Earth and establish a dataset specifically for old landslide detection. Compared to the original model, our proposed YOLOv8-CW model achieves an increase in detection accuracy of 10.9%, recall rate of 6%, and F1 score from 0.66 to 0.74, respectively. These results demonstrate that our improved model exhibits excellent performance in detecting old landslides within the Sichuan–Tibet area.



Citation: Li, Y.; Ding, M.; Zhang, Q.; Luo, Z.; Huang, W.; Zhang, C.; Jiang, H. Old Landslide Detection Using Optical Remote Sensing Images Based on Improved YOLOv8. *Appl. Sci.* **2024**, *14*, 1100. <https://doi.org/10.3390/app14031100>

Academic Editors: Ke Wu, Yuxiang Zhang and Yi Wang

Received: 15 November 2023

Revised: 22 January 2024

Accepted: 26 January 2024

Published: 28 January 2024



Copyright: © 2024 by the authors. 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/>).

Keywords: old landslide detection; optical remote sensing images; YOLOv8; Wise-IOU loss function; attention module; YOLOv8-CW

1. Introduction

An old landslide is the result of prolonged and intricate geological processes occurring on slopes [1]. Although the majority of old landslides exhibit long-term stability, they possess the potential for reactivation and renewed sliding. This reactivation is triggered by factors, notably seismic events, precipitation, and human-engineered activities, which exert their influence on pre-existing slide accumulations. The slide mass resulting from old landslides is a site for human activities, primarily influenced by the underlying topography and landforms [2]. In recent years, the escalating human engineering activities and ever-changing global climatic conditions have led to a sharp increase in the frequency of old landslide reactivations. This phenomenon has resulted in significant harm to both human life and property safety, as well as the natural environment [3–5]. Therefore, to safeguard the safety of human lives and property, it is imperative to undertake extensive detection and monitoring of old landslides on a large scale.

The existing landslide detection methods mainly fall into the following three categories: (1) Visual interpretation method. This approach relies excessively on expert experience, demanding a substantial investment of time and effort, leading to relatively low efficiency [6,7]. Nevertheless, its high accuracy in identifying old landslides compensates for these limitations; (2) Machine learning method. In contrast to the visual interpretation

method, this approach exhibits a higher degree of automation [8–12]. However, it necessitates the extraction of a large number of image features and involves conducting extensive feature selection and hyperparameter tuning experiments based on the feature data, which incurs a substantial workload; (3) Deep learning method. Over the past few years, Convolutional Neural Networks have witnessed rapid advancements and achieved remarkable milestones in the domain of image processing [13–18]. In contrast to traditional machine learning methods, deep learning obviates the need for manual feature engineering and selection when dealing with landslide characteristics. Moreover, deep learning is amenable to larger sample sizes and is well suited for landslide detection in more expansive scenes.

The detection of nascent landslides has experienced rapid advancements in the field of deep learning [19]. However, the detection of old landslides still faces technical challenges that need to be solved in relevant research areas. The presence of old landslides is characterized by their long history and considerable time interval, leading to various degrees of transformation over time [20]. Subsequent to the occurrence of a landslide, vegetation often reestablishes itself over several years, blending with the surrounding environment. As a result, differentiating old landslides from recent ones primarily relies on scrutinizing the inherent morphological characteristics of the landslide itself and identifying certain traces of human-induced alteration [21]. Zili et al. [20] designed an iterative classification and semantic segmentation network to classify and segment old landslides on the Loess Plateau, and the results show that the designed network is extremely effective for old landslides that are difficult to identify there. For the semantic segmentation task, the F1 score increased from 0.5054 to 0.5448 and the detection accuracy of the old landslide improved to 0.9 compared to the basic network. Yuanzhen et al. [22,23] used Mask R-CNN to automatically identify old landslides in the loess area. The results show that the two-stage algorithm has a better ability to detect old landslides in the Loess Plateau. Zhaoying et al. [24] used CNNs and DEM data to identify old landslides in the Loess Plateau, with a detection rate of 95.7% and a recall rate of 100%.

However, the relevant algorithms for the detection of old landslides in the southwest river basin of the Sichuan–Tibet area are extremely lacking in this field and need to be improved [25]. In the existing research, the landslide detection model is relatively complex, with limited accuracy and generalization ability for the location and range of the old landslides in the southwest mountainous area.

To solve the above problems, this paper proposes a detection method based on the improved YOLOv8 [26,27] (YOLOv8-CW). This method improves the detection accuracy of old landslides in various geomorphic environments and exhibits robust performance in detecting complex old landslides. The primary contributions of this paper can be summarized as follows: Firstly, an old landslides dataset is compiled using Google Earth images, with a focus on the southwest river basin of the Sichuan–Tibet area in China as the designated research area [23]. Then, the boundary box loss function Complete- IoU (CIoU) loss function is replaced by the Wise- IoU (WIoU) loss function [28–30], which solves the problem of model training caused by low-quality samples and introduces a CBAM attention mechanism to improve model detection ability [31–35]. Finally, we compare the WIoU loss functions of different versions and different attention mechanisms. The experimental results clearly demonstrate that the improved method has superior detection ability and generalization and has great application value to the detection of old landslides in the Sichuan–Tibet area of China.

2. Study Area and Dataset

The research area is located in the southwest river basin of the Sichuan–Tibet area, which is renowned for its abundance of hydraulic resources, as shown in Figure 1. The study area encompasses 359,945 km² across the river basins of the Dadu River, Jinsha River, Nujiang River, and Minjiang River in the Sichuan–Tibet area. The diverse plane morphologies of old landslides encompass long-tongue shapes, ovals, trumpets, irregular formations, and more. These old landslides have typically evolved through the gradual

accumulation of rockfalls, resulting in an inverted triangle, conical, or fan-shaped appearance on the plane (Figure 2). Notably, some large and extra-large old landslides have distinctive armchair landforms with clear boundaries, which have now become inhabited areas, serving as the main living places for present-day residents [36,37]. The dataset used in this study comprised Google Earth images with a spatial resolution of 2 m and covered a time span from 2012 to 2022. According to the analysis provided by geological experts, a total of 329 landslides have been identified within the southwest river basin of the Sichuan–Tibet area. Leveraging the manipulation of Google Earth’s three-dimensional perspective and adjusting the timing of image acquisition, a multitude of landslide remote sensing images were successfully captured. Subsequently, the collected image data underwent processing using the LabellImg tool to construct the dataset. To facilitate model training and evaluation, the dataset was divided into training and validation sets in an 8:2 ratio. Additionally, data augmentation techniques, such as image cropping and rotation, were employed to enhance the diversity and robustness of the data. The augmented training dataset comprised 2500 images, while the validation dataset encompassed 600 images.

