IEEEAccess®
Multidisciplinary | Rapid Review | Open Access Journal

APPLIED RESEARCH

LSI-YOLOv8: An Improved Rapid and High Accuracy Landslide Identification Model Based on YOLOv8 From Remote Sensing Images

XINBAO CHEN[1,2], **CHANG LIU**[1], **SHAN WANG**[3], **AND XINPING DENG**[3]

[1]School of Earth Sciences and Spatial Information Engineering, Hunan University of Science and Technology, Xiangtan, Hunan 411201, China
[2]Sanya Research Institute of Hunan University of Science and Technology, Sanya, Hainan 572025, China
[3]Hunan Institute of Geological Disaster Investigation and Monitoring, Changsha, Hunan 410004, China

Corresponding authors: Shan Wang (wangshan_hn@163.com) and Chang Liu (liuchang@mail.hnust.edu.cn)

ABSTRACT The rapid identification of landslides in mountainous terrain through remote sensing imagery after earthquakes or heavy rainfall is crucial for assessing the severity of the situation and providing timely assistance for post-disaster rescue operations. However, traditional manual interpretation methods and computer vision techniques tend to be inefficient and time-consuming. To address this problem, we propose a novel algorithm called LSI-YOLOv8, which builds upon the lightweight YOLOv8 to enable fast and accurate landslide identification using remote sensing imagery. In this study, we introduce optimization techniques integrated into the YOLOv8 model. Firstly, we incorporate an expandable residual attention Dilation-wise Residual Segmentation (DWRSeg) module to reconstruct its neck structure, reducing parameters and computational load. Additionally, we enhance the positioning ability of the model by replacing the bounding box loss function Complete Intersection over Union (CIoU) with Efficient Intersection over Union (EIoU), optimizing it to improve computing speed. Through comparative experiments with other state-of-the-art models, the accuracy of our LSI-YOLOv8 significantly outperforms Mask-RCNN, YOLOv5, YOLOv7, YOLOv8, and YOLOv9 model, with mAP@0.5 being 9.0%,7.7%, 4.0%, 6.3%, and 2.9% higher than other models, while mAP@0.5:0.95 is 3.1%,3.6%, 2.7%, 4.8%, and 0.3% higher than others, respectively. Furthermore, the detecting speed of the LSI-YOLOv8 model is 73.2 f/s, which is notably faster by 15.5 f/s, 3.5 f/s, 5.1 f/s, 5.9 f/s, and 4.3 f/s respectively, compared to Mask-RCNN, YOLOv5, YOLOv7, YOLOv8, and YOLOv9. The proposed method provides valuable insights for advancing landslide identification methodologies. The focus of subsequent work will be more on multi-source data fusion, such as Digital Elevation Model (DEM) data and Synthetic Aperture Radar (SAR) data to remove the influence of some background information to improve the accuracy of landslide identification.

INDEX TERMS Landslide identification, remote sensing imagery, deep learning, LSI-YOLOv8, DWR segmentation, EIoU.

I. INTRODUCTION

Over the past few decades, there has been a significant increase in the incidence of geological disasters worldwide.

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Amitrano.

This increase can be attributed to frequent human activity, natural actions like earthquakes and river erosion, and climate change, such as extremely heavy rainfall, especially in mountainous regions [1]. In China, landslides are the most common geological disasters, posing a threat to the safety, survival, and sustainable development of the national economy.

With the rapid advancements of modern technology, these methods such as GNSS [2], [3], InSAR [4], [5], [6], LiDAR [7], [8] and high resolution optical remote satellites or drones [9], [10], have demonstrated promising results in monitoring and interpreting landslides and geological disasters within the local or small regions. However, it is very important to identify landslides quickly and accurately, determine their geographical location and evaluate their impacts from large regions after a heavy disaster. This information is vital for prompt local emergency rescue operations and minimizing further losses. Furthermore, the rapid and accurate identification of large-scale landslides from satellite or drone remote sensing imagery following earthquakes or heavy rainfall enables real-time monitoring and management in regions affected by severe weather or complex mountainous terrain. This technology also aids in statistical analysis of landslide disasters, allows for quick updates to landslide inventories, and enhances the precision and effectiveness of disaster relief efforts for affected communities.

Traditionally, field investigation is a commonly used and reliable method for exploring landslides. However, it has several limitations including a limited field of view, low efficiency, high cost, and high risk. With the advancement of remote sensing technology, remote sensing images became available and have become a valuable tool for interpreting geological disasters like landslides [11]. Currently, there are three main methods for identifying landslides from remote sensing images: visual interpretation [12], pixel-based analysis [13] and object-oriented analysis [14]. Visual interpretation, although prone to being influenced by expert experience and personal subjective impressions, is often used. However, it suffers from drawbacks such as low recognition efficiency, lengthy duration, weak pertinence, and the potential for omitting important information when dealing with large-scale areas [15]. One obstacle to the practical application of pixel-based classification is the unsatisfactory results obtained from pixel-by-pixel analysis. This can often be attributed to the fact that the image's geometric and contextual information is largely ignored [16]. In the process of object-oriented image segmentation, the segmented landslide objects may include other features. The threshold settings of the extraction rules involve subjective factors. Additionally, the segmentation process of object-oriented image method and the subsequent calculation of the topological relationships between the objects can consume significant computer memory and time [17]. However, landslide identification as an image processing problem has been extensively studied using statistical and machine learning methods, resulting in relatively satisfactory outcomes [18].

Considering landslide identification as an image processing problem, statistical and machine learning methods have been widely utilized and have shown promising results [19]. Various machine learning algorithms such as Artificial Neural Network (ANN) [20], Support Vector Machine (SVM) [21], Random Forest (RF) [22], eXtreme Gradient Boosting (XGBoost) [23], among others, have been employed. However, these traditional machine learning methods necessitate manual construction and selection of feature layers during landslide feature processing, posing a challenging task. Moreover, the selected features may not generalize well beyond the test area [24]. In recent years, there has been significant advancement in target identification through the adoption of deep learning technology. This transition has shifted from conventional feature-based manual algorithms to automatic identification methods based on deep neural networks [25]. Deep learning algorithms can be broadly categorized into two methods: (1) two-stage methods, which involve identifying multiple target candidate areas in the image and then classifying and regressing the boundaries of each candidate area. Notable two-stage algorithms include region based Fast R-CNN, region-based fully convolutional network (R-FCN), and mask region-based Mask R-CNN. (2) Single-stage algorithms directly locate and classify all detected targets across the entire image without explicitly marking candidate areas. Well- known algorithms in this category include the You Only Look Once (YOLO) [26] series, Single Shot MultiBox Detector (SSD) [27], and RetinaNet [28]. Both methods offer distinct advantages, with the single stage algorithm being faster and the two-stage algorithm demonstrating higher accuracy.

Meanwhile, recent advancements in deep learning techniques have shown significant progress in remote sensing image processing, leading researchers to explore the application of deep learning for landslide identification. For instance, Yu et al. [29] proposed an intelligent landslide identification method that combines a deep convolutional neural network (CNN) and an improved region growing algorithm (RSG_R), achieving the identification accuracy of 94.74%. Wu et al. [30] combined SMOTE with convolutional neural networks and reported an overall accuracy of 84.59% in landslide susceptibility assessment. Xiaoting et al. [31] used the Faster R-CNN algorithm to build an automatic landslide hazard identification model based on the Gaofen-1 image. The final test results showed that the AP value reached 92.42%. Another approach by Jiang et al. [32] used the Mask R-CNN method to simulate difficult samples under complex backgrounds, and the model identification accuracy reached 92.31%. However, it should be noted that the CNN model may take longer to process for large data sets and may require further structural adjustments. Pang et al. [33] developed a neural network model using the YOLOv3 algorithm specifically for landslide identification from remote sensing satellite image features, achieving an average accuracy of 0.98 for co-seismic landslide identification. Wang et al. [34] improved upon the YOLOv5 framework by enhancing the feature splicing method and introducing adaptive spatial feature fusion (ASFF) and the convolutional block attention module (CBAM), resulting in a 1.64% improvement in model performance. Overall, these studies and their methods have

demonstrated stable and reliable performance in landslide identification tasks.

However, despite the contributions of the aforementioned study to landslide identification tasks using remote sensing images, deep learning methods for landslide identification till face challenges. These challenges include missed detection and incorrect identifications, which highlight the need for further improvements in accurately extracting landslide spatial information. While single stage methods offer slightly lower accuracy compared to two-stage approaches, they tend to have faster processing speeds. This speed advantage allows for the separation of specific categories and estimation of boundaries. Additionally, in the case of massive landslides triggered by heavy rainfalls or major seismic events, it is crucial to quickly and accurately detect and map regional landslides for disaster relief and statistical analysis during emergencies after a heavy disaster. In this context, single-stage detectors with their significantly better identification speeds and efficiency are more suitable for real-time identification in emergency scenarios. As a classic single-stage identification algorithm, the YOLO algorithm has evolved to YOLOv8, which is also the stable and offers significant improvements in both identification accuracy and speed [35]. Moreover, there have been limited studies that have utilized YOLOv8 for landslide identification. Therefore, we have decided to optimize the model using the YOLOv8 framework to further enhance the accuracy of the algorithm.

