LS-YOLO: A Novel Model for Detecting Multiscale Landslides With Remote Sensing Images

Wenjie Zhang, Zhiheng Liu, *Member, IEEE*, Suiping Zhou, Wenjuan Qi, Xinjun Wu, Tianyu Zhang, and Ling Han

***Abstract*—The landslide is a widespread and devastating natural disaster, posing serious threats to human life, security, and natural assets. Investigating efficient methods for accurate landslide detection with remote sensing images has important academic and practical implications. In this article, we proposed an LS-YOLO, a novel and effective model for landslide detection with remote sensing images. We first built a multiscale landslide dataset (MSLD) and introduced random seeds in the data augmentation to increase data robustness. Considering the multiscale characteristic of landslides in remote sensing images, a multiscale feature extraction module is designed based on efficient channel attention, average pooling, and spatial separable convolution. To increase the receptive field of the model, dilated convolution is employed to the decoupled head. Specifically, the context enhancement module consisting of dilation convolutions is added to the decoupled head regression task branch, and then the improved decoupled head is to replace the coupled head in YOLOv5s. Extensive experiments show that our proposed model has high performance for multiscale landslide detection and outperforms other object detection models (faster RCNN, SSD, EfficientDet-D0, YOLOv5s, YOLOv7, and YOLOX). Compared with the baseline model YOLOv5s, the AP of the LS-YOLO for detecting landslides has increased by 2.18%–97.06%.**

Manuscript received 20 September 2023; revised 7 November 2023 and 2 January 2024; accepted 2 February 2024. Date of publication 7 February 2024; date of current version 22 February 2024. This work was supported in part by the Open Fund of the Key Laboratory of Mine Geological Hazards Mechanism and Control, Ministry of Natural Resources, under Grant 2022-08, in part by the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People's Republic of China, under Grant KLSMNR-G202303, in part by the Open Fund of Beijing Key Laboratory of Advanced Optical Remote Sensing Technology under Grant AORS20238, in part by the Project of Natural Science Basic Research Program of Shaanxi under Grant 2023-JC-QN-0299, in part by the Fundamental Research Funds for the Central Universities, CHD, under Grant 300102353502, and in part by the Fundamental Research Funds for the Central Universities, XDU, under Grant XJS221307. *(Corresponding authors: Zhiheng Liu; Suiping Zhou.)*

Wenjie Zhang and Zhiheng Liu are with the School of Aerospace Science and Technology, Xidian University, Xi'an 710126, China, and also with the Key Laboratory of Mine Geological Hazards Mechanism and Control, Ministry of Natural Resources, Xi'an 710054, China (e-mail: 21131213387@stu.xidian.edu.cn; liuzhiheng@xidian.edu.cn).

Suiping Zhou, Wenjuan Qi, and Xinjun Wu are with the School of Aerospace Science and Technology, Xidian University, Xi'an 710126, China (e-mail: spzhou@xidian.edu.cn; 21131213413@stu.xidian.edu.cn; 15529559608@163.com).

Tianyu Zhang is with the Key Laboratory of Mine Geological Hazards Mechanism and Control, Ministry of Natural Resources, Xi'an 710054, China, and also with the Shaanxi Institute of Geological Survey, Xi'an 710068, China (e-mail: 15388611751@163.com).

Ling Han is with the Shaanxi Key Laboratory of Land Reclamation Engineering, Chang'an University, Xi'an 710064, China, and also with the School of Land Engineering, Chang'an University, Xi'an 710064, China (e-mail: hanling@chd.edu.cn).

The code and MSLD will be available at https://github.com/wenjieo/LS-YOLO.

Digital Object Identifier 10.1109/JSTARS.2024.3363160

***Index Terms*—Deep learning, dilated convolution, landslide detection, remote sensing images, spatial separable convolution, YOLOv5s.**

I. Introduction

LANDSLIDES are one of the most common and destructive disasters in mountainous regions [1], [2], [3], [4]. Often triggered by earthquakes or heavy rains, they not only damage infrastructure, such as roads, bridges, and electricity lines, but also destruction of vegetation and soil, causing land degradation [5]. In China, the vastness of the country, the large number of mountains, and the abundance of year-round rainfall in a large part of the country have led to a significantly higher frequency of landslide disasters than in other countries. In particular, the mountainous regions of Yunnan, Guizhou, and Sichuan in China often experience large numbers of landslides. Thus, the accurate detection of landslides is essential to prevent and respond to landslide disasters [6].

Remote sensing imaging allows for real-time imaging of the Earth due to its high spatial resolution and extensive coverage. Landslide detection methods with remote sensing images can be roughly divided into two categories: manual visual interpretation and methods based on computer vision [1]. The manual visual interpretation is performed by geological experts who discriminate landslides based on their geometric, textural, and other features in remote sensing images [7]. Although manual visual interpretation of landslides has a high accuracy, it has the disadvantages, such as being time-consuming, dependent on the expertise of experts, and erroneous quantitative descriptions. Thus, the manual visual interpretation makes meeting the requirement for rapid landslide detection challenging [8], [9], [10]. The use of computer vision methods to identify natural disasters, such as landslides, earthquakes, and glacial movements, with remote sensing images has become a research hotspot with the rapid development of remote sensing technology [11], [12], [13], [14].

The methods based on computer vision provide promising solutions to address the above problems and deep learning being the most representative approach [15]. In recent years, deep learning, a type of artificial neural network model, has made remarkable strides in the field of object detection [16]. As deep learning advances, methods based on computer vision have become the most widely used technique in the area of landslide detection [12], [17], [18], [19], [20]. The deep learning models have two categories: one-stage and two-stage models,

depending on whether they require the generation of proposal regions. Two-stage models make use of either a selective search algorithm or the region proposal network to generate region proposals, which are then employed for object detection. Representative two-stage models are RCNN [21], fast RCNN [22], faster RCNN [23], and mask RCNN [24]. One-stage models consider all the positions of the input image as potential objects and, subsequently, endeavor to classify each interesting area as either the background or an object class. One-stage models are faster and more efficient than two-stage models. Representative models of one stage include SSD [25], EfficientDet [26], and the series of YOLO [27], [28], [29], [30]. These techniques have shown a remarkable enhancement in the precision and speed of object detection when compared with the conventional approaches.