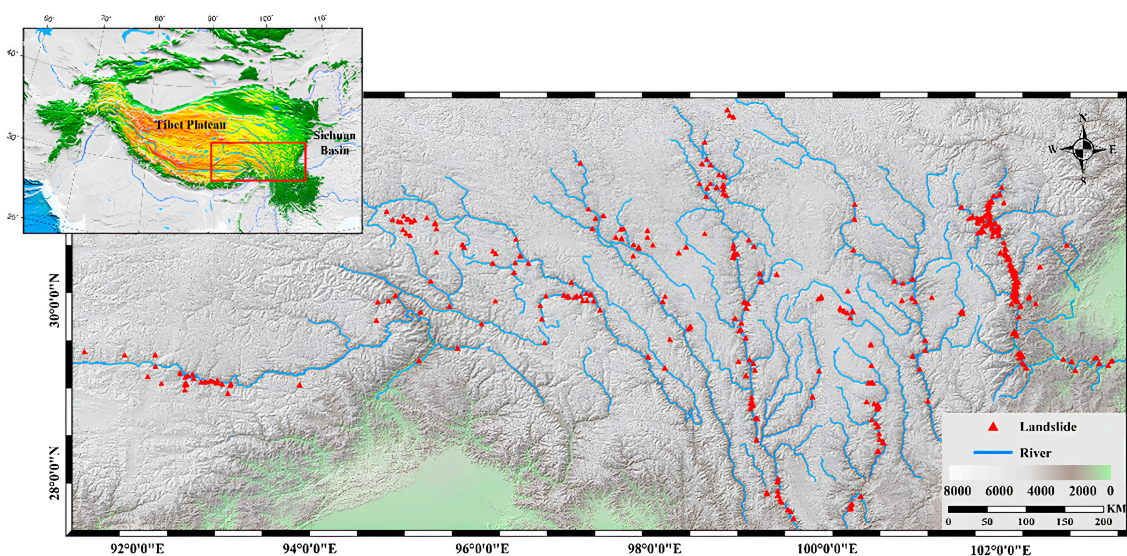


Figure 1. The Sichuan–Tibet study area.

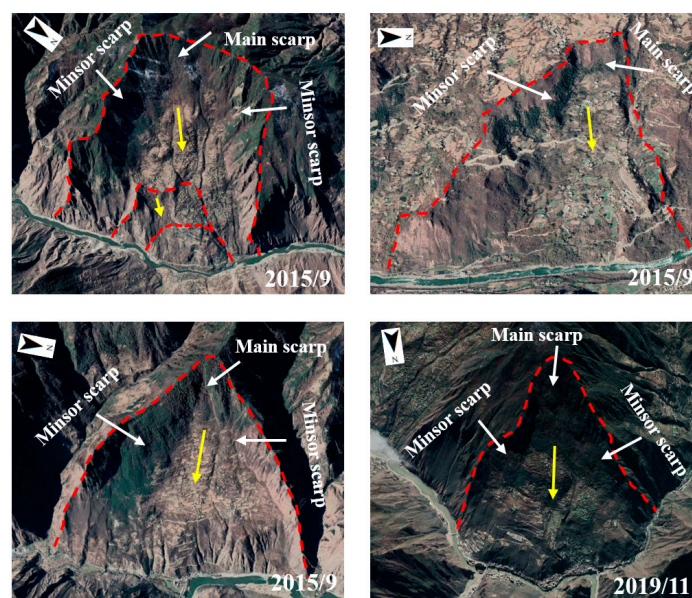


Figure 2. Examples of old landslides. Red lines indicate the old landslide boundary and the yellow arrow indicates the direction of slide.

3. Method

Firstly, high-resolution old landslide image data in the study area are collected based on Google Earth images. These images are subsequently segmented into sizes of 640×640 , labels are marked, and data enhancement is performed to complete the dataset production. Next, the network structure of the YOLOv8 model is improved by adding a CBAM attention mechanism to the backbone network and modifying the IoU loss function at the detection head to enhance model accuracy. Finally, the finished dataset was trained on the improved model for accuracy assessment. Figure 3 illustrates the old landslide detection process using the improved YOLOv8 algorithm.

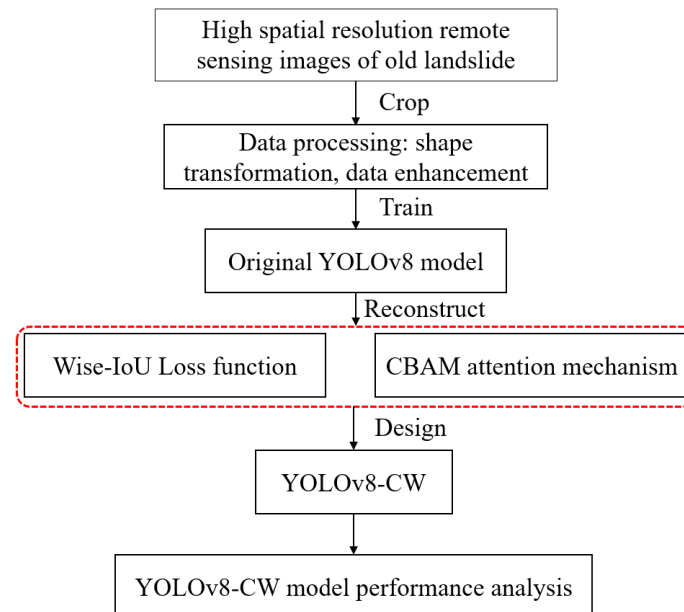


Figure 3. Experimental flowchart. Red dashed lines represent model improvements.

3.1. YOLOv8 Model

The YOLOv8 model is a one-stage object detection model, which was proposed by Ultralytics in 2023 [27]. It is an improvement upon the YOLOv5 algorithm [26], featuring a more streamlined network structure with fewer parameters and higher detection precision. The basic code for the YOLOv8 algorithm can be obtained from the GitHub website at <https://github.com/ultralytics/ultralytics>, (accessed on 24 April 2023). The network architecture of the YOLOv8 model is composed of three key components: the backbone network, the neck network, and the head network. It differs slightly from the overall structure of the previous YOLO model. The main network structure is illustrated in Figure 4.

In the backbone network part, YOLOv8 made a significant improvement by replacing the C3 module used in YOLOv5 (Figure 5) with the C2f structure shown in Figure 4. The C2f structure offers a more abundant gradient flow, enhancing the flow of information throughout the network. Additionally, YOLOv8 adjusted the channel numbers differently for different scale models, which further contributed to the overall performance improvement. This enhancement in the backbone network played a crucial role in significantly improving the performance of the YOLOv8 model, leading to better object detection accuracy and more robust feature extraction capabilities [38–40].

In the head network section, the previous YOLO model utilized a coupled head structure (Figure 6) that directly inputted the feature output from the convolutional layer into the fully connected layer to output target position and category. However, compared to this structure, the decoupled head (shown in Figure 4) calculates classification and position loss functions separately for obtained feature graphs. This approach effectively reduces parameters and computational complexity while enhancing the generalization ability and robustness of the model. Additionally, YOLOv8 adopts an anchor-free detection method

that removes preset anchor frames to predict target boundary frames directly. Although this method has stronger generalization ability, simpler framework design, and better abnormal scale target detection than anchor frame-based methods, it is not suitable for general object detection.