In response to the aforementioned issues, an identification model called LSI-YOLOv8 is introduced for the rapid identification of landslides using post disaster remote sensing images. This model enhances both the accuracy and efficiency of landslide identification. The main contributions of this paper are considered as follows:

• The DWR Segmentation (DWRSeg) [36] module was introduced into the C2f module to reduce the number of parameters in the model and enhance its generalization ability. The inclusion of C2f-DWR enables better adaptation to diverse input data, improving the model's expression ability and performance. This allows the network to flexibly adapt to features of varying scales, accurately identify and segment objects in images, and optimize the identification accuracy of the model.

• To enhance the convergence speed and positioning accuracy of the algorithm, we will utilize the Efficient Intersection over Union (EIoU) [37] method to replace the original bounding box regression loss in the model. The EIoU loss directly minimizes the disparity in width and height between the target box and anchor box, leading to faster convergence speed and improved positioning results.

• Extensive experiments were conducted on real remote sensing datasets to evaluate the proposed method. The experimental results confirm its superiority. Additionally, detailed ablation experiments were performed to analyze the impact of different settings of loss functions and parameter values.

II. METHODOLOGY

Real time object identification has become a crucial element in numerous applications across various fields, including autonomous vehicles, robotics, video surveillance, and augmented reality [38]. Among the different object identification algorithms, the YOLO framework has gained prominence due to its impressive combination of speed and accuracy. It allows for the quick and dependable identification of objects in images.

A. BASIC YOLOV8 MODEL

The YOLO model has made significant advancements in the field of computer vision. Consequently, researchers have continuously improved and expanded upon this method by introducing various classical models. Among the exceptional models in the YOLO series (such as YOLOv5 and YOLOv7), YOLOv8 stands out as an advanced and cutting-edge model that offers superior identification accuracy and speed [39]. The YOLOv8's basic structure, which consists of an input segment, a backbone, a neck, and an output segment. The input segment performs mosaic data augmentation, adaptive anchor calculation, and adaptive grayscale padding on the input image. The backbone network and neck module are the central structures in the YOLOv8 network. The input image undergoes processing by multiple Conv and C2f modules to extract feature maps at different scales. The C2f module, an improved version of the original C3 module, functions as the primary residual learning module. It incorporates the benefits of the ELAN structure in YOLOv7 [40], reducing one standard convolutional layer and making full use of the Bottleneck module to enhance the gradient branch. This approach not only preserves the lightweight characteristics but also captures more abundant gradient flow information. The output feature maps are then processed by the SPPF module, which employs pooling with varying kernel sizes to combine the feature maps before passing the results to the neck layer.

B. IMPROVED YOLOV8 MODEL: LSI-YOLOV8

The YOLOv8 is a state-of-the-art object identification model that considers the multiscale nature of objects. It incorporates three scale-identification layers to accommodate objects of different scales. However, the data obtained through remote sensing for landslide identification poses challenges due to complex background and a heavy workload for identifying landslides, especially in drone imagery. Consequently, the identification structure of YOLOv8 fails to meet the demands for quick and precise landslide identification. To address these challenges, this paper proposes a solution that builds upon YOLOv8 as the base model and optimizes it through improvements in the loss function, attention mechanism, and multi-scale feature fusion. To achieve rapid and high-accuracy recognition, improvements have been made to the YOLOv8 model. The specific enhancements are as follows: Firstly, considering the varying shapes and sizes of targets

in remote sensing images, the aspect ratios of bounding boxes often differ significantly. To address this, an additional penalty term called EIoU is introduced based on Complete Intersection over Union (CIoU) [41]. The EIoU separates the aspect ratio image factors of the predicted frame and the real frame, and calculates the length and width information of each frame individually. This solves the issue of proportional changes in aspect ratio that CIoU fails to handle. In this article, CIoU is replaced with EIoU. Secondly, the background of remote sensing images is often complex, and the identification target occupies only a small area in the image. This leads to a large proportion of redundant features during image iteration. To enhance the selectivity of the model, considering that the dataset consists of low-resolution and high-complexity images, we propose the introduction of the DWRSeg module in C2f. The DWRSeg module utilizes an adaptive weight allocation strategy to dynamically adjust the weights of the convolution kernel based on the characteristics of the input data. This approach improves the model's performance by making it more flexible in adapting to different input data. Lastly, a landslide identification method called LSIYOLOv8 is designed to achieve a balance between lightweight implementation and algorithmic performance. The overall enhancements shown as Figure 1 in red rectangle.

FIGURE 1. The network structure of LSI-YOLOv8.

1) EIOU LOSS FUNCTION

The Yolov8 utilizes the CIoU as the bounding box regression function, which is an indicator employed to calculate the similarity between bounding boxes in target identification. CIoU is an improvement over Intersection over Union (IoU) [42], which is traditionally used to measure the degree of overlap between predicted boxes and ground-truth boxes in object identification tasks. However, IoU only considers the position information of the box and neglects the size and aspect ratio of the box. Consequently, there are cases where IoU fails to accurately reflect the similarity between two boxes. On the other hand, CIoU incorporates the center point distance,

aspect ratio difference, and box size into its calculation. Given a predicted box B and a target box B^{gt}, b and b^{gt} denote the central points of B and B^{gt} respectively. The CIoU loss function is defined as follows.

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (1)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (2)$$

$$v = \frac{4}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h})^2 \quad (3)$$

where $\rho(\cdot) = \| b - b^{gt} \|^2$ indicates the Euclidean distance. c is the diagonal length of the smallest enclosing box covering the two boxes. The gradients of v, w, r, t, and h, are calculated as follows.

$$\frac{\partial v}{\partial w} = \frac{8}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}) * \frac{h}{w^2 + h^2} \quad (4)$$

$$\frac{\partial v}{\partial h} = -\frac{8}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}) * \frac{w}{w^2 + h^2} \quad (5)$$

Since v only reflects the discrepancy of aspect ratio, rather than the real relationship between the predicted boxes of the real box, CIoU may optimizes the similarity in an unreasonable way. To address this problem, we further change CIoU to a more efficient version of IoU, namely EIoU, which is defined as follows.

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} \quad (6)$$

$$L_{EIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{(w^c)^2 + (h^c)^2} + \frac{\rho^2(w, w^{gt})}{(w^c)^2}$$
$$+ \frac{\rho^2(h, h^{gt})}{(h^c)^2} \quad (7)$$

We divide the loss function into three parts: the IoU loss L_{IoU}, the distance loss L_{dis} and the aspect loss L_{asp} (Overlapping area, center point example, aspect ratio). The height-width loss directly minimizes the height and width differences between the predicted target bounding box and the real bounding box, enabling faster convergence speed and more enhanced localization results.

2) DWR SEGMENTATION

The backbone network and neck architecture of YOLOv8 may have been influenced by the design concept of Efficient Layer Aggregation Networks (ELAN) from YOLOv7. Additionally, the C3 structure of YOLOv5 has been replaced by a C2f structure to enhance gradient flow, with adjustments made to channel numbers for different scale models. By incorporating additional branches and layer links, YOLOv8 can capture more comprehensive gradient flow information without compromising its lightweight nature. The C2f module is a specialized convolution module that merges feature maps of varying scales, thereby enhancing the model's receptive field and identification accuracy. In YOLOv8, the C2f module has been used in the identification head to merge feature maps of different scales,
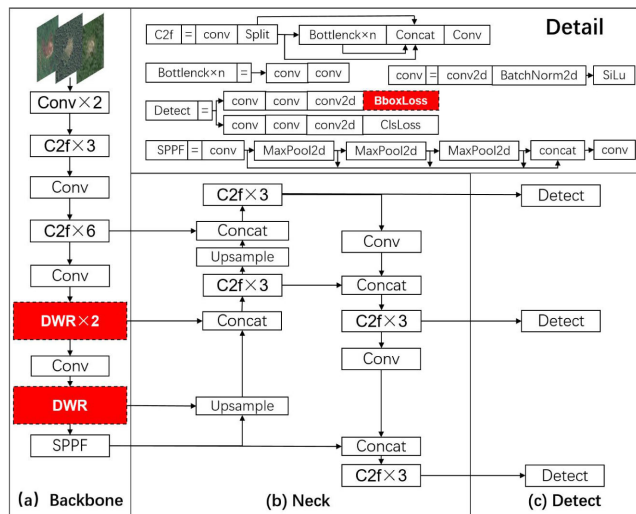
enhancing identification accuracy. However, the C2f module (as shown in the Figure 2) also introduces additional computational complexity to the model, leading to increased training and time costs. Additionally, it also increases the number of parameters in the YOLOv8 model. Therefore, we replace some c2f modules with DWR modules to reduce the overall complexity and parameter count of the model.
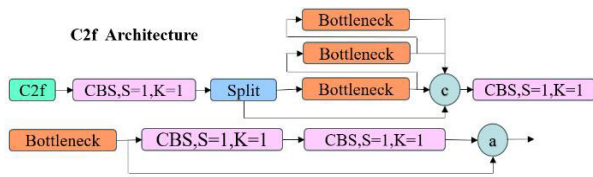

FIGURE 2. C2f module structure.