Many scholars have explored the application of deep learning methods to identify landslides. Li et al. [12] designed an innovative graph convolutional network to detect landslides. Their approach draws inspiration from the attention mechanism's capacity to focus on specific information, complemented by different channels. Niu et al. [11] proposed the Reg-SA-Unet++ for landslide detection, which was based on Unet++. Their model was improved by removing deep supervised pruning, employing RegNet to replace the convolutional blocks, and adding attention modules at each layer. Ghorbanzadeh et al. [31] proposed a ResU-Net model and constructed a basic rule-based OBIA to detect landslides. Ullo et al. [32] exploited the mask R-CNN to detect landslides. Sameen and Pradhan [33] conducted a comparison between a one-layer convolutional neural network (CNN) and two corresponding deeper counterparts, as well as residual networks to identify landslides. Cheng et al. [34] proposed the YOLO-SA based on the group convolution, ghost bottleneck, and attention mechanism to detect landslides.

Research on landslide detection using computer vision methods has achieved great progress. However, there are still some problems that need to be further investigated. First, in the existing publicly accessible landslide datasets, the landslide samples are limited, the intraclass variation is small, and the background is simple, making it challenging to effectively train the models. Second, previous research on detecting landslides using computer vision methods did not fully integrate landslide features, such as a wide-scale range of landslide scale and diverse shapes, resulting in poor accuracy of landslide detection and difficulty meeting practical needs.

A novel landslide detection model (LS-YOLO) is proposed in this article, which results in excellent performance in multiscale landslide detection. Our major contributions can be summarized as follows.

1) A multiscale landslide dataset (MSLD) is created, which has the advantages of a significant number of landslide samples, a high degree of intraclass variation, a wide range of landslide sizes, and complex backgrounds.
2) We proposed the novel multiscale feature extraction (MSFE) module to adequately extract landslide features from several receptive fields by five parallel branches. These branches consist of average pooling or spatial separable convolution, which increases the depth of the network and improves the landslide detection accuracy of the model.
3) We improve the existing decoupled head. We applied several dilated convolutions in the regression task of the decoupled head. The combination of different dilated rate convolutions helps to capture multiscale context information and improve the landslide location accuracy of the model.

The rest of this article is organized as follows. Section II outlines the specifics of MSLD generation. Section III presents the LS-YOLO model for detecting landslides. In Section IV, comprehensive experimental results and analysis are presented. Section V discusses the advantages and disadvantages of LS-YOLO and future research work. Finally, Section VI concludes this article.

II. Landslide Dataset

The methods based on deep learning to detect landslides require a lot of landslide image data. However, the accessible landslide datasets ineffectively train deep learning models due to limited and similar landslide samples. For this reason, this article built an MSLD. First, we collected two public landslide datasets, the first being the Bijie landslide dataset created by Ji et al. [17]. The study area is located in Bijie, a northwestern part of Guizhou Province, China. The region spans across an area of 26 853 km^2 and its altitude ranges from 457 to 2900 m. Bijie is located in the transitional slope zone between the Tibetan plateau and the eastern highlands and is one of the most severe landslide areas in China. The Bijie landslide dataset consists of 770 positive sample images containing landslides and 2003 negative sample images without landslides, acquired by the TripleSat series of satellites between May and August 2018. The images have a ground resolution of 0.8 m. As shown in Fig. 1(a), the red dots indicate the spatial location distribution of the landslides.

The labels in the Bijie landslide dataset are used for instance segmentation. Therefore, we research generated landslide bounding boxes automatically based on the boundary points in the mask files. This method ensures that generated labels are suitable for landslide detection. The generation process of the landslide detection labels is shown in Fig. 2.

The Bijie landslide dataset contains not only a large number of landslide samples but also abundant negative samples, such as mountains, villages, roads, rivers, forests, and agricultural lands. However, all the positive sample images in the Bijie landslide dataset contain one landslide and no images containing multiple landslides, making it challenging to assess the performance of models in multiple landslides detection.

Therefore, we gathered another available landslide dataset to enrich the diversity and robustness of landslide data (https://github.com/YhQIAO/LandSlide_Detection_Faster-RCNN), which is named Southwest landslide dataset. As shown in Fig. 1(b), this study area is widely distributed and covers five provinces in southwestern China, including Gansu, Sichuan, Guizhou, Yunnan, and Tibet. It is located between 90°23′–106°39′E and 22°27′–33°56′N, and has typical subtropical monsoon climate and alpine climate. The

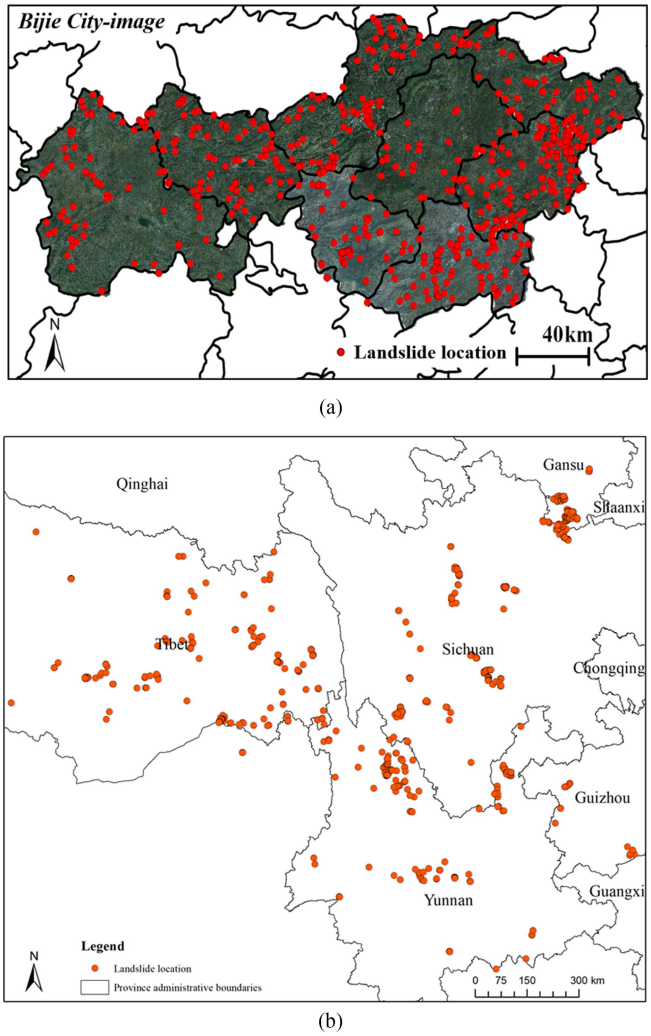

Fig. 1. Spatial distribution of landslides in the two datasets, respectively. (a) Landslides location in the Bijie dataset [17]. (b) Landslides location in the Southwest dataset.