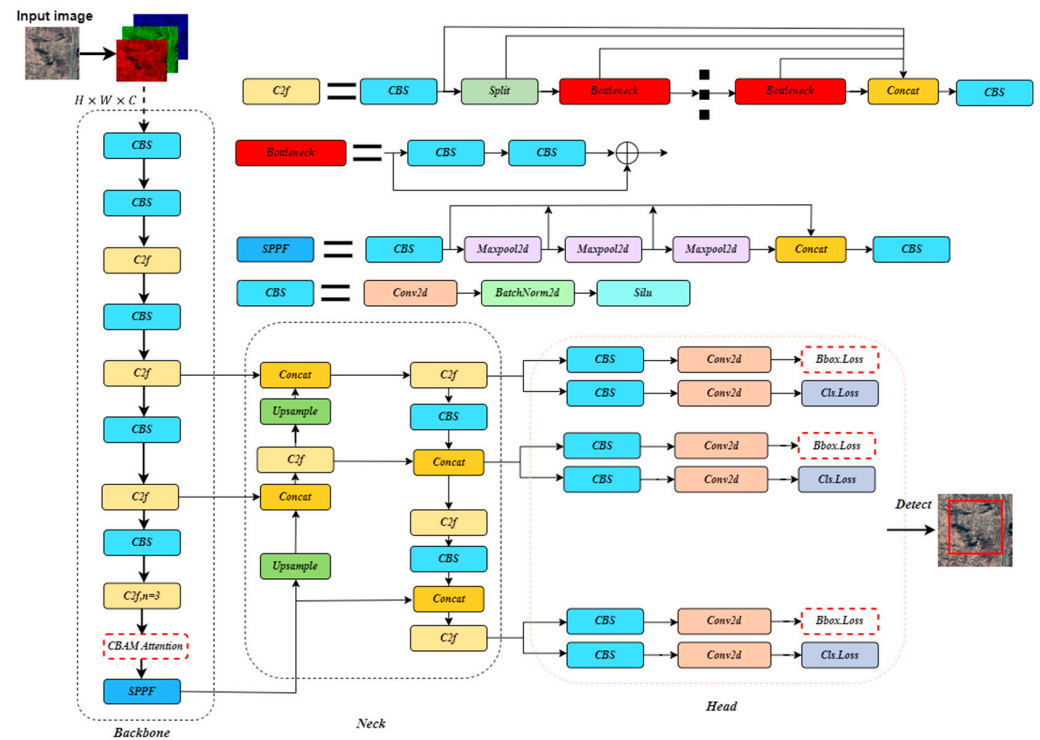


Figure 4. YOLOv8 model. The red dotted box in the backbone network is the location for adding the CBAM attention module and the red dotted box in the head network represent the loss function improved in this paper. Conv2d, BatchNorm2d, Maxpool2d, and Silu denote Two-dimensional Convolution, Batch Normalization for 2D, 2D Max Pooling, and Silu activation function, respectively.

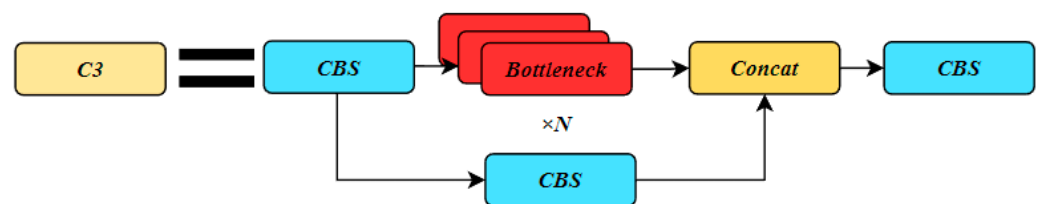


Figure 5. C3 module diagram.

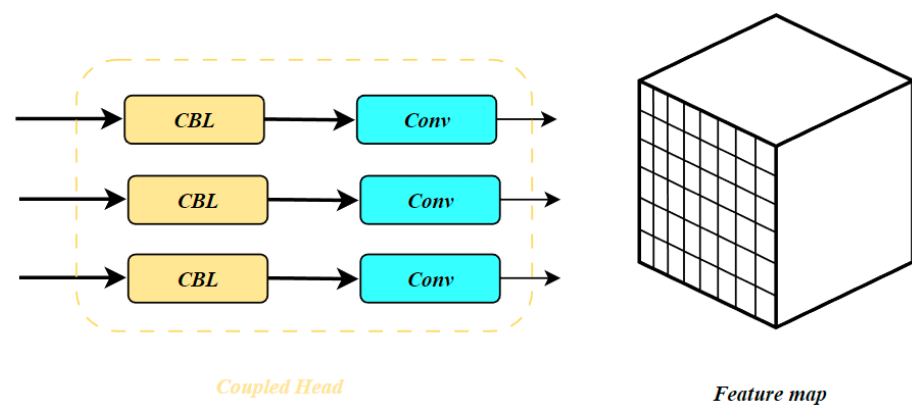


Figure 6. Coupled head structure. The CBL consists of Convolution, Batch Normalization, and LeakyRelu loss function. Conv denotes Convolution.

Regarding loss functions, the YOLO series adopted a static allocation strategy. However, recognizing the superior performance of dynamic allocation strategies, the YOLOv8 algorithm directly incorporates Task Aligned Assigner. The core concept is to select positive samples based on scores weighted by both classification and regression scores. When calculating the loss function, the Distribution Focal Loss is introduced to tackle the challenge of highly imbalanced quantities of positive and negative samples.

3.2. CBAM Attention Mechanisms

CBAM is an attention mechanism designed for Convolutional Neural Networks to enhance their performance [33]. It consists of two essential submodules: channel attention and spatial attention (Figure 7).

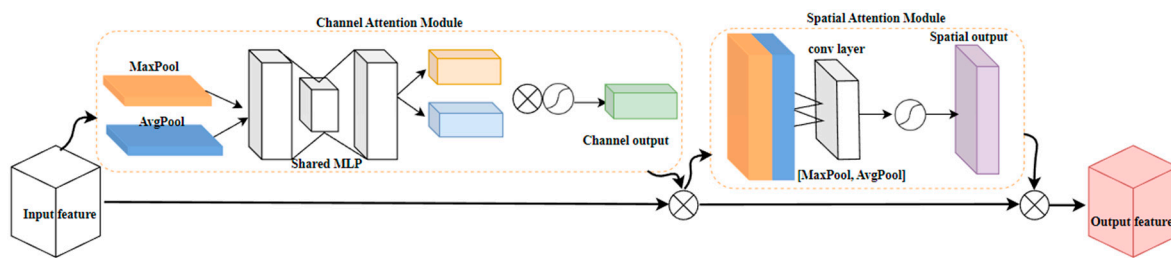


Figure 7. CBAM attention mechanism.

Channel attention involves weighted average pooling of the feature for each channel. The resultant vector undergoes two fully connected layers to obtain a weight vector using the Sigmoid function. This weight vector is then multiplied with the input original feature and its residual feature, thereby providing attention weighting to distinct channels. To sum up, the channel attention is computed as:

$$M_C(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

In Equation (1), $M_C(F)$ denotes channel attention mechanism operation, where F is the input feature, σ indicates the Sigmoid function, and MLP is a multi-layer perceptron.

Spatial attention involves weighted average pooling of pixel features, followed by two fully connected layers to yield a weight vector. This weight vector is multiplied with the original feature to produce a weighted feature. CBAM attention mechanisms enhance model performance by exploiting the interplay between channels and spatial dimensions. This mechanism has demonstrated substantial success in various large-scale image classification tasks. The spatial attention is computed as:

$$M_S(F) = \sigma(f^{7 \times 7}([AvgPool(F_1); MaxPool(F_1)])) \quad (2)$$

where $M_S(F)$ denotes spatial attention mechanism operation, $f^{7 \times 7}$ denotes a convolution operation with the filter size of 7×7 , and F_1 denotes the feature output through the channel attention operation.