As known, the attention mechanism in image processing is designed to highlight key information by assigning higher importance to specific areas. This facilitates the fusion and interaction of information across different levels or scales, selectively boosting feature channels that are rich in target information. Ultimately, this enables the model to prioritize important areas and minimize the influence of background information [43].

Here, the DWRSeg network utilizes an attention mechanism, incorporating a novel Dilation-wise Residual (DWR) module and a Simple Inverted Residual (SIR) module for both high and low levels. The overall architecture, depicted in Figure 3, is characterized by its simplicity and efficiency. This network follows an encoder-decoder setup, with the encoder comprising four stages: the stem block, the low stage of the SIR modules, and two high stages of the DWR module. The DWR module is structured in a residual fashion, employing a two-step method within the residual to efficiently gather multi-scale contextual information and fuse feature maps with multi-scale receptive fields. By decomposing the previous single-step multi-scale contextual information acquisition method into a two-step approach, the acquisition difficulty is reduced. The Simple Inverted Residual (SIR) module, derived from the DWR module, is tailored to maintain high feature extraction efficiency by meeting the requirement for a smaller receptive field size in the lower stage. The role of multi-rate deep expansion convolution in feature extraction is from difficult to simple: from obtaining as much complex contextual information as possible from complexly expressed feature maps to performing simple morphological filtering of feature maps with the desired expansion rate for each succinct expression. Simple region featural maps form clear, simplified learning expansion. The goal of deep convolution makes the learning process more organized.

III. RESULTS AND ANALYSIS
A. EXPERIMENTAL ENVIRONMENT
To assess the performance enhancements, we utilized PyTorch as the deep learning framework and employed
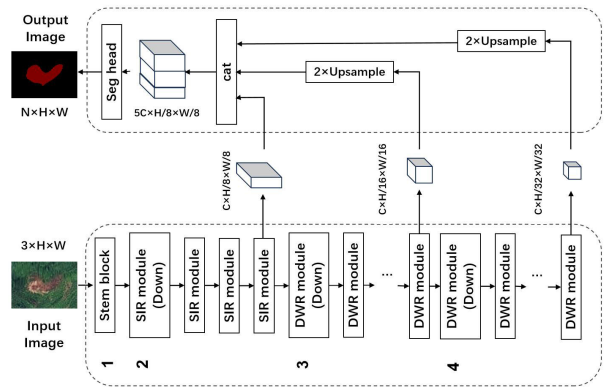

FIGURE 3. DWR segmentation model structure [36].

YOLOv8s as the baseline network model. The hardware platform and environmental parameters utilized during the experimental training phase are detailed in Table 1.

TABLE 1. Configuration of experimental environment.

Parameters	Configuration
Operating systems	Windows10
CPU	Inter Xeon (ZQ)E5-2686
GPU	NVIDIA GeForceRTX4060Ti
GPU memory size	32G
Deep learning architecture	torch-2.1.2+cu118-cp39

Consistent hyperparameters were also applied throughout the training process across all experiments. Table 2 displays the precise hyperparameters employed during the training process. The practice code for YOLOv8 can be found at https://gitee.com/.

TABLE 2. Configuration of experimental training process.

Hyper parameters	Value
Epoch	200
Momentum	0.937
Initial learning rate	0.01
GPU mem	4.30
GFLOPs	8.2
Input image size	256×256
Initial learning rate	0.01
Weight decay	0.0005
Optimizer	SGD
Data enhancement strategy	Mosaic

B. LANDSLIDE IMAGE DATAEST
In this experiment, the remote image of landslides primarily originates from the Bijie Area in the northwest of Guizhou Province, China. Bijie City is situated in the transitional slope zone between the Qinghai-Tibet Plateau and the eastern highlands, making it one of the regions most affected by landslides in China. These landslides frequently occur in mountainous areas abundant in vegetation, road slopes susceptible to landslides, and locations with significant human

engineering activities, often triggered by heavy rainfall. In the transition from spring to summer each year, heavy and extremely heavy rains in this region significantly increase, leading to a surge in mountain landslides and other geological disasters. The government annually invests substantial manpower and material resources in conducting surveys and censuses of geological hazards in the area to compile a comprehensive database. Traditional manual geological hazard surveys are time-consuming and labor-intensive. There is an urgent need to leverage satellite remote sensing and other technologies to detect landslides and other geological hazards in large-scale remote sensing images. This will enhance the efficiency of geological hazard surveys and contribute to the digital transformation of the process.

Currently, the Bijie Landslide Dataset [19] is open-source into research community, and now has been freely accessed and downloaded as following link (http://study.rsgis.whu.edu.cn/pages/download/). The data set consists of 770 positive sample images containing landslides and 2003 negative sample images excluding landslides collected by the TripleSat series satellites from May to August 2018. The ground resolution of the image is 0.8 meters. During experimental preprocessing, the slice size is set to 256×256 and the label name is "landslide". Considering practical factors, image augmentation (rotation, flipping, brightness enhancement, contrast enhancement, introduction of Poisson noise, Gaussian noise, salt and pepper noise) is applied to these images.

Conventionally, to accurately annotate landslide remote sensing image samples and meet the training, testing, and validation requirements of deep learning models, the research dataset was divided into a training set, a test set, and a validation set in a ratio of 8:1:1. Figure 4 and 5 show examples of images from the training and validating set, respectively.
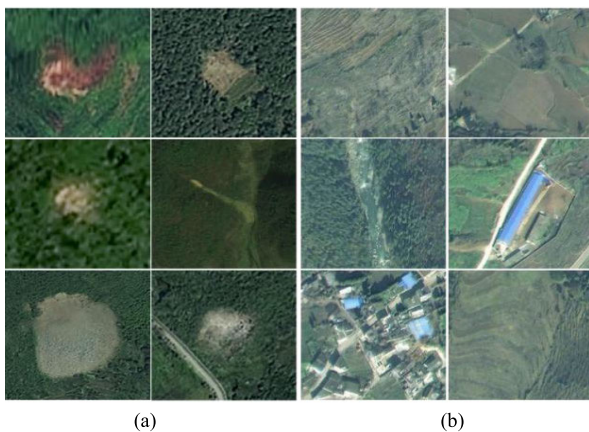

FIGURE 4. Landslide and non-landslide images in train dataset. (a) The images of landslide;(b) The images of non-landslide.

C. EVALUATION METRICS

To test the identification performance of our improved LSI-YOLOv8 model, we use precision, recall, F1-score, mAP@0.5 and mAP@0.5:0.95 as evaluation metrics. The calculation formula for the various evaluation are as
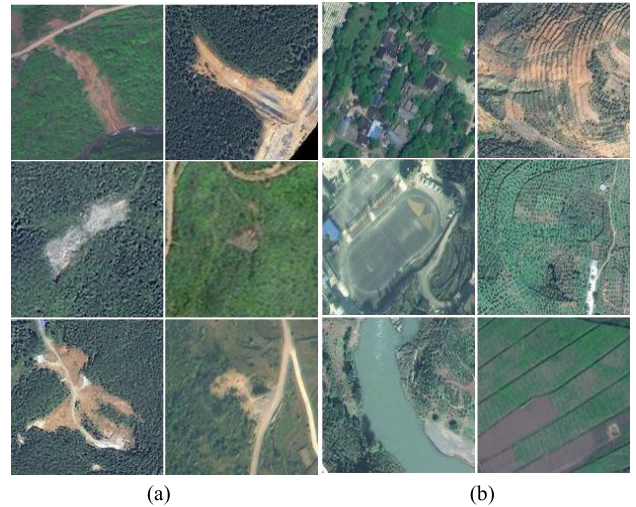

FIGURE 5. Landslide and non-landslide images in validated dataset. (a) Landslide images; (b) Non – landslide images.

follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

$$\text{F1} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

The F1 value is calculated as the arithmetic mean divided by the geometric mean, and a higher value is considered better. By considering the formulas of Precision and Recall, it can be observed that when the F1 value is small, there is a relatively higher increase in True Positive and a decrease in false, resulting in an increase in both Precision and Recall. Therefore, F1 effectively balances both Precision and Recall.

$$\text{AP} = \int_0^1 \text{Precision (Recall)} \, d \, (\text{Recall}) \tag{11}$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} = \text{AP}_i \tag{12}$$

The AP_i in Equation (12) denotes the AP value with category index value i, and N denotes the number of categories of the samples in the training dataset (in this paper, N is 1). mAP@0.5 denotes the average accuracy when the IoU of the identification model is set to 0.5, and mAP@0.5:0.95 denotes the average accuracy when the IoU of the identification model is set from 0.5 to 0.95 (with values taken at intervals of 0.05).