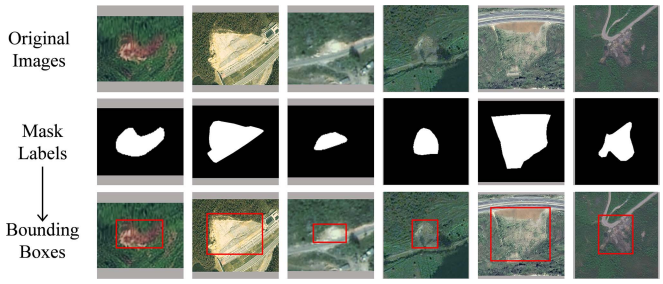

Fig. 2. Landslide detection labels generation process for the Bijie landslide dataset.

terrain's elevation ranges from 500 to 4500 m, characterized by interlaced mountains, towering peaks, deep valleys, and an intricate network of rivers. Heavy rains account for more than 50% of the annual total precipitation, resulting in frequent landslide disasters. This landslide dataset consists of 500 positive sample images, each containing one or multiple landslides. These images are labeled by experts in the field of remote sensing using the LabelImg software to generate labels in VOC format.

There are only 1270 images containing landslide in the mentioned two landslide datasets, which makes it difficult to train landslide detection models with excellent performance. Unfortunately, it is difficult to access big data in the landslide detection domain. To cover these problems, we use a combination of offline augmentation and online augmentation [35], [36]. First, we employed offline augmentation on 1270 positive sample images to significantly increase the diversity of the data and the size of the dataset. The offline augmentation techniques include geometric transformations (the clockwise rotation consists of 90°, 180°, 270°, flip operation, Salt noise injection, and Gaussian noise injection) and color space transformations (image darkening, image brightening, and contrast adjustment). Meanwhile, to prevent the model from learning augmented data patterns, this article introduces a random seed in the offline augmentation process. Each positive sample image randomly selects five of the above techniques for augmentation. As shown in Fig. 3, the a-original, b-original, c-original, and d-original are the original images. The first four rows of images are generated by using offline augmentation on the four original images. The images e–j are negative sample images without landslides in MSLD. Then, we used online augmentation (Mosaic, Mixup, and Copy_paste) during model training to train landslide detection models better.

The MSLD contains 7620 positive sample images with multiple landslide types and 2003 negative sample images. As shown in Fig. 4, the first row includes the loess landslide cases and the second row includes the rocky landslide cases in MSLD. It can be seen that the color and texture features of landslides in MSLD are markedly dissimilar. Meanwhile, as shown in Fig. 5, the shape and size of landslides in MSLD are notably different.

In summary, the MSLD has significant advantages, a large number of landslide samples, a high degree of intraclass variation, a wide range of landslide sizes, and complex backgrounds.

III. Methodology

YOLO is a typical one-stage network model that was proposed in 2015 by Redmon et al. [37]. After a series of development, YOLOv2, YOLOv3, YOLOv4, and YOLOv5 were derived [38]. YOLOv5 has four versions, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. The model size and detection accuracy increase sequentially. The four versions differ only in the setting of width and depth factors in the model. YOLOv5 consists of four parts: input, backbone, neck, and head, respectively. The input carries out data augmentation, adaptive anchor calculation, and adaptive image scaling on input images before feeding them into the backbone. The backbone transforms the input image into multiscale feature maps and extracts the feature information of targets. The neck fuses the different scale feature information extracted from the backbone. The head performs object detection and confidence calculation and outputs the prediction results of the network.

Considering the favorable accuracy–size–speed tradeoff, this article selects the YOLOv5s as the baseline model for