3.3. WIoU Loss Function

Due to the influence of weather changes, geological movements, and human activities, the morphological characteristics of old landslides have significantly changed compared to their initial formation. Climate warming and accelerated vegetation growth have led to extensive vegetation cover on the slopes and walls of old landslides. Consequently, when constructing the old landslide dataset, a large number of low-quality samples can adversely affect the detection performance. To avoid the detrimental impact of low-quality samples on the model's training process and to emphasize the significance of high-quality old landslide samples, this study replaced the boundary box loss function, Complete-IoU loss function, in the original YOLOv8 model with the Wise-IoU loss function. This adjustment

is intended to improve the model's detection performance when working with historical landslide datasets.

There are three versions of WIoU, representing three different ways of constructing loss functions. WIoUv1 constructs the attention-based boundary box loss, while WIoUv2 and WIoUv3 add a focusing mechanism by constructing the gradient gain (focusing coefficient) calculation method.

3.3.1. Wise-IoUv1

When training an old landslide dataset, encountering numerous low-quality samples, commonly referred to as difficult samples, is inevitable. Geometric measures, such as aspect ratio and distance, can exacerbate the penalty for these challenging samples, resulting in a reduction in the model's generalization performance. An effective loss function should alleviate the geometric penalty when the predicted bounding box aligns well with the ground truth bounding box, while also emphasizing less intervention training to enhance the model's overall generalization ability. Building upon this foundation, distance attention is crafted based on distance measurements, yielding a WIoUv1 loss function with a two-layer attention mechanism.

$$\mathcal{L}_{WIoUv1} = \mathcal{R}_{WIoU} \mathcal{L}_{IoU} \quad (3)$$

Following Equation (3), \mathcal{R}_{WIoU} can be calculated by Equation (4).

$$\mathcal{R}_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right), \mathcal{R}_{WIoU} \in [1, e) \quad (4)$$

In Equation (4), the parameters are shown in Figure 8. Since the \mathcal{R}_{WIoU} is always greater than 1, this will effectively enhance the \mathcal{L}_{IoU} of the ordinary quality anchor frame. On the contrary, because the \mathcal{L}_{IoU} is less than 1, it will notably diminish the \mathcal{R}_{WIoU} for high-quality anchor frames and reduce the emphasis on the center point distance when the anchor frame and the target frame are well aligned. If the anchor frame and the target frame are well aligned, to prevent \mathcal{R}_{WIoU} from creating gradients that hinder convergence, W_g and H_g are separated from the calculation step (* indicates this operation). Since it effectively removes the factors that impede convergence, it does not introduce any new metrics, such as aspect ratio, normalized length, or other geometric measurements.

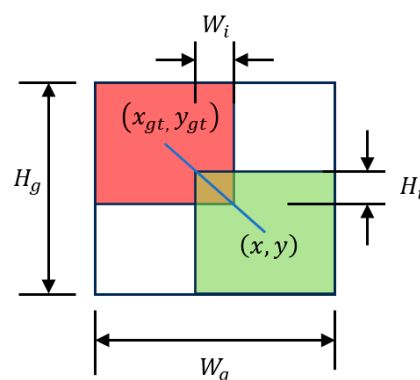


Figure 8. IoU diagram. The green box denotes the real box and the red box denotes the prediction box.

3.3.2. Wise-IoUv2

Lin et al. [41] designed a monotony focusing mechanism for cross-entropy, which effectively reduces the contribution of low-quality samples to loss values, makes the model pay more attention to difficult samples, and improves classification performance. Similarly,

the monotonic focusing coefficient, $\mathcal{L}_{IoU}^{\gamma*}$ of Wise-IoUv1 is first constructed, and the loss function of Wise-IoUv2 is then calculated by Equation (5).

$$\mathcal{L}_{WIoUv2} = \mathcal{L}_{IoU}^{\gamma*} \mathcal{L}_{WIoUv1}, \gamma > 0 \quad (5)$$

During the training of the model, the focusing coefficient $\mathcal{L}_{IoU}^{\gamma*}$ decreases with the decrease of \mathcal{L}_{IoU} , resulting in a slow convergence rate in the later training period. Therefore, the mean value of \mathcal{L}_{IoU} is introduced as a normalization factor, and the Wise-IoUv2 loss function is calculated by Equation (6).

$$\mathcal{L}_{WIoUv2} = \left(\frac{\mathcal{L}_{IoU}^*}{\overline{\mathcal{L}_{IoU}}} \right)^\gamma \mathcal{L}_{WIoUv1}, \gamma > 0 \quad (6)$$

The gradient gain is kept at a high level by dynamically updating the normalization factor ($r = \left(\frac{\mathcal{L}_{IoU}^*}{\overline{\mathcal{L}_{IoU}}} \right)^\gamma$), which solves the problem of slow convergence in the late training period.

3.3.3. Wise-IoUv3

The Wise-IoUv3 loss function directly replaces the IoU value with the outlier degree to describe the mass of the anchor frame. The outlier degree is calculated by Equation (7).

$$\beta = \frac{\mathcal{L}_{IoU}^*}{\overline{\mathcal{L}_{IoU}}}, \beta \in [0, +\infty) \quad (7)$$

As observed in the given formula, the mass of the anchor frame increases as the outlier becomes smaller, and a non-monotonic focusing coefficient, β , is constructed. This coefficient, when multiplied by Wise-IoUv1, yields Wise-IoUv3 (Equation (8)). The introduced coefficient allows the regression of bounding frames to focus on anchor frames with ordinary mass. By assigning smaller gradient gains to anchors with larger outliers, the model can effectively prevent harmful gradients from low-quality samples. This approach helps in improving the overall performance and robustness of the model during the training process.

$$\mathcal{L}_{WIoUv3} = \theta \mathcal{L}_{WIoUv1}, \theta = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (8)$$

In Equation (8), α and δ are hyperparameters, which need to be selected according to specific experiments. Since $\overline{\mathcal{L}_{IoU}}$ is in a dynamic process, the quality division criteria of the anchor frame are also dynamic, which enables the Wise-IoUv3 loss function to create the gradient gain allocation strategy that best meets the current situation during each training process.

3.4. Model Evaluation Methods

In this study, precision, recall, F1 score, and mean average precision are used to evaluate the prediction ability of the model [42]. The above indexes can be calculated by the confusion matrix (Table 1).

Table 1. The confusion matrix.

Real	Predicate	
	Landslide	Background
Landslide	TP	FN
Background	FP	TN

The precision, recall, and F1 score in Table 1 are computed using Equations (9)–(11), respectively. Based on these calculations, a precision–recall rate curve can be plotted, with the recall on the horizontal axis and the precision on the vertical axis. The AP represents the area between this curve and the axis, which can be determined using Equation (12). After obtaining the AP for each category, the mean average precision is calculated as Equation (13).