D. EXPERIMENTAL RESULTS

1) EIOU+DWR SEGAMENTATION

To validate the effectiveness of introducing EIoU, we maintained all other training conditions unchanged. We performed comparative experiments on YOLOv8 using EIoU along with other mainstream loss functions (CIoU, DIoU [41], SIoU [44], GIoU [45]). The experimental results are presented in Table 3.

TABLE 3. Comparison of identification results for different loss functions introduced by YOLOv8.

Methods	Precision	Recall	mAP@0.5	mAP@0.5:0.95
CIoU	81.0%	74.8%	82.3%	48.5%
DIoU	81.2%	75.9%	81.7%	49.5%
SIoU	83.2%	77.8%	84.0%	50.3%
GIoU	84.7%	71.5%	80.7%	49.3%
EIoU	85.3%	70.8%	81.3%	48.1%

According to the experimental results in Table 3, when utilizing the EIoU as the bounding box regression loss, the YOLOv8 model demonstrates a significantly high accuracy in identifying the target, thereby enhancing the positioning accuracy of YOLOv8. However, a challenge arises in terms of the low recall rate, and lower mAP@0.5 value. To address this issue, the C2f-DWR attention mechanism is introduced. This mechanism ensures a high accuracy in the recognition results, enabling the network to effectively adapt to features of varying scales and achieve more precise object segmentation in images.

To evaluate the performance of the YOLOv8 with DWRSeg module, we maintained consistent training conditions while comparing it to the original YOLOv8 model. We conducted comparative experiments on loss functions (CIoU, DIoU, SIoU, GIoU, EIoU) in the improved YOLOv8 model with different improvement strategies. Comparative results are detailed in Table 4. The findings indicate that the accuracy of all loss functions with the attention mechanism has also slightly improved. Particularly, when integrating EIoU with DWRSeg, the model's identification accuracy significantly increased by 6.1%. Moreover, the recall increased by 4.1% and the mAP@0.5 showed an improvement of 7.3%. Consequently, the model achieved superior identification performance.

TABLE 4. Comparison of identification results for different loss functions introduced by the improved YOLOv8.

Methods/Strategies	Precision	Recall	mAP@0.5	mAP@0.5:0.95
CIoU+ DWRSeg	81.1%	75.4%	82.7%	49.7%
DIoU+ DWRSeg	82.2%	81.2%	83.8%	50.6%
SIoU+ DWRSeg	84.4%	76.1%	84.9%	50.7%
GIoU+ DWRSeg	84.1%	77.5%	84.3%	50.3%
EIoU+ DWRSeg (ours)	91.5%	74.9%	88.6%	53.3%

By comprehensively and intuitively analyzing the experimental results presented in Figure 6, the performance and comparison results of each improved method with YOLOv8 can be summarized as follows. The EIoU method achieves a higher accuracy of 85.3% in the bounding box function results under the YOLOv8 model, albeit with a relatively low recall rate of 70.8%, resulting in an average accuracy of 81.3% for the EIoU model. Conversely, the SIoU method
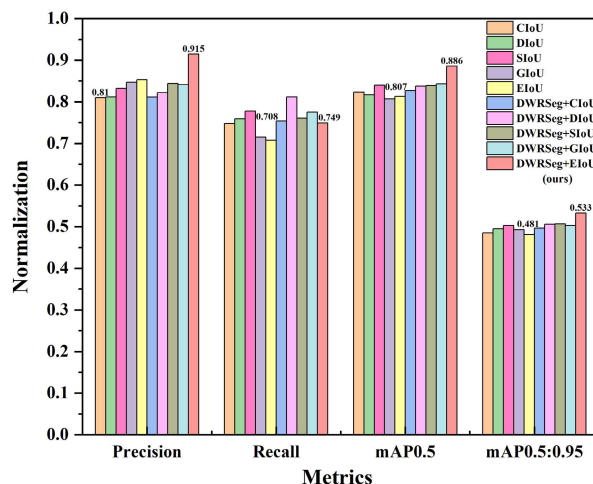

FIGURE 6. The normalization effect of experimental indicators.

exhibits a lower accuracy of 83.2%, while displaying a higher recall rate at 77.8%. Overall, the SIoU method attains the highest average precision of 84.0%, surpassing the EIoU method within the YOLOv8 model. However, considering the improved method that incorporating the C2f-DWRSeg attention mechanism, the performance of each combination for the YOLOv8 model has slightly increased. Notably, the EIoU+ DWRSeg method excels with an impressive accuracy of 91.5%. Despite its generally lower recall rate of 74.9%, the average accuracy (mAP@0.5) of 88.6% still outperforms the average identification accuracy of alternative bounding box loss functions.

2) DIFFERENT ATTENTION MECHANISM

In order to further evaluate the effectiveness of integrating the DWRSeg module while maintaining the same Loss Function of EIoU, three improvement methods from different attention mechanism were selected for comparison: Asymptotic Feature Pyramid Network (AFPN) [46], Deformable Attention (D-Attention) [47], and Diverse Branch Block (DBB) [48] were tested on YOLOv8. The metrics involved in the table, Frames Per Second (FPS): the number of frames transmitted per second, that is, the network can detect how many images per second. Giga Floating-point Operations PER Second (GFLOPs): the number of billion floating-point operations per second. The results, presented in Table 5, show that the identification accuracy with the DWRSeg module outperforms the other enhancement modules. The DWRSeg module demonstrates an accuracy improvement of 11.1%, 7.2%, and 7.3% over other module, showcasing its superior recognition capabilities. In the comprehensive accuracy evaluation, the ranking is as follows: DWRSeg > D-Attention > DBB > AFPN.

3) ABLATION EXPERIMENT

To assess the impact of various enhancement strategies outlined in this research, ablation experiments were conducted

TABLE 5. Comparison results from different attention by the improved YOLOv8.

Modules	AFPN	D-Attention	DBB	DWRSeg
Precision/%	80.4	84.3	84.2	91.5
Recall/%	74.4	78.4	76.6	74.9
mAP@0.5/%	81.2	85.4	84.3	88.6
mAP@0.5:0.95/%	48.2	51.8	51.2	53.3

on the baseline model utilizing the same of Bijie Landslide Datasets. The results of these experiments are detailed in Table 6. The study primarily focused on improving the baseline model (YOLOv8) by incorporating better modules such as DWRSeg and D-Attention into the C2f module and replacing CIOU with EIoU as the bounding box loss function within the final loss function. The performance improvements of each individual model component were analyzed, including the baseline model, the enhanced DWRSeg model (backbone), the enhanced EIoU model (loss function), and the combined EIoU+ DWRSeg model (backbone + loss function). Three evaluation metrics, including the identification average accuracy of mAP@0.5, F1-Score, and FPS (frames per second, which means the higher, the better), were quantitatively assessed to observe the impact of these enhancements.

TABLE 6. Ablation experiment of different components in LSI-YOLOV8.

CIoU	EIoU	D-Attention	mAP@0.5	F1–Score	FPS(f/s)
√	-	-	82.3%	78.4%	67.3
-	√	-	81.3%	78.5%	68.1
√	-	√	85.4%	78.9%	71.9
-	√	√	87.1%	83.1%	72.6

(a)

CIoU	EIoU	DWRSeg	mAP@0.5	F1–Score	FPS(f/s)
√	-	√	82.7%	78.8%	72.7
-	√	√	88.6%	84.2%	73.2

(b)

Table 6 illustrates that the YOLOv8 model's overall accuracy is improved by integrating both the EIoU and DWRSeg modules. The ablation experiment emphasizes the importance and effectiveness of incorporating each module. It is clear that EIoU and DWRSeg modules outperformed other components. The combination of EIoU and DWRSeg modules maintain a high recognition speed (FPS:73.2 f/s) while achieving the highest target recognition accuracy (mAP@0.5: 88.6%).

Figure 7 also demonstrates the comparison of accuracy between the original and improved models. The analysis of Figure 7 indicates that our LSI-YOLOv8s model proposed in this study exhibits significantly higher values for

mAP@0.5 and mAP@0.5:0.95 in comparison to the original YOLOv8s model.