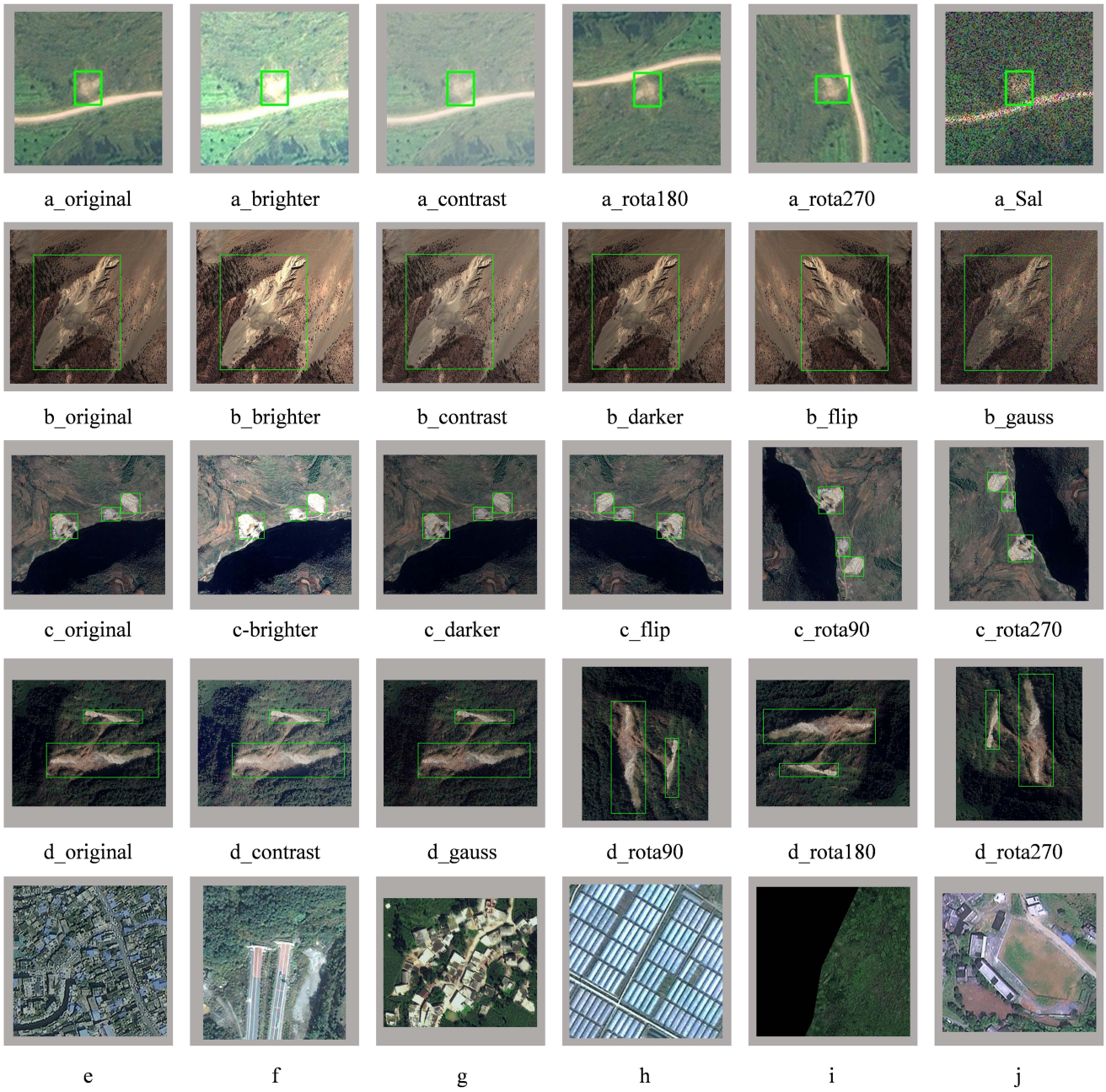

Fig. 3. Various instances in the MSLD.

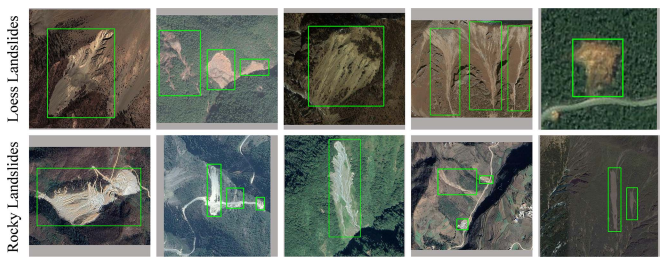

Fig. 4. Different landslide cases in MSLD.

optimization and designs a novel model for detecting multiscale landslides with remote sensing images, called LS-YOLO.

A. LS-YOLO

Given the diversity of landslides in remote sensing images, directly using the generic object detection model YOLOv5s to high-precision detect landslides is challenging. To address this issue, we designed the LS-YOLO model that has outstanding performance in landslide detection.

As shown in Fig. 6, the features of LS-YOLO include the following two main areas.

1) An MSFE module was proposed to improve the accuracy of the model for landslide detection. The MSFE consists of efficient channel attention (ECA) [39], average pooling, and spatial separable convolution. We introduced the MSFE into the neck of YOLOv5s.

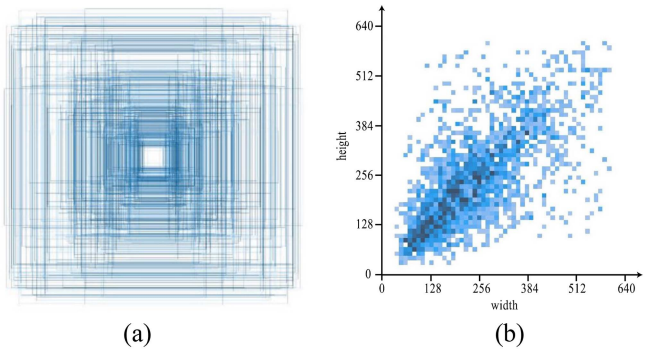

Fig. 5. Statistical of landslides shape and size in MSLD. (a) Landslides shape. (b) Landslides size.

2) The decoupled head was improved to increase the accuracy of the model for landslide locations. The context enhancement module (CEM) consisting of dilated convolution was employed to replace the 3×3 convolution in the decoupled head regression task branch. The coupled head in the YOLOv5s was replaced by our improved decoupled head.

B. MSFE Module

The complexity and diversity of landslides in remote sensing images make it difficult to effectively extract landslide feature information by deep learning models. In this article, we designed an MSFE module that effectively enhances the adaptability to scales and nonlinearity of the model, ultimately improving the performance of YOLOv5s for detecting multiscale targets. As shown in Fig. 7, MSFE contains two branches, the first one is a residual connection, which could mitigate gradient vanishing and speed up model training. The second branch consists of ECA, average pooling, and spatial separable convolution, which could fully extract landslide feature information.

1) Efficient Channel Attention: ECA is a versatile plug-and-play block that enhances the performance of CNNs [39]. ECA consists of a squeeze module to condense global spatial information and an excitation module to achieve channel interactions. As shown in Fig. 8, after using global average pooling without dimensionality reduction to aggregate features, the ECA adaptively determines kernel size k. Then, the one-dimensional (1-D) convolution with kernel size k is conducted on the feature maps, followed by the Sigmoid function to obtain the channel attention vector. The output feature maps are obtained by scaling each channel of the input feature maps by multiplying the corresponding element in the attention vector. The formulations of the ECA are given as follows:

$$w = F_{\mathrm{ECA}}(X) = \sigma(\mathrm{Conv1D}(\mathrm{GAP}(X))) \tag{1}$$

$$Y = wX \tag{2}$$

where σ is the Sigmoid function, and Conv1D denotes the 1-D convolution with kernel size k. The kernel size k is determined based on the channel dimensionality C adaptively. w denotes the channel attention vector. X and Y represent the input feature maps and the output feature maps, respectively

$$k = \psi(C) = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{\mathrm{odd}} \tag{3}$$

where $|x|_{\mathrm{odd}}$ represents the nearest odd number of x. The γ and b are the hyperparameters. We assigned the values of 2 and 1 to γ and b.

2) Average Pooling: Average pooling works out the mean of the pixels in the input feature map relevant area based on the kernel size. The advantage of average pooling is that there is no parameter to optimize; thus, overfitting is avoided. As shown in Fig. 9, the computation of average pooling can be expressed as follows:

$$F_{\mathrm{AP}} = \frac{1}{s \times s} \sum_{i=1}^{s} \sum_{j=1}^{s} X_{ij} \tag{4}$$

where s denotes the kernel size, and X_{ij} indicates the pixel value of the point (i,j) in the computational region.