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TN + FP} \quad (10)$$

$$F1 = \frac{2 \times P \times R}{(P + R)} \quad (11)$$

$$AP = \int_0^1 \text{precision}(\text{recall})d(\text{recall}) \quad (12)$$

$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (13)$$

4. Results

4.1. Experimental Setup

In this study, experiments are carried out on a workstation featuring an Intel i5-13400F processor (Intel, Santa Clara, CA, USA), 32 GB of Random Access Memory (RAM) (Dell, Round Rock, TX, USA), and an NVIDIA RTX 4080 graphics processor (NVIDIA, Santa Clara, CA, USA) equipped with 16 GB of video memory. The experiments are conducted using the PyTorch 3.8 deep learning framework and implemented in the Python programming language.

The Stochastic Gradient Descent (SGD) optimizer is employed for all training procedures, with an initial learning rate of 0.001 and a batch size of 32. The training consists of 200 epochs, incorporating a weight decay factor of 0.005 and a momentum factor of 0.937. The prediction box threshold is set at 0.7. Within the optimizer, the learning rate decay is managed through the cosine annealing scheduler. During the training phase, the input images are normalized within the range of (0, 1).

4.2. Model Assessment

4.2.1. Model Results for Old Landslide Detection

After the experiment, the improved YOLOv8 model was compared with the original YOLOv8 model, and the results are shown in Table 2. The YOLOv8 model, after adding a CBAM attention mechanism and replacing the IoU loss function, performs better than the original YOLOv8 model. Among them, the YOLOv8-CW2 model using the WIoUv2 loss function performs the best. The F1 score is 0.74, and the precision rate is increased by 10.9%, the recall rate by 6%, and mAP by 11% compared to the original model. Compared with the YOLOv8-CW1 model and the YOLOv8-CW3 model using WIoUv1 and WIoUv3 functions, respectively, the YOLOv8-CW2 model also achieved better experimental results. The improved YOLOv8 is compared with classic models such as YOLOv5 and Retinanet, demonstrating its superior detection ability. The test results reveal a significant increase in F1 scores of 9.4% and 10.2%, respectively, validating the suitability of the YOLOv8-CW2 model for old landslide detection.

Figure 9 shows a comparison of the detection results between the original YOLOv8 algorithm and the improved YOLOv8 algorithm applied to the old landslide dataset in the study area. From the graph, it is evident that the original YOLOv8 model exhibits differences in detection performance compared to the improved YOLOv8 model, both for small and large old landslides. The complexity of the background environment in high-resolution remote sensing images is much greater than that of natural images, leading

to the original YOLOv8 model incorrectly identifying certain objects in the background environment, such as exposed soil that resembles landslide image characteristics.

Table 2. Comparison of detection performance for different models.

Model	Precision (%)	Recall (%)	mAP (%)	F1 (%)
YOLOv8	69.2%	63.1%	62.8%	66.0%
YOLOv8-CW1	76.7%	65.9%	68.2%	71.0%
YOLOv8-CW2	80.1%	69.1%	73.8%	74.0%
YOLOv8-CW3	79.0%	66.8%	70.3%	72.0%

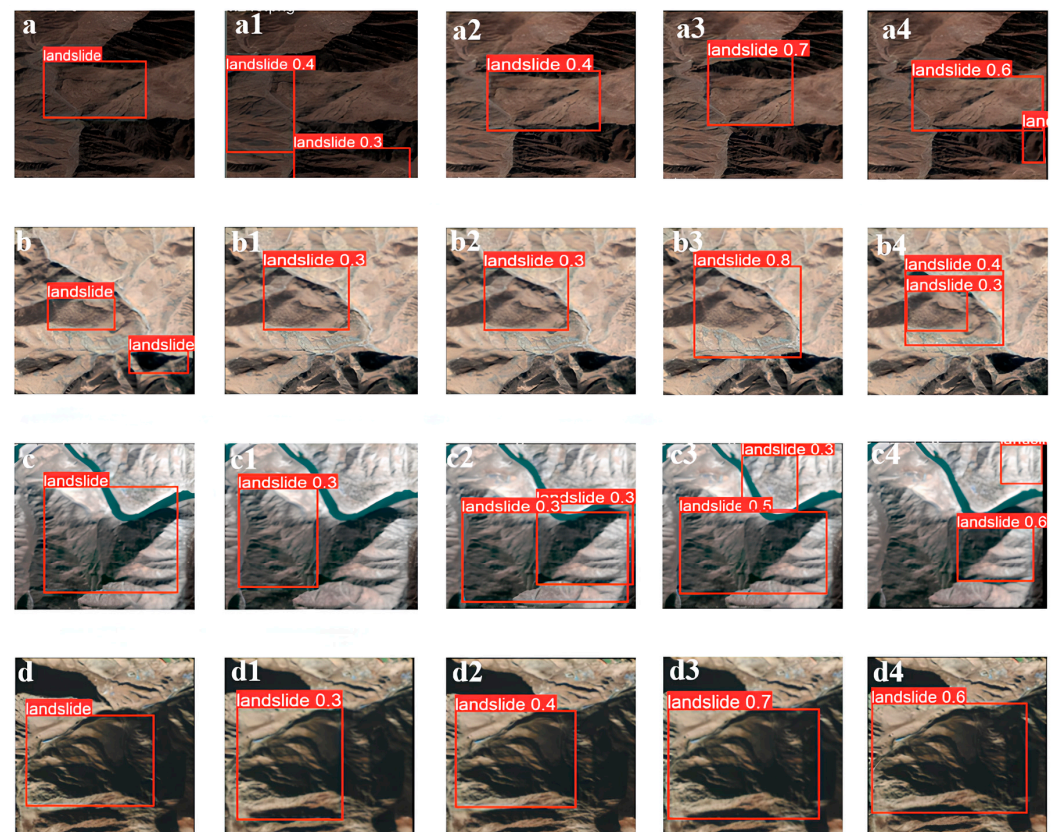


Figure 9. Landslide detection results ((a–d) is the mark location of old landslides; (a1–d1) indicates the detection results of the original YOLOv8 model; (a2–d2) indicates the detection results of the YOLOv8-CW1 model; (a3–d3) indicates the detection results of the YOLOv8-CW2 model; (a4–d4) indicates the detection results of the YOLOv8-CW3 model).

Comparing the detection results of the three improved YOLOv8 models, it is observed that the YOLOv8-CW2 model achieves highly accurate detection for both small and large old landslides. In the case of large old landslides, the YOLOv8-CW3 model performs exceptionally well, with higher accuracy compared to the other two improved models. However, the YOLOv8-CW3 model shows a higher number of overlapping detection boxes when detecting small old landslides, leading to inaccurate detection of the actual locations of the landslides.

Overall, the improved YOLOv8 models demonstrate better performance in detecting old landslides in the study area, especially for small and large landslides, indicating their effectiveness in addressing the complexities of high-resolution remote sensing images.

4.2.2. Comparison of Different Attention Modules

In this study, two attention modules, SE and SK, are integrated into the baseline model for comparative analysis. Subsequently, the most suitable attention module is selected to

construct the optimal model [34,35]. Table 3 shows the results of the experiment using the same old landslide dataset.

Table 3. The effects of different attentional mechanisms.