(a)

(b)
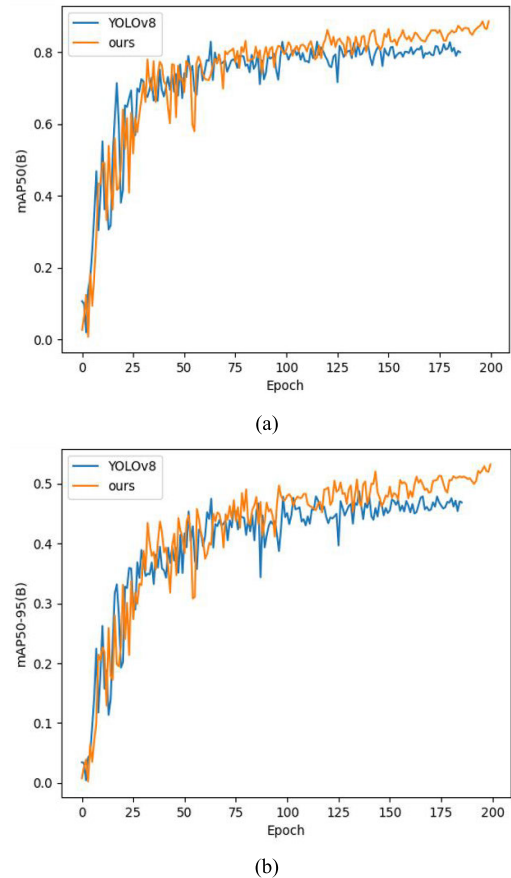
FIGURE 7. A comparison was conducted to evaluate the mAP values of the LSI-YOLOv8s model against the original. (a) mAP@0.5; (b) mAP@0.5:0.95.

4) VISUAL INTERPRETABILITY

Deep learning models are commonly considered black-box models, making it difficult to interpret their decision-making and reasoning processes [35]. Hence, it is crucial to thoroughly explore the interpretability of deep learning models to gain a clear understanding of their performance. In this experiment, we chose to verify the LSI-YOLOv8 model and the YOLOv8 model by examining their confusion matrices.

The confusion matrices of the two models are shown in Figure 8. The confusion matrices indicate that both models exhibit high rates of false negatives (i.e., misidentify targets as background categories) and false positives. The original YOLOv8 exhibited subpar identification performance for landslide identification, achieving an accuracy rate of only 82.0% and higher misidentification rate of 18.0%. In contrast, the LSI-YOLOv8 significantly enhanced the identification accuracy of landslide, with improvement of 3.0% and reduction of misidentification rate 3.0%.

To explain the difference more intuitively, XGrad-CAM [49] was often utilized to generate heat maps from various models for visual displays. Heat maps are visual
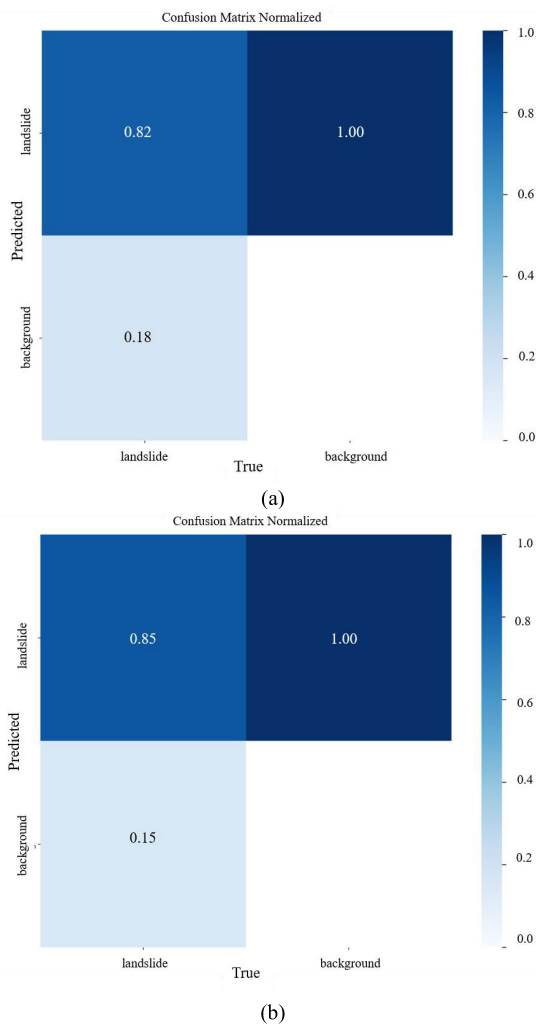
(a)

(b)

FIGURE 8. Confusion matrix of the YOLOv8 model and LSI-YOLOv8 model. (a) Confusion matrix of YOLOv8; (b) Confusion matrix of LSI-YOLOv8.

FIGURE 9. XGrad-CAM rendering of different model validation datasets (a) Origin images; (b) YOLOv8; (c)LSI-YOLOv8.

IV. DISCUSSION

To test the robustness of the proposed LSI-YOLOv8 model, we aim to compare the accuracy of algorithms in different models to determine their progress and complexity. Furthermore, we will assess how well these models can generalize by testing their performance on diverse datasets. This thorough evaluation aims to confirm the effectiveness and versatility of the proposed model, establishing a strong basis for future enhancements and applications in the field.

A. COMPARISON OF VARIOUS MODELS

In order to enhance the demonstration of the LSI-YOLO model, this experiment selected YOLOv5, YOLOv7, and different versions of YOLOv8 to compare parameters, training time, and accuracy. For instance, the four versions of YOLOv8 (YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, YOLOv8x) vary in parameter number from small to large. The comparison results are illustrated in Fig. 10, indicating that the 'n' model in the YOLO series has a relatively small number of parameters, ensuring high accuracy in detection. All four versions of the LSI-YOLOv8 model used in this experiment exhibit efficient landslide identification and detection capabilities, yielding rapid results for detection.

To further validate the effectiveness of the proposed LSI-YOLOv8, this study conducts comparative experiments with prominent other state-of-the-art models such as Mask-RCNN [50], YOLOv5, YOLOv7, YOLOv8, and YOLOv9 [51], known for their superior accuracy in other domains. The outcomes of these comparative experiments are detailed in Table 7. Mask R-CNN is a simple and effective model with two-stage identification algorithm that extends Faster R-CNN to detect objects and generate masks for each instance. The YOLOv5 is a single-stage object

representations that highlight the areas in the feature map that the model emphasizes. The gradient value is obtained by back-propagating the model's output category confidence through XGrad-CAM. In the heat map, pixels with higher gradients in the feature map are depicted as darker shades of red, indicating a stronger focus. Conversely, pixels with lower gradients are represented by darker shades of blue.

The comparative experimental results for displays are depicted in Figure 9. The findings suggest that the original YOLOv8 model has limited attention when faced with complex backgrounds and struggles to accurately identify landslides. In contrast, our proposed LSI-YOLOv8 model demonstrates a superior ability to filter out background noise and gives priority to detecting landslides. The model's focus is primarily on the central point of the object, leading to more accurate predictions of bounding boxes and a more thorough identification of landslides. This methodology effectively reduces instances of missed or incorrect identifications, ultimately improving the overall performance of the model in landslide detection.
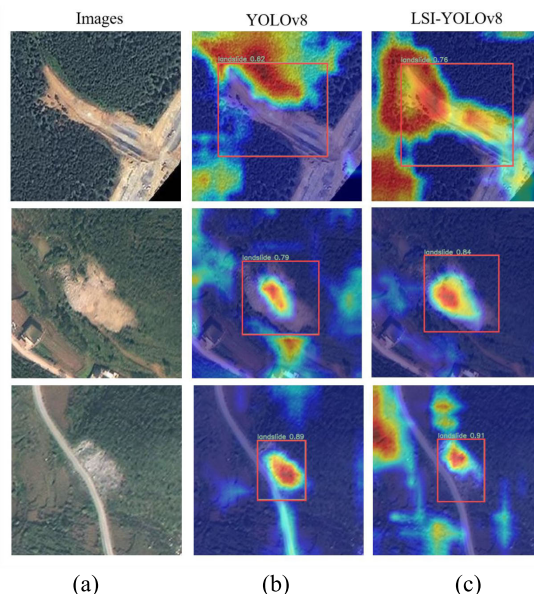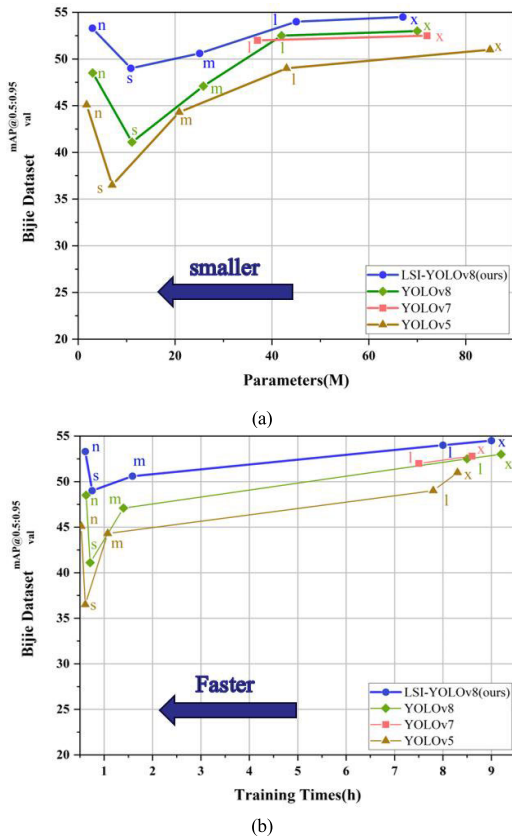
(a)

(b)

FIGURE 10. Comparison between different YOLO models within the validated dataset. (a) Parameters numbers; (b) Training Times.