3) Spatial Separable Convolution: Spatial separable convolution splits a standard convolution operation into multiple small kernel convolution operations in the spatial dimension. As shown in (5) and (6), the $k \times k$ convolution is equivalent to the branch consisting of $k \times 1$ convolution and $1 \times k$ convolution

$$\begin{bmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{k1} & \cdots & w_{kk} \end{bmatrix}_{kk} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \times \begin{bmatrix} y_1 & \cdots & y_k \end{bmatrix} \tag{5}$$

$$w_{ij} = x_i \times y_j, (1 \le i, j \le k) \tag{6}$$

where w_{ij} denotes the value of the point (i, j) in the $k \times k$ convolution, x_i represents the value of the point $(i, 1)$ in the $k \times 1$ convolution, and y_j denotes the value of the point $(1, j)$ in the $1 \times k$ convolution. The number of parameters of $k \times 1$ convolution and $1 \times k$ convolution is $(k + 1) + (1 + k) = 2k$, which is smaller than the number of parameters of the $k \times k$ convolution when $k > 2$. Spatial separable convolution reduces the module parameters, speeds up the model computation, and increases the model depth. As shown in Fig. 10, the 3×3 convolution was decomposed into a 3×1 convolution and a 1×3 convolution.

In the MSFE module, the feature maps first go through ECA, which has small parameters but brings significant performance improvement. Then, the multiscale features of landslides are fully extracted by five parallel lines, which are as follows:

1) average pooling;
2) 1×1 convolution;
3) continuous 3×1 convolution and 1×3 convolution;
4) continuous 3×1 convolution, 1×3 convolution, 3×1 convolution, and 1×3 convolution;
5) continuous 5×1 convolution, 1×5 convolution, 5×1 convolution, and 1×5 convolution.

The feature maps output from these five lines are then spliced by channel, followed by a 1×1 convolution to fuse the information and downscale channels of feature maps. Finally, the output feature maps of the two branches are summed in the spatial dimension, which is considered as the output of the MSFE.

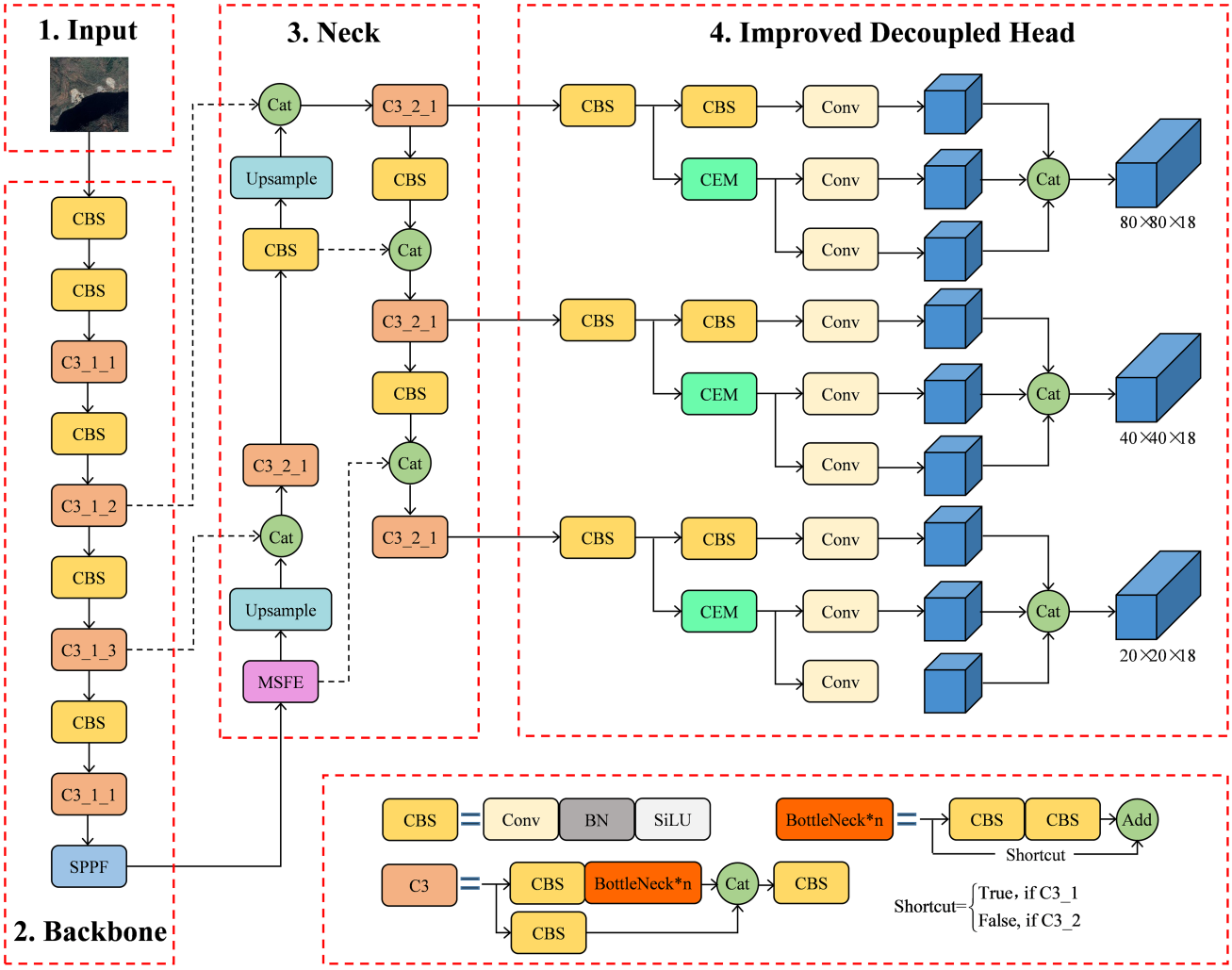

Fig. 6. Overall framework of the proposed LS-YOLO.

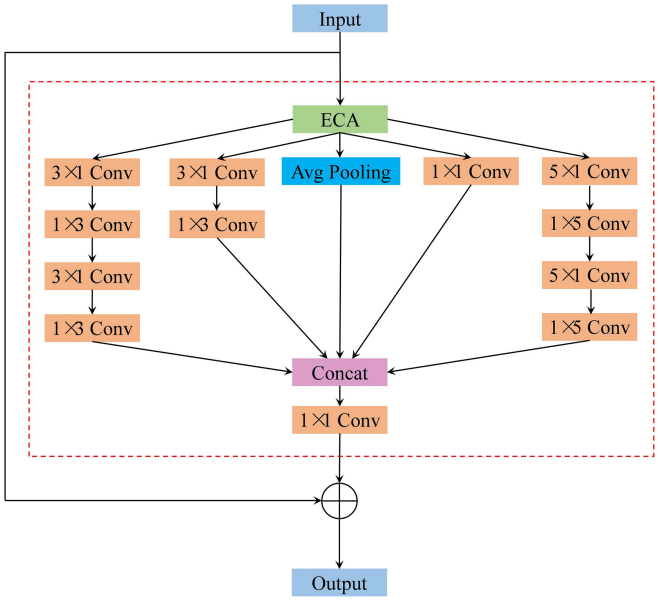

Fig. 7. Structure of MSFE.