Model	Precision (%)	Recall (%)	mAP (%)	F1 (%)
YOLOv8-CW1	76.7%	65.9%	68.2%	71.0%
YOLOv8-CW2	80.1%	69.1%	73.8%	74.0%
YOLOv8-CW3	79.0%	66.8%	70.3%	72.0%
YOLOv8-SEW1	73.2%	68.3%	63.4%	71.0%
YOLOv8-SEW2	75.6%	71.5%	70.6%	74.0%
YOLOv8-SEW3	76.6%	63.4%	66.1%	69.0%
YOLOv8-SKW1	74.9%	70.0%	70.9%	72.0%
YOLOv8-SKW2	74.4%	70.3%	70.7%	72.0%
YOLOv8-SKW3	73.3%	66.8%	64.7%	70.0%

It is obvious from the table that the model with a CBAM attention mechanism has better performance. The accuracy and mAP of the YOLO-CW2 model are much higher than those with the SE and SK attention mechanism. In terms of recall rate, the model with the SE attention mechanism has better performance. In short, the YOLOv8-CW2 model is selected as the improved model in this experiment.

5. Discussion

5.1. The Effects of the WIoU Loss Function on the Model

The effects of using the WIoU loss function to detect old landslides are obviously better than those of using the CIoU loss function to detect old landslides, as shown in Table 4. After replacing the loss function, the model is significantly higher than the original model in all evaluation indexes. Figure 10 shows the training loss curves of the CIoU, WIoUv1, WIoUv2, and WIoUv3 networks, respectively. The loss curves decreased significantly at the beginning of the training and the decline trend began to slow down after 100 epochs. The total loss function converges to 0.39787, compared to the original loss function of YOLOv8. The improved loss function after the introduction of WIoU can make the network converge faster, so the improvement of the network loss function in this paper is reasonable.

Table 4. Comparison of ablation experiment results.

	CBAM	SE	SK	CIoU	WIoUv1	WIoUv2	WIoUv3	mAP (%)
Base line	-	-	-	✓	-	-	-	62.8%
1	-	-	-	-	✓	-	-	68.3%
2	-	-	-	-	-	✓	-	67.2%
3	-	-	-	-	-	-	✓	67.9%
4	✓	-	-	✓	-	-	-	65.8%
5	✓	-	-	-	✓	-	-	68.2%
6	✓	-	-	-	-	✓	-	73.8%
7	✓	-	-	-	-	-	✓	70.3%
8	-	✓	-	✓	-	-	-	63.7%
9	-	✓	-	-	✓	-	-	63.4%
10	-	✓	-	-	-	✓	-	70.6%
11	-	✓	-	-	-	-	✓	66.1%
12	-	-	✓	✓	-	-	-	64.6%
13	-	-	✓	-	✓	-	-	70.9%
14	-	-	✓	-	-	✓	-	70.7%
15	-	-	✓	-	-	-	✓	64.7%

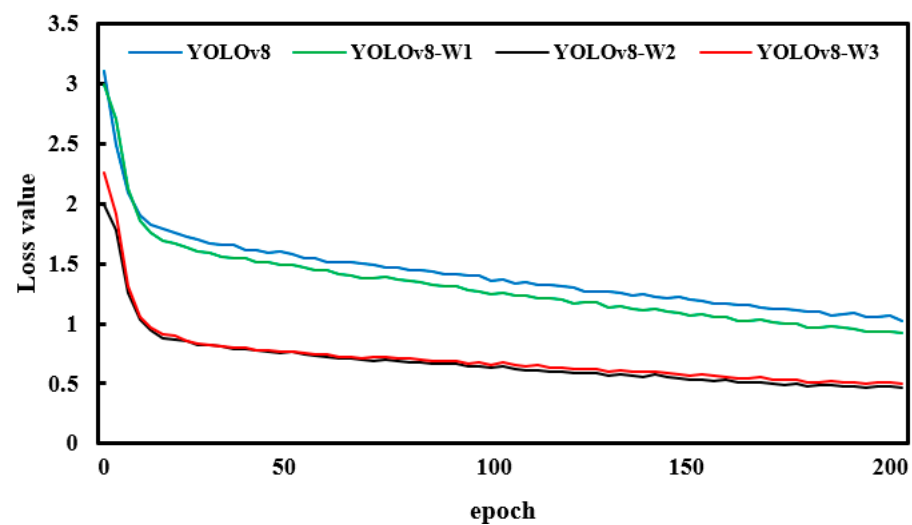


Figure 10. The effects of different Wise-IoU loss functions.

5.2. The Effects of the Attention Mechanism on the Model

To investigate the impact of the attention mechanism on the YOLOv8 model, this experiment compares the YOLOv8 model without the addition of the attention mechanism, and the research results are presented in Table 5. From Table 5, it is evident that the improved model with the CBAM attention mechanism performs better; in particular, the YOLOv8-CBAM model with the attention mechanism shows a 1.7% increase in recall rate and a 3% increase in mean average precision, with its F1 score rising from 0.66 to 0.71.

Table 5. The effects of attention mechanisms.

Model	Precision (%)	Recall (%)	mAP (%)	F1 (%)
YOLOv8	69.2%	63.1%	62.8%	66.0%
YOLOv8-CBAM	77.6%	64.8%	65.8%	71.0%
YOLOv9-SE	74.4%	64.6%	63.7%	70.0%
YOLOv8-SK	72.2%	65.4%	64.6%	69.0%

It is observed that the attention mechanism enables the model to focus more on the detection target and suppress irrelevant information (Figure 11), by analyzing the heat map of the feature layer with the added attention mechanism. The CBAM attention mechanism, which incorporates the spatial attention mechanism in addition to the channel attention mechanism, exhibits a heightened focus on the landslide surface and enables more accurate localization of old landslides. By combining both spatial and channel information, the CBAM attention mechanism ensures that no location information is lost and facilitates improved detection performance.

5.3. Limitations and Future Challenges

Although the improved YOLOv8 model has demonstrated superior detection performance in the study of old landslides, there remains three major limitations within this field.

- (1) Limited sample size: Deep learning models, including YOLOv8-CW2, require extensive datasets for optimal training. However, there is a critical shortage of old landslide data in the Sichuan and Tibet regions of China. It is anticipated that the performance of the model could be further improved with the availability of more comprehensive datasets.
- (2) Limited data type: The complexity of detecting old landslides in the Sichuan and Tibet regions is exacerbated by diverse topography, landforms, climate variations,

and soil types, as well as the high level of vegetation coverage due to minimal human activity. These factors make the use of optical image data for landslide detection somewhat reductive. Therefore, we should integrate high-precision DEM (Digital Elevation Model) data and InSAR (Interferometric Synthetic Aperture Radar) data to enable a more thorough assessment of old landslides in these areas.

- (3) Limited model transferability: The transferability of the improved YOLOv8 model to other regions and landslide types has not been fully validated. To improve the model's robustness and generalization capabilities, it is essential to expand the dataset with varied types of old landslide data, such as those from the Loess Plateau of China. This will allow the model to learn additional features relevant to different landslide characteristics and environments.