TABLE 7. Comparison of identification results for different models based on Bijie landslide datasets.

Models	Mask-RCNN	YOL Ov5	YOL Ov7	YOLO v8n	YOL Ov9c	LSI-YOLO v8(Ours)
Precision/%	94.6	80.9	89.1	81.0	80.2	91.5
Recall/%	65.9	73.1	73.7	74.8	78.2	74.9
mAP@0.5/%	79.6	80.9	84.6	82.3	85.7	88.6
mAP@0.5:0.95/%	50.2	49.7	50.6	48.5	53.0	53.3
Training Time/h	5.31	0.60	2.10	0.68	4.11	0.61
FPS(f/s)	57.7	69.7	68.1	67.3	68.9	73.2

identification algorithm that segments the image into multiple grids, with each grid responsible for predicting one or more objects. The network structure of YOLOv5 is fully convolutional. YOLOv7 employs a unique design that transforms the target identification task into a single convolution operation, enabling effective identification of diverse targets, particularly small ones. On the other hand, YOLOv9, as a latest version of YOLO series model, incorporates advanced techniques like the feature pyramid network and attention mechanism, enhancing the accuracy of target identification and demonstrating superior performance in detecting dense targets.

According to the findings presented in Table 7, it can be observed that Mask R-CNN exhibits the highest detection accuracy and is capable of accurately segmenting the

shape and size of landslides even in complex backgrounds. However, it should be noted that the training time for Mask R-CNN is relatively long and the detection time with lowest FPS means the detection is not fast, which may hinder rapid post-disaster analysis. The YOLOv5 demonstrates the fastest training speed but lower average accuracy compared to other YOLO models. Among the original YOLO models, YOLOv7 excels in identification accuracy, although its recall rate, mAP@0.5, and mAP@0.5:0.95 metrics are surpassed by YOLOv9. Despite this, the overall accuracy of YOLOv7 exceeds that of YOLOv8. While YOLOv8 does not exhibit clear superiority in the model, its identification time significantly outperforms other YOLO models. This study focuses on enhancing overall identification performance using YOLOv8 and introduces the LSI-YOLOv8 model. The accuracy of LSI-YOLOv8 significantly outperforms other Mask-RCNN, YOLOv5, YOLOv7, YOLOv8, and YOLOv9 model, with mAP@0.5 being 9.0%,7.7%, 4.0%, 6.3%, and 2.9% higher than other models, while mAP@0.5:0.95 is 3.1%, 3.6%, 2.7%, 4.8%, and 0.3% higher than others, respectively. Furthermore, the detection speed of the LSI-YOLOv8 model is 73.2 f/s, which is notably faster by 15.5 f/s, 3.5 f/s, 5.1 f/s, 5.9 f/s, and 4.3 f/s respectively, compared to Mask-RCNN, YOLOv5, YOLOv7, YOLOv8, and YOLOv9.

Based on the results presented in Table 7, it is evident intuitively from Figure11 that both YOLOv5 (Fig.11c) and YOLOv8 (Fig.11e) are prone to missing identifications in complex backgrounds, with YOLOv5 exhibiting a higher rate of missed identifications. In contrast, YOLOv9 (Fig.11f) demonstrates the ability to accurately identify the shape and size of landslides even in challenging backgrounds. However, the identification time of YOLOv9 is relatively long, potentially impeding timely post-disaster analysis. While YOLOv8 shows improved identification and recognition compared to YOLOv7 (Fig.11d), it still faces challenges in landslide identification in complex backgrounds, leading to unclear localization and false identifications in landslide identification scenarios. On the other hand, LSI-YOLOv8 (Fig.11g) achieves higher identification accuracy than other YOLO models, albeit with a noticeable issue of identification omissions.

B. COMPARISON OF VARIOUS DATASETS

To further validate the robustness of our approach, we conducted generalization experiments using the original YOLOv8 and the improved LSI-YOLOv8 on various datasets comprising landslide images from open-source community in recent years. Our experiment involved four well-known landslide including Bijie Datasets [19], two datasets (Moxi Town and Jiuzhai Valley) of CAS Landslide Dataset [52] and Southwest landslide dataset [53] (https://github.com/YhQIAO/LandSlide_Detection_Faster-RCNN). These four datasets are detailed in the Table 8. These experimental results effectively demonstrate the generalization ability of the improved model.
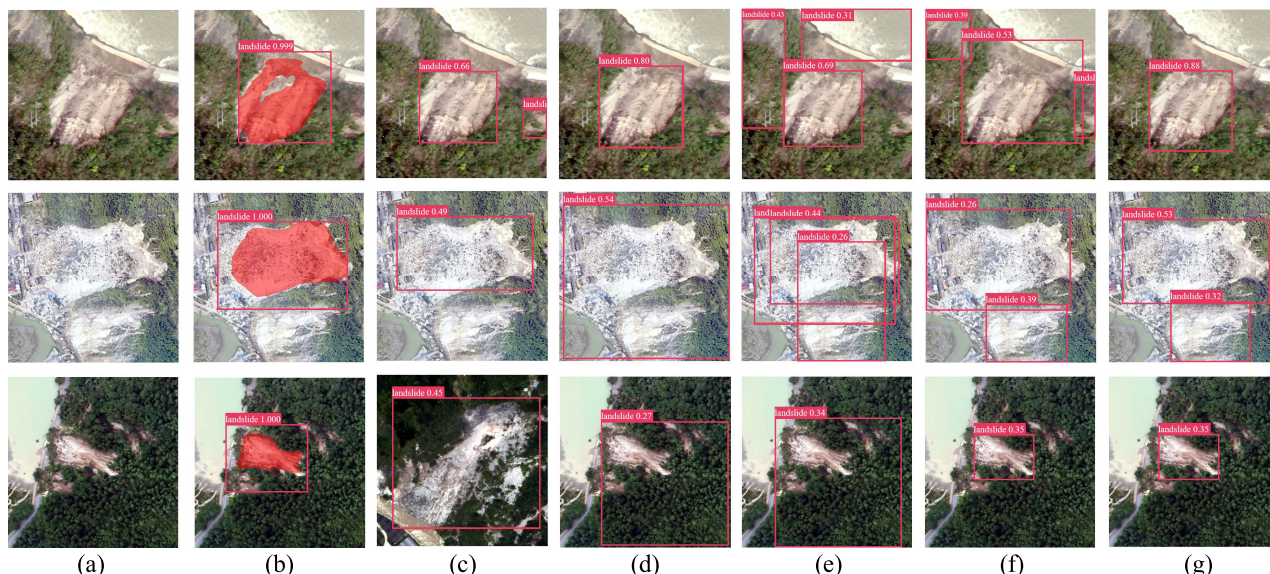
FIGURE 11. Detect results of test dataset for six models. (a) Origin images; (b) Mask R-CNN; (c) YOLOv5; (d) YOLOv7; (e) YOLOv8; (f) YOLOv9; (g) our LSI-YOLOv8.

TABLE 8. Comparison of identification results for different datasets.

Datasets	Descriptions
Bijie Dataset [19]	Number of images of landslides:770 Number of non-landslide images：2003
Moxi Town [52]	sample size (statistics)：1635
Jiuzhai Valley [52]	sample size (statistics)：250
Southwest Landslide Dataset [53]	sample size (statistics)：501

TABLE 9. Comparison of identification results for different datasets.