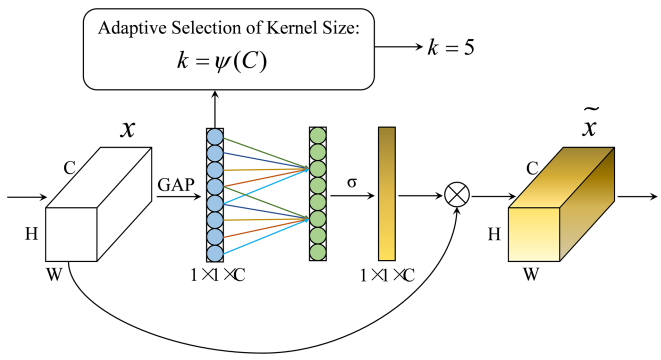

Fig. 8. Structure of ECA.

C. Improved Decoupled Head

As shown in Fig. 11, YOLOv5s shares parameters between classification and regression tasks. However, some research has shown that conflict exists between classification and regression tasks [40], [41]. The classification task focuses on the texture

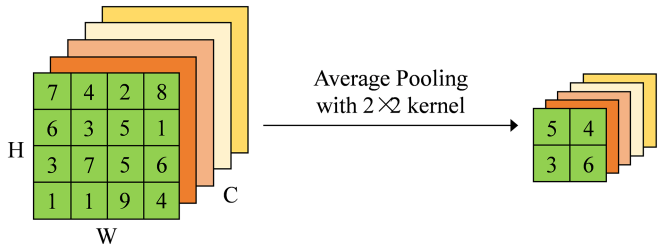

Fig. 9. Calculation process of average pooling.

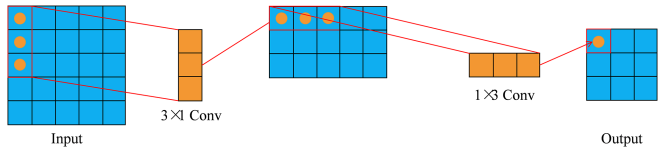

Fig. 10. Calculation process in 3 × 3 spatial separable convolution.

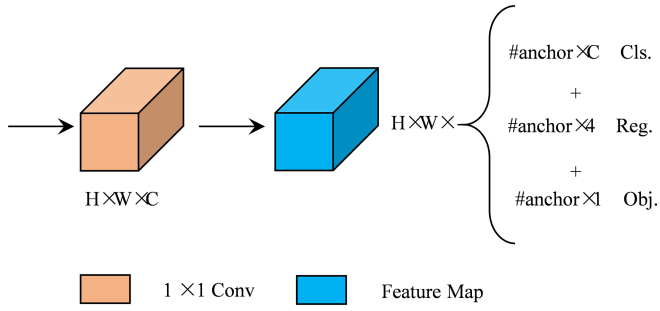

Fig. 11. Structure of coupled head in YOLOv5s.

information of the target and identifies the extracted features as more comparable to those of the existing categories. Conversely, the regression task prioritizes the edge information of the target and adjusts the anticipated bounding box parameters based on the positional coordinates of the ground truth. Thus, decoupling the classification and regression tasks will improve the detection performance and convergence speed of the model.

In this article, the decoupled head was improved. The structure of the decoupled head before and after the improvement is shown in Fig. 12. In the original decoupled head [42], two parallel branches are used to decouple the classification and the regression tasks to mitigate conflict. Each branch comprises a 3 × 3 convolution and a 1 × 1 convolution. Considering the multiscale characteristic of landslides in remote sensing images, this article used the CEM [43] consisting of dilated convolutions to substitute the 3 × 3 convolution on the regression task branch to improve the model's location accuracy for multiscale landslides.

As shown in Fig. 13(a), the CEM comprises 1 × 1 convolution and 3 × 3 dilated convolutions with dilation rates of 1, 3, and 5, respectively. The CEM significantly enhances the receptive field of the model. Fig. 13(b), (c), and (d) demonstrate the three different fusion scenarios: adaptive fusion, concatenation fusion, and weighted fusion, respectively.

1) Adaptive Fusion: As shown in (7) and (8), the adaptive fusion initially conducts separate 1 × 1 convolution on the three feature maps, followed by concatenation of the feature maps along channel dimensions. Subsequently, another 1 × 1 convolution and the Softmax are implemented to obtain weights for the three feature maps. These weights are multiplied with their corresponding feature maps, and the resulting three feature maps are summed in the spatial dimension to produce the output

$$W_{[1,2,3]} = \delta(\mathrm{Conv}(\mathrm{Cat}[\mathrm{Conv}(F_1), \mathrm{Conv}(F_2), \mathrm{Conv}(F_3)])) \tag{7}$$

$$F_{\mathrm{out}} = \sum_{i=1}^{3} W_i \times F_i \tag{8}$$

where δ denotes the Softmax, Cat stands for the concatenation operation along the channel dimension, Conv is the 1 × 1 convolution, and F_{out} denotes the output feature map.

2) Concatenation Fusion: Concatenation fusion concatenates the feature maps of the three branch outputs in the channel dimension. The number of channels in the output feature map is three times greater than that of the input feature map. Thus, we added an extra 1 × 1 convolution after concatenation fusion to ensure that the input and output feature map channels are consistent.

3) Weighted Fusion: Weighted fusion directly adds the output feature map of the three branches in the spatial dimension.

The three fusion scenarios improve the generalization and robustness of CEM. Specifically, the improved decoupled head first performs a 1 × 1 convolution on the input feature maps to realize the information interaction between channels and then carries out the classification task and the regression task, respectively. The classification task branch consists of a 3 × 3 convolution and a 1 × 1 convolution. The regression task branch consists of the CEM and a 1 × 1 convolution.

IV. Experimental Results and Analysis

All experiments in this article were conducted under the same hardware and software environments, employing PyTorch 1.9 and the 64-bit Windows 10 operating system. The GPU utilized was NVIDIA GeForce RTX 3070 with a video memory size of 8 GB and 31.7 GB of RAM. All models were trained for 300 epochs. During the model's training and testing, both IOU and confidence threshold were set to 0.5. The three online augmentation techniques, Mosaic, Mixup, and Copy_Paste, are applied in real time during model training to improve the model's generalization ability. The hyperparameter configurations of the proposed model in this article are shown in Table I.