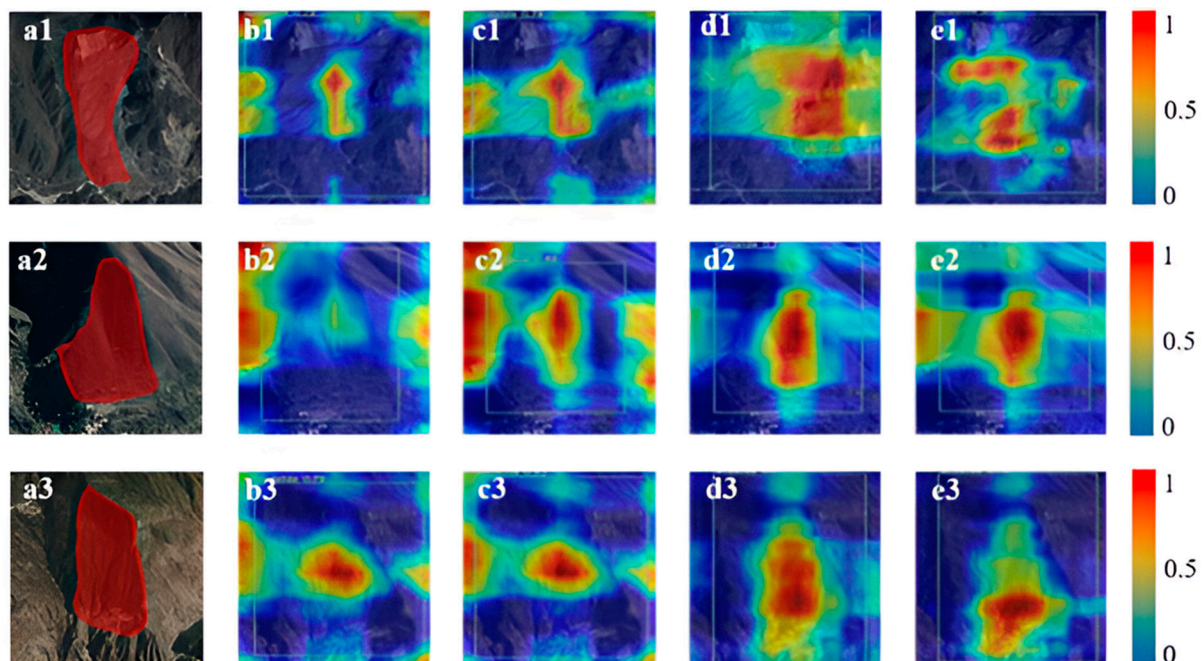


Figure 11. Heat map comparison ((a1–a3) indicates the actual boundary of the old landslide; (b1–b3) represents the YOLOv8; (c1–c3) represents the YOLOv8-SK; (d1–d3) represents the YOLOv8-CBAM; (e1–e3) represents the YOLOv8-SE.).

6. Conclusions

This study utilizes Google Earth images as the data source to establish an old landslide dataset for the Sichuan–Tibet river basin. Leveraging the YOLOv8 deep learning model, this paper replaces the Wise-IoU bounding box loss function and introduces the CBAM attention mechanism, constructing the YOLOv8-CW network that incorporates three different bounding box loss functions. A general detection method for old landslides is proposed, demonstrating strong detection capabilities for old landslides in the mountainous region of Southwest China and presenting promising application prospects. The main conclusions are as follows:

1. In the detection of old landslides in the Sichuan–Tibet river basin, the improved YOLOv8-CW2 model achieves a detection accuracy of 80.1%, recall rate of 69.1%, mAP of 73.8%, and F1 score of 0.74. Compared to the original YOLOv8 model, the accuracy and recall rate are increased by 10.9% and 6%, respectively. The F1 score is increased from 0.66 to 0.74.
2. Comparing the detection results of the YOLOv8-CW2 model and the original YOLOv8 model on the old landslide dataset reveals substantial improvements in the accuracy of both model types. This indicates the effectiveness and feasibility of detecting old landslides in the Sichuan–Tibet river basin using the proposed optimization method.

3. To further enhance the stability and recognition ability of the model, multiple types of landslide data sources can be utilized for training, ultimately achieving accurate identification of multi-source data and various types of landslides. This approach provides timely and precise data support for landslide disaster rescue and disaster assessment.

Despite achieving improvements in old landslide detection accuracy, this study acknowledges certain limitations. Due to the non-obvious characteristic form of old landslides, there are instances of missed detections, necessitating more precise classification and segmentation of old landslides. When detecting old landslides, it is recommended to combine InSAR and other applications, superimpose DEM and other multi-source data to construct comprehensive datasets, and utilize deep learning and other methods for accurate landslide detection. As geological disaster detection and identification enters the era of artificial intelligence, the use of automation technologies such as deep learning can significantly improve identification efficiency, holding significant implications for landslide prevention and mitigation across generations.

Author Contributions: Conceptualization, Q.Z.; data analysis, Y.L.; experimentation, Y.L.; funding support, M.D.; illustration painting, Z.L. and H.J.; image data collection, Y.L.; research design, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, M.D., W.H. and C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key Research and Development Program of China (2021YFC3000400), the National Natural Science Foundation of China (42374027), the Opening Fund of Key Laboratory of Smart Earth (KF2023YB04-01), the Application and Demonstration of Comprehensive Governance and Scale Industrialization in the Sichuan–Tibet Region under the High-resolution Satellite Project (87-Y50G28-9001-22/23), Key R&D Program Projects in Zhejiang Province (2023C03177), and the Fundamental Research Funds for the Central Universities (300102262203).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: We are very grateful to Google Earth for the image data and platform support provided by ArcGIS, GMT, and Ultralytics.

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

References

1. Zhang, Y.; Ren, S.; Liu, X.; Guo, C.; Li, J.; Bi, J.; Ran, L. Reactivation mechanism of old landslide triggered by coupling of fault creep and water infiltration: A case study from the east Tibetan Plateau. *Bull. Eng. Geol. Environ.* **2023**, *82*, 291. [\[CrossRef\]](#)
2. Zhang, Y.; Wu, R.; Guo, C.; Wang, L.; Yao, X.; Yang, Z. Research Progress and Prospect on Reactivation of Ancient Landslides. *Adv. Earth Sci.* **2018**, *33*, 728–740.
3. Huang, W.; Ding, M.; Li, Z.; Yu, J.; Ge, D.; Liu, Q.; Yang, J. Landslide susceptibility mapping and dynamic response along the Sichuan-Tibet transportation corridor using deep learning algorithms. *Catena* **2023**, *222*, 106866. [\[CrossRef\]](#)
4. Huang, W.; Ding, M.; Li, Z.; Zhuang, J.; Yang, J.; Li, X.; Meng, L.e.; Zhang, H.; Dong, Y. An Efficient User-Friendly Integration Tool for Landslide Susceptibility Mapping Based on Support Vector Machines: SVM-LSM Toolbox. *Remote Sens.* **2022**, *14*, 3408. [\[CrossRef\]](#)
5. Xu, W.-J.; Xu, Q.; Liu, G.-Y.; Xu, H.-Y. A novel parameter inversion method for an improved DEM simulation of a river damming process by a large-scale landslide. *Eng. Geol.* **2021**, *293*, 106282. [\[CrossRef\]](#)
6. Huang, R.Q.; Li, W.L. Analysis of the geo-hazards triggered by the 12 May 2008 Wenchuan Earthquake, China. *Bull. Eng. Geol. Environ.* **2009**, *68*, 363–371. [\[CrossRef\]](#)
7. Fariz, T.R.; Jatmiko, R.H.; Mei, E.T.W.; Lutfiananda, F. Interpretation on aerial photography for house identification on landslide area at Bompon sub-watershed. *AIP Conf. Proc.* **2023**, *2683*, 030013.
8. Stumpf, A.; Kerle, N. Object-oriented mapping of landslides using Random Forests. *Remote Sens. Environ.* **2011**, *115*, 2564–2577. [\[CrossRef\]](#)
9. Van Den Eeckhaut, M.; Kerle, N.; Poesen, J.; Hervás, J. Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data. *Geomorphology* **2012**, *173*, 30–42. [\[CrossRef\]](#)