Datasets	YOLOv8			LSI-YOLOv8 (Ours)		
	mAP @0.5	F1-Score	FPS (f/s)	mAP @0.5	F1-Score	FPS (f/s)
Bijie Dataset[19]	82.3%	78.4%	67.3	88.6%	84.2%	73.2
Moxi Town [52]	68.5%	67.0%	68.3	68.0%	65.0%	60.2
Jiuzhai valley[52]	76.8%	74.5%	69.0	77.7%	76.3%	61.2
Southwest landslide dataset [53]	57.4%	58.1%	68.1	54.4%	55.7%	62.4

The comparative experiments results, as presented in Table 9, indicate that the improved LSI-YOLOv8 model outperformed the original YOLOv8 in various datasets used in similar configurations on three metrics, exhibiting the highest accuracy for landslide recognition and algorithmic efficiency from remote sensing imagery. However, the model performance varied across different datasets; while Bijie Datasets performed relatively well, the Southwest Landslide Datasets showed the worst performance. These findings strongly highlight the advantages of utilizing the number of sample images and emphasize the importance of high-quality imagery. Overall, the experiments further validated the robustness of our improved LSI-YOLOv8 model on various datasets from remote sensing imagery.

V. CONCLUSION

In recent years, there has been a significant increase in landslide studies utilizing deep learning algorithms. These algorithms have the advantage over traditional machine learning methods as they can automatically extract the most efficient features by leveraging deep convolutional layers from large datasets. Unlike machine learning, deep learning eliminates the requirement for manual construction and selection of feature layers for land-slide feature processing. Furthermore, deep learning is capable of handling larger sample sizes, making it well-suited for landslide identification in larger scenes [54].

To improve the efficiency of landslide identification in complex backgrounds using remote sensing imagery and reduce missed identifications, this study presents the LSI-YOLOv8 model as an enhancement of the YOLOv8 model. Several improvement methods are proposed for the YOLOv8 model in this study. Firstly, the EIoU loss function replaces the original CIoU bounding box loss function to enhance both the convergence speed and positioning accuracy of the algorithm. Additionally, the DWRSeg module, based on the original C2F module, is introduced. The C2F-DWR module aims to reduce the model's parameters while enhancing its generalization ability, making it more adaptable to various input data and improving its overall performance. The LSI-YOLOv8 model demonstrates improved recognition accuracy compared to the original YOLOv8 model, showing a 4.8% enhancement in mAP@0.5 and a 6.3% improvement in mAP@0.5:0.95. Moreover, the training time is reduced by 0.07hours and the detection speed is notably faster by 5.9 f/s than the original YOLOv8. However, there are still some undetected landslides, which can be attributed to factors such as background complexity, long-distance image capture leading to blurriness, and limitations in detecting landslides

that closely resemble the surrounding environment despite the enhanced capabilities of the LSI-YOLOv8 model.

Additionally, the proposed LSI-YOLOv8 model offers valuable features including lightweight parameters and fast recognition accuracy, making it suitable for deployment in UAV remote sensing monitoring for landslide identification. Compared to the original YOLOv8 model, the LSI-YOLOv8 model effectively reduces computational complexity and improves identification speed. This enables rapid landslide identification post-disaster and facilitates the planning of evacuation routes. Moreover, the model can be trained using a landslide database from various geological disasters to enhance its accuracy and generalization capabilities. Additionally, it can be applied to identify other geological hazards such as loess landslides and rainfall landslides.

Conclusively, this study presents LSI-YOLOv8, an enhanced model for landslide detection using YOLOv8, designed for efficient and accurate identification in optical remote sensing images. Through innovative modifications, LSI-YOLOv8 surpasses the current state-of-the-art approaches in both speed and precision of identification landslides from remote-sensing imagery. However, there are still limitations and opportunities for further investigation. The current datasets are restricted to optical remote sensing data, suggesting potential enhancement by incorporating SAR training datasets. SAR data offers benefits in areas with cloud cover or poor lighting conditions, improving the adaptability and reliability of the model. Future developments could involve integrating factors such as DEM and spectral characteristics to enhance accuracy by providing comprehensive contextual information. By considering these aspects, the goal is to enhance the model's capability to detect subtle terrain and surface variations, resulting in more accurate predictions. In summary, while LSI-YOLOv8 represents a significant advancement in landslide detection, there is scope for refinement and expansion. Ongoing research in these areas shows promise for the development of more precise and dependable systems for landslide detection and response in the future.

ACKNOWLEDGMENT

REFERENCES

[1] A. Mohan, A. K. Singh, B. Kumar, and R. Dwivedi, "Review on remote sensing methods for landslide detection using machine and deep learning," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 7, p. e3998, Jul. 2021.

[2] G. Huang, S. Du, and D. Wang, "GNSS techniques for real-time monitoring of landslides: A review," *Satell. Navigat.*, vol. 4, no. 1, p. 5, Dec. 2023.

[3] C. Zhao, X. Liu, Q. Zhang, J. Peng, and Q. Xu, "Research on loess landslide identification, monitoring and failure mode with InSAR technique in heifangtai, Gansu," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 44, no. 7, pp. 996–1007, 2019, doi: 10.13203/j.whugis20190072.

[4] O. Sarychikhina, D. G. Palacios, L. A. Delgado Argote, and A. G. Ortega, "Application of satellite SAR interferometry for the detection and monitoring of landslides along the Tijuana—Ensenada scenic Highway, Baja California, Mexico," *J. South Amer. Earth Sci.*, vol. 107, Apr. 2021, Art. no. 103030.

[5] Z. Jianjun, H. Jun, L. Zhiwei, S. Qian, and Z. Wanji, "Recent progress in landslide monitoring with InSAR," *Acta Geodaetica et Cartographica Sinica*, vol. 51, no. 10, p. 2001, 2022.

[6] M. H. Derron and M. Jaboyedoff, "Preface 'LiDAR and DEM techniques for landslides monitoring and characterization,'" *Natural Hazards Earth Syst. Sci.*, vol. 10, no. 9, pp. 1877–1879, 2010, doi: 10.5194/nhess-10-1877-2010.

[7] K. Yikai, "Application of airborne LiDAR technology in geological disaster monitoring," *Geomatics Spatial Inf. Technol.*, vol. 40, no. 9, pp. 117–119, 2017.

[8] P. Dalei, X. Qiang, D. Xiujun, J. Yuanzhen, Q. Xing, and T. Yeqing, "Application of unmanned aerial vehicles low-altitude photogrammetry in investigation and evaluation of loess landslide," *Adv. Earth Sci.*, vol. 32, no. 3, p. 319, 2017.

[9] J. Wang, Y. Wei, A. Mei, J. Guan, X. Ding, and S. Wang, "Multidimensional extraction of UAV tilt photography-based information of loess landslide and its application," *Geol. China*, vol. 48, no. 2, pp. 388–401, 2021, doi: 10.12029/gc20210204.

[10] R. Ye, Q. Deng, and H. Wang, "Landsides identification based on image classification: A case study on Guizhoulaocheng landslide in the Three Gorges Reservoir Area," *Chin. J. Eng. Geophys.*, vol. 4, no. 6, pp. 574–577, 2007.

[11] Z. Wang, "Remote sensing for landslide survey, monitoring and evaluation," *Remote Sens. Land Resour.*, vol. 1, pp. 10–17, Jan. 2007.

[12] F. Fiorucci, F. Ardizzone, A. C. Mondini, A. Viero, and F. Guzzetti, "Visual interpretation of stereoscopic NDVI satellite images to map rainfall-induced landslides," *Landslides*, vol. 16, no. 1, pp. 165–174, Jan. 2019.

[13] B. Adriano, N. Yokoya, H. Miura, M. Matsuoka, and S. Koshimura, "A semiautomatic pixel-object method for detecting landslides using multitemporal ALOS-2 intensity images," *Remote Sens.*, vol. 12, no. 3, p. 561, Feb. 2020.

[14] J. Dou, K.-T. Chang, S. Chen, A. Yunus, J.-K. Liu, H. Xia, and Z. Zhu, "Automatic case-based reasoning approach for landslide detection: Integration of object-oriented image analysis and a genetic algorithm," *Remote Sens.*, vol. 7, no. 4, pp. 4318–4342, Apr. 2015. [Online]. Available: https://www.mdpi.com/2072-4292/7/4/4318

[15] P. Lu, A. Stumpf, N. Kerle, and N. Casagli, "Object-oriented change detection for landslide rapid mapping," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 701–705, Jul. 2011.

[16] D. Hongqiang, H. Sixi, J. Wenbin, and W. Hao, "Research on rapid identification technology of highway landslide in Mountainous areas of Southeast Fujian based on remote sensing data," *J. Natural Disasters*, vol. 32, no. 1, pp. 217–227, 2023.

[17] A. Stumpf and N. Kerle, "Object-oriented mapping of landslides using random forests," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2564–2577, Oct. 2011.

[18] T. G. Whiteside, G. S. Boggs, and S. W. Maier, "Comparing object-based and pixel-based classifications for mapping savannas," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, no. 6, pp. 884–893, Dec. 2011.