The MSLD was divided into the training validation set and the test set in a 7 : 3 ratio and the training validation set was divided into the train and the validation sets in a 9 : 1 ratio. The division of MSLD is shown in Table II. The MSLD includes 9623 images. The train set, validation set, and test set include 6062 images, 674 images, and 2887 images, respectively.

A. Evaluation Metrics

In this article, Precision (P), Recall (R), and AP were used to quantitatively evaluate the accuracy of the models in landslide

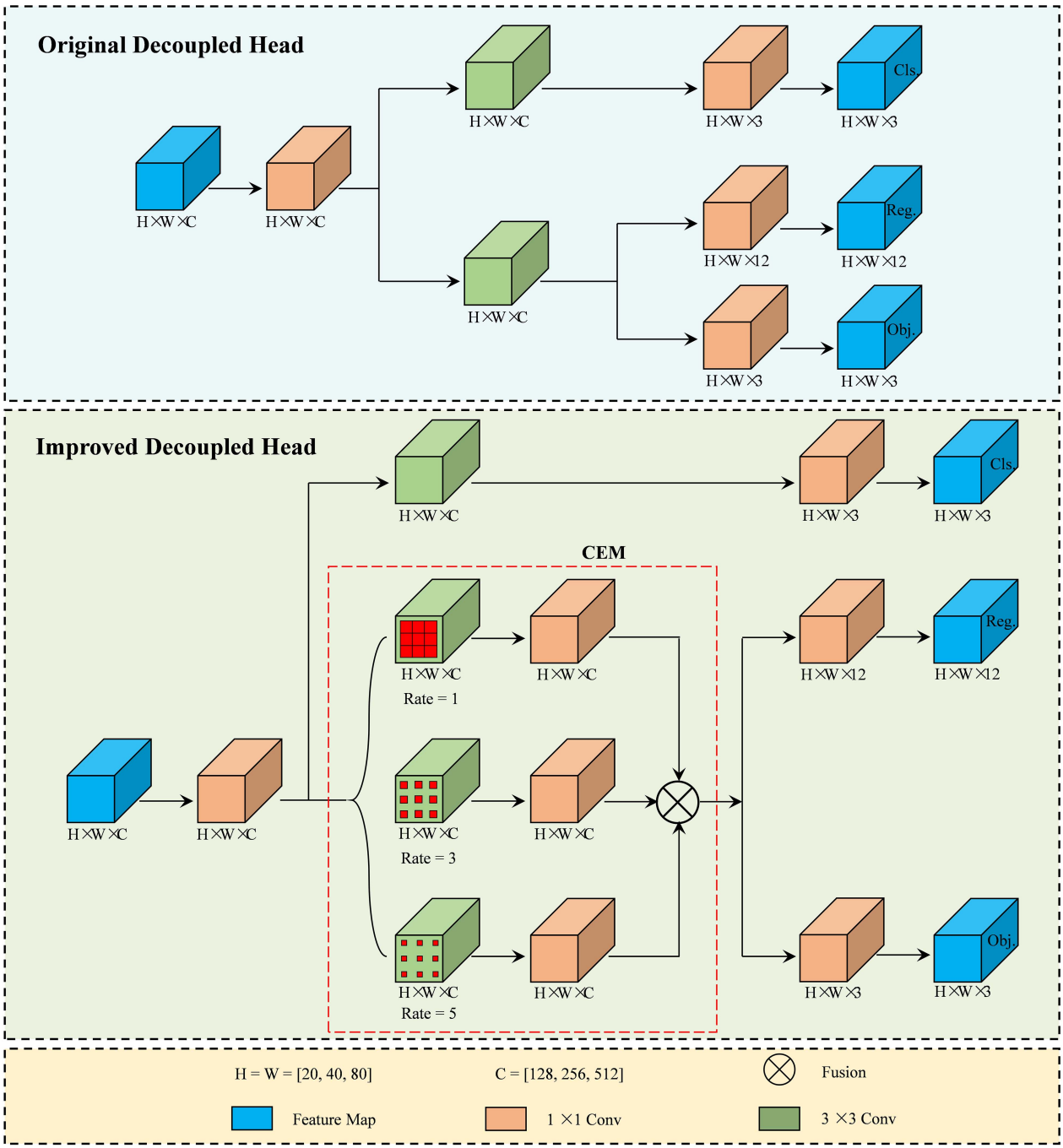

Fig. 12. Structure of the original decoupled head and the improved decoupled head.

detection. The formulae are given as follows:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{9}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{10}$$

$$\mathrm{AP} = \int_0^1 PdR \tag{11}$$

where TP represents the amount of accurately identified landslides, FP denotes the amount of incorrectly identified landslides, and FN represents the number of missed landslides. AP is the area under the *P*--*R* curve, which offers a comprehensive comparison of the Precision and Recall of the model. Greater values of *P*, *R*, and AP represent improved performance in detecting landslides.

In addition, the number of parameters (Params), floating-point operations (FLOPs), and frames per second (FPS) were used to measure the model's size and the requirement of

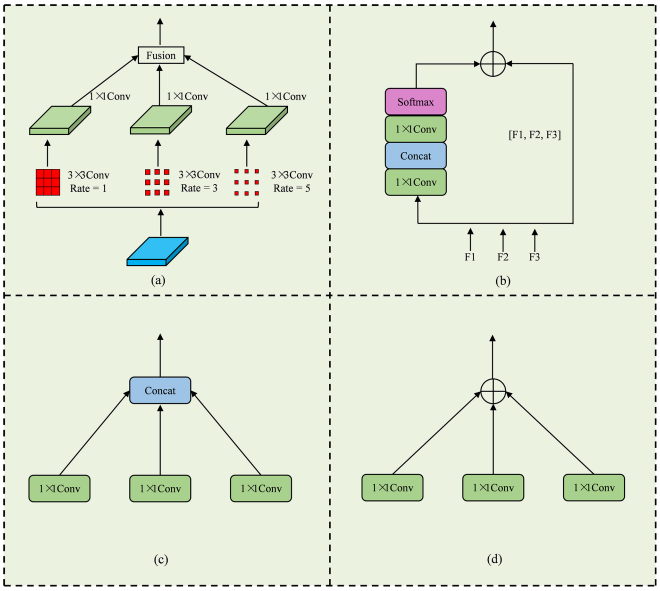

Fig. 13. (a) Structure of CEM. (b) Adaptive fusion. (c) Concatenation fusion. (d) Weighted fusion.