10. Chen, W.; Li, X.; Wang, Y.; Chen, G.; Liu, S. Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China. *Remote Sens. Environ.* **2014**, *152*, 291–301. [\[CrossRef\]](#)
11. Gorsevski, P.V.; Brown, M.K.; 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. [\[CrossRef\]](#)
12. Mezaal, M.R.; Pradhan, B. An improved algorithm for identifying shallow and deep-seated landslides in dense tropical forest from airborne laser scanning data. *Catena* **2018**, *167*, 147–159. [\[CrossRef\]](#)
13. 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. [\[CrossRef\]](#)
14. Ghorbanzadeh, O.; Meena, S.R.; Blaschke, T.; Aryal, J. UAV-Based Slope Failure Detection Using Deep-Learning Convolutional Neural Networks. *Remote Sens.* **2019**, *11*, 2046. [\[CrossRef\]](#)
15. Sameen, M.I.; Pradhan, B. Landslide Detection Using Residual Networks and the Fusion of Spectral and Topographic Information. *Ieee Access* **2019**, *7*, 114363–114373. [\[CrossRef\]](#)
16. 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. [\[CrossRef\]](#)
17. Prakash, N.; Manconi, A.; Loew, S. Mapping Landslides on EO Data: Performance of Deep Learning Models vs. Traditional Machine Learning Models. *Remote Sens.* **2020**, *12*, 346. [\[CrossRef\]](#)
18. Yu, B.; Chen, F.; Xu, C. Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015. *Comput. Geosci.* **2020**, *135*, 104388. [\[CrossRef\]](#)
19. Casagli, N.; Intrieri, E.; Tofani, V.; Gigli, G.; Raspini, F. Landslide detection, monitoring and prediction with remote-sensing techniques. *Nat. Rev. Earth Environ.* **2023**, *4*, 51–64. [\[CrossRef\]](#)
20. Lu, Z.; Peng, Y.; Li, W.; Yu, J.; Ge, D.; 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*, 1–13. [\[CrossRef\]](#)
21. Zhang, W.; Ding, B.; Zhang, W.; Zhang, G. Analysis on the Cause and Failure Mechanism of the Jiangdingya Large Landslide in Zhouqu, Gansu Province. *J. Disaster Prev. Mitig. Eng.* **2022**, *42*, 714–722.
22. Ju, Y.; Xu, Q.; Jin, S.; Li, W.; Dong, X.; Guo, Q. Automatic Object Detection of Loess Landslide Based on Deep Learning. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 1747–1755.
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. Yang, Z.; Han, L.; Zheng, X.; Li, W.; Feng, L.; Wang, Y.; Yang, Y. Landslide identification using remote sensing images and DEM based on convolutional neural network: A case study of loess landslide. *Remote Sens. Nat. Resour.* **2022**, *34*, 224–230.
25. Yan, Y.; Guo, C.; Li, C.; Yuan, H.; Qiu, Z. The Creep-Sliding Deformation Mechanism of the Jiaju Ancient Landslide in the Upstream of Dadu River, Tibetan Plateau, China. *Remote Sens.* **2023**, *15*, 592. [\[CrossRef\]](#)
26. Qi, J.; Liu, X.; Liu, K.; Xu, F.; Guo, H.; Tian, X.; Li, M.; Bao, Z.; Li, Y. An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease. *Comput. Electron. Agric.* **2022**, *194*, 106780. [\[CrossRef\]](#)
27. Wang, G.; Chen, Y.; An, P.; Hong, H.; Hu, J.; Huang, T. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. *Sensors* **2023**, *23*, 7190. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Zhaohui, Z.; Ping, W.; Dongwei, R.; Wei, L.; Rongguang, Y.; Qinghua, H.; Wangmeng, Z. Enhancing Geometric Factors in Model Learning and Inference for Object Detection and Instance Segmentation. *IEEE Trans. Cybern.* **2022**, *12*, 8574–8586.
29. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D.; Association for the Advancement of Artificial Intelligence. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 12993–13000.
30. Tong, Z.; Chen, Y.; Xu, Z.; Yu, R. Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. *arXiv* **2023**, arXiv:2301.10051.
31. Tu, Z.; Lu, Z.; Liu, Y.; Liu, X.; Li, H. Modeling Coverage for Neural Machine Translation. In Proceedings of the 54th Annual Meeting of the Association-for-Computational-Linguistics (ACL), Berlin, Germany, 07–12 August 2016; pp. 76–85.
32. Zhou, Q.; Yang, N.; Wei, F.; Zhou, M. Selective Encoding for Abstractive Sentence Summarization. In Proceedings of the 55th Annual Meeting of the Association-for-Computational-Linguistics (ACL), Vancouver, BC, Canada, 30 July–4 August 2017; pp. 1095–1104.
33. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the 15th European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
34. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023. [\[CrossRef\]](#)
35. Zhang, S.; Liu, Z.; Chen, Y.; Jin, Y.; Bai, G. Selective kernel convolution deep residual network based on channel-spatial attention mechanism and feature fusion for mechanical fault diagnosis. *Isa Trans.* **2023**, *133*, 369–383. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Wu, R.; Zhang, Y.; Guo, C.; Yang, Z.; Ren, S.; Tong, B. Reactivation characteristics and dynamic hazard prediction of an ancient landslide in the east margin of Tibetan Plateau. *Environ. Earth Sci.* **2018**, *77*, 573. [\[CrossRef\]](#)
37. Dai, Z.; Yang, L.; Zhang, N.; Zhang, C.; Zhang, Z.; Wang, H. Deformation characteristics and reactivation mechanism of an old landslide induced by combined action of excavation and heavy rainfall. *Front. Earth Sci.* **2023**, *10*, 1009855. [\[CrossRef\]](#)

38. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [[CrossRef](#)]
39. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 448–456.
40. Wang, C.-Y.; Liao, H.-Y.M.; Wu, Y.-H.; Chen, P.-Y.; Hsieh, J.-W.; Yeh, I.H. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 14–19 June 2020; pp. 1571–1580.
41. Tsung-Yi, L.; Goyal, P.; Girshick, R.; Kaiming, H.; Dollar, P. Focal loss for dense object detection. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2999–3007.
42. Hou, H.; Chen, M.; Tie, Y.; Li, W. A Universal Landslide Detection Method in Optical Remote Sensing Images Based on Improved YOLOX. *Remote Sens.* **2022**, *14*, 4939. [[CrossRef](#)]

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.