[19] S. Ji, D. Yu, C. Shen, W. Li, and Q. Xu, "Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks," *Landslides*, vol. 17, no. 6, pp. 1337–1352, Jun. 2020.

[20] S. Gameiro, E. S. Riffel, G. G. de Oliveira, and L. A. Guasselli, "Artificial neural networks applied to landslide susceptibility: The effect of sampling areas on model capacity for generalization and extrapolation," *Appl. Geography*, vol. 137, Dec. 2021, Art. no. 102598.

[21] M. Marjanović, M. Kovačević, B. Bajat, and V. Voženílek, "Landslide susceptibility assessment using SVM machine learning algorithm," *Eng. Geol.*, vol. 123, no. 3, pp. 225–234, 2011.

[22] Y. Long, W. Li, R. Huang, Q. Xu, B. Yu, and G. Liu, "Automatic extraction and evolution trend analysis of landslides in Mianyuan River Basin in the 10 years after Wenchuan Earthquake," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 45, no. 11, pp. 1792–1800, 2020.

[23] N. Lin, S. Feng, B. Wang, F. Tang, H. Zhu, D. Zhang, P. Pan, and J. He, "Research on rapid landslide extraction and analysis based on XGBoost from high resolution remote sensing," *Geomatics Inf. Sci. Wuhan Univ.*, to be published, doi: 10.13203/j.whugis20220296.

[24] W. Qi, M. Wei, W. Yang, C. Xu, and C. Ma, "Automatic mapping of landslides by the ResU-net," *Remote Sens.*, vol. 12, no. 15, p. 2487, Aug. 2020.

[25] X. Chen, C. Liu, L. Chen, X. Zhu, Y. Zhang, and C. Wang, "A pavement crack detection and evaluation framework for a UAV inspection system based on deep learning," *Appl. Sci.*, vol. 14, no. 3, p. 1157, Jan. 2024.

[26] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[27] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision— ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 21–37.

[28] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[29] H. Yu, Y. Ma, L. Wang, Y. Zhai, and X. Wang, "A landslide intelligent detection method based on CNN and RSG_R," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Aug. 2017, pp. 40–44.

[30] X. Wu, J. Yang, and R. Niu, "A landslide susceptibility assessment method using SMOTE and convolutional neural network," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 45, no. 8, pp. 1223–1232, 2020.

[31] J. Xiaoting, Z. Kang, Z. Xiaoqing, Z. Qi, and Z. Wen, "Landslides risk identification based on Faster R-CNN-a case study in Fugong County," *Ind. Minerals Process./Huagong Kuangwu Yu Jiagong*, vol. 51, no. 12, pp. 1–10, 2022.

[32] W. Jiang, J. Xi, Z. Li, M. Ding, L. Yang, and D. Xie, "Landslide detection and segmentation using mask R-CNN with simulated hard samples," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 48, no. 12, pp. 1931–1942, 2023.

[33] D. Pang, G. Liu, J. He, W. Li, and R. Fu, "Automatic remote sensing identification of co-seismic landslides using deep learning methods," *Forests*, vol. 13, no. 8, p. 1213, Aug. 2022.

[34] T. Wang, M. Liu, H. Zhang, X. Jiang, Y. Huang, and X. Jiang, "Landslide detection based on improved YOLOv5 and satellite images," in *Proc. 4th Int. Conf. Pattern Recognit. Artif. Intell. (PRAI)*, Aug. 2021, pp. 367–371.

[35] X. Wang, H. Gao, Z. Jia, and Z. Li, "BL-YOLOv8: An improved road defect detection model based on YOLOv8," *Sensors*, vol. 23, no. 20, p. 8361, Oct. 2023.

[36] H. Wei, X. Liu, S.-T. Xu, Z. Dai, Y. Dai, and X. Xu, "DWRSeg: Rethinking efficient acquisition of multi-scale contextual information for real-time semantic segmentation," 2022, *arXiv:2212.01173*.

[37] Y.-F. Zhang, W. Ren, Z. Zhang, Z. Jia, L. Wang, and T. Tan, "Focal and efficient IOU loss for accurate bounding box regression," *Neurocomputing*, vol. 506, pp. 146–157, Sep. 2022.

[38] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach. Learn. Knowl. Extraction*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023.

[39] G. Wang, Y. Chen, P. An, H. Hong, J. Hu, and T. Huang, "UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios," *Sensors*, vol. 23, no. 16, p. 7190, Aug. 2023.

[40] C.-Y. Wang, A. Bochkovskiy, and H.-Y.-M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 7464–7475.

[41] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: Faster and better learning for bounding box regression," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 2020, no. 7, pp. 12993–13000.

[42] J. Yu, Y. Jiang, Z. Wang, Z. Cao, and T. Huang, "UnitBox: An advanced object detection network," in *Proc. 24th ACM Int. Conf. Multimedia*, Oct. 2016, pp. 516–520.

[43] X. Zhang, T. Shen, and D. Xu, "Remote-sensing image object detection based on improved YOLOv8 algorithm," *Laser Optoelectron. Prog.*, vol. 61, no. 10, 2024, Art. no. 1028001.

[44] Z. Gevorgyan, "SIoU loss: More powerful learning for bounding box regression," 2022, *arXiv:2205.12740*.

[45] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 658–666.

[46] G. Yang, J. Lei, Z. Zhu, S. Cheng, Z. Feng, and R. Liang, "AFPN: Asymptotic feature pyramid network for object detection," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2023, pp. 2184–2189.

[47] Z. Xia, X. Pan, S. Song, L. E. Li, and G. Huang, "Vision transformer with deformable attention," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 4784–4793.

[48] X. Ding, X. Zhang, J. Han, and G. Ding, "Diverse branch block: Building a convolution as an inception-like unit," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 10881–10890.

[49] R. Fu, Q. Hu, X. Dong, Y. Guo, Y. Gao, and B. Li, "Axiom-based grad-CAM: Towards accurate visualization and explanation of CNNs," 2020, *arXiv:2008.02312*.

[50] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.

[51] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "YOLOv9: Learning what you want to learn using programmable gradient information," 2024, *arXiv:2402.13616*.

[52] Y. Xu, C. Ouyang, Q. Xu, D. Wang, B. Zhao, and Y. Luo, "CAS landslide dataset: A large-scale and multisensor dataset for deep learning-based landslide detection," *Sci. Data*, vol. 11, no. 1, p. 12, Jan. 2024, doi: 10.1038/s41597-023-02847-z.

[53] W. Zhang, Z. Liu, S. Zhou, W. Qi, X. Wu, T. Zhang, and L. Han, "LS-YOLO: A novel model for detecting multiscale landslides with remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 4952–4965, 2024.

[54] Y. Ju, Q. Xu, S. Jin, W. Li, X. Dong, and Q. Guo, "Automatic object detection of loess landslide based on deep learning," *J. Wuhan Univ.*, vol. 45, no. 11, pp. 1747–1755, 2020, doi: 10.13203/j.whugis20200132.

XINBAO CHEN was born in Hunan, China, in 1980. He received the Ph.D. degree in cartography and geographic information engineering from Central South University, in 2011, under the guidance of Prof. Li Songnian and Prof. Zhu Jianjun. Prior to this, he studied with Ryerson University, Canada, from 2007 to 2010. Currently, he works with the School of Earth Sciences and Spatial Information Engineering, Hunan University of Science and Technology. He has led various research projects funded by organizations, such as the National Natural Science Foundation and China Postdoctoral Foundation, and participated in multiple national projects. He has an extensive publication record, including more than 30 SCI/EI/CSCD academic articles, eight software copyrights, two invention patents, one academic monograph, and one proposed textbook. His research interests include spatial and temporal data modeling, spatial and temporal big data modeling and mining, spatial and temporal patterns of regional economy, regional innovation differences, impact factor response models, and geographic information system software development.

CHANG LIU was born in Heilongjiang, China, in 1998. She is currently studying with the School of Earth Science and Spatial Information Engineering, Hunan University of Science and Technology, under the supervisor of Asst. Prof. Chen Xinbao. Her main research interest includes disaster remote sensing.

SHAN WANG was born in Hunan, China, in 1998. She received the degree from the School of Earth Sciences and Spatial Information Engineering, Hunan University of Science and Technology, in 2023. She is currently with Hunan Province Geological Disaster Survey and Monitoring Institute, Changsha, Hunan. Her main research interest includes disaster remote sensing.

XINPING DENG born in Hunan, China, in 1983. He is currently employed at Hunan Institute of Geological Disaster Investigation and Monitoring, Changsha, Hunan, China. His main responsibilities include hydrogeology, urban geology, environmental geology, ecological geology survey, and evaluation. He also contributes to natural resources, land, and spatial surveys, monitoring, and other related tasks.

. . .