TABLE I
HYPERPARAMETERS OF OUR MODEL

Hyperparameters	Number	Hyperparameters	Number
lr0	0.01	fl_gamma	0.0
lrf	0.1	hsv_h	0.015
momentum	0.937	hsv_s	0.7
weight_decay	0.0005	hsv_v	0.4
warmup_epochs	3.0	degrees	0.0
warmup_momentum	0.8	translate	0.1
warmup_bias_lr	0.1	scale	0.9
box	0.05	shear	0.0
cls	0.3	perspective	0.0
cls_pw	1.0	flipud	0.0
obj	0.7	fliplr	0.5
obj_pw	1.0	mosaic	1.0
iou_t	0.20	mixup	0.1
anchor_t	4.0	copy_paste	0.1

TABLE II
DIVISION OF MSLD

Dataset	Train	Val	Test	Total
Number	6,062	674	2,887	9,623

hardware facilities. The Params refers to the total number of parameters required for model training, which measures the space complexity of the model. The FLOPs count the number of FLOPs performed by the model, measuring the time complexity of the model. The FPS represents the number of images that the model can process in a second, which is a metric to quantify the operational speed of the model.

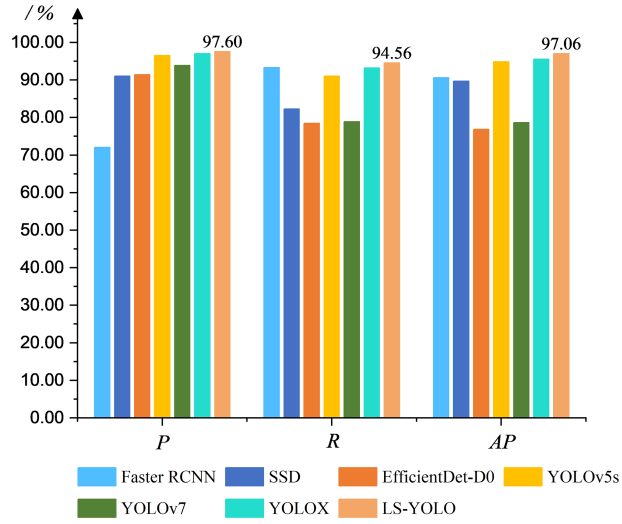

Fig. 14. Experimental results of different models.

B. Ablation Experiment

We have conducted ablation studies of the proposed model. Major experimental results are shown in Table III.

Adding the MSFE to the neck of YOLOv5s could improve the AP to 95.93%, which is an improvement of 1.05%. It proved that the designed MSFE is effectively improving the performance of YOLOv5s in landslide detection.

The successive 1×1 convolutions and Softmax in adaptive fusion result in a large number of parameters and computations. The concatenation operation in concatenation fusion is simple; however, the 1×1 convolution to ensure that the input and output feature map channels are consistent leads to extra parameters and computations. The structure of weighted fusion is simple with no extra parameters and little computation.

As shown in Table III, replacing the coupled head of YOLOv5s with the original decoupled head improves the *R* and AP by 1.79% and 1.05%, respectively, while keeping the *P* nearly unchanged. It demonstrated that decoupling the classification and regression tasks can improve the landslide detection performance of the model. Then, we applied the three different kinds of fusion techniques of CEM to improve decoupled head, respectively. Using adaptive fusion to improve the original decoupled head enhances the *P*, *R*, and AP by 0.25%, 0.42%, and 0.25% to 96.69%, 93.30%, and 96.18%. Using concatenation fusion to improve the original decoupled head enhances the *R* and AP by 0.99% and 0.63% to 93.87% and 96.56%; however, the *P* decreases by 0.15%–96.29%. Using weighted fusion to improve the original decoupled head achieves the best result. It enhances the *P*, *R,* and AP by 0.21%, 0.53%, and 0.36% to 96.65%, 93.41%, and 96.29%. These results demonstrated the effectiveness of using the three fusion techniques to improve the original decoupled head.

In addition, we conducted sufficient experiments to verify the compatibility between the improvements in this article. Using a combination of MSFE and improved decoupled head to improve YOLOv5s has a greater enhancement than each of the two modules is used separately. Using a combination of MSFE and

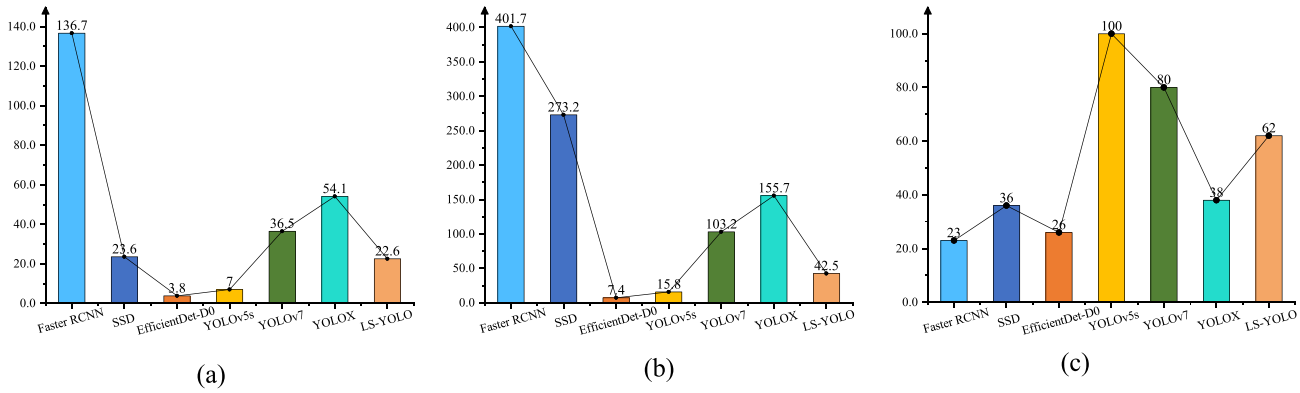

Fig. 15. Comparison of params, FLOPs, and FPS of different models. (a) Params comparison. (b) FLOPs comparison. (c) FPS comparison.

TABLE III
RESULTS OF ABLATION EXPERIMENTS

YOLOv5s	MSFE	O-DH	A-DH	C-DH	W-DH	*P*(%)	*R*(%)	*AP*(%)
✓						96.57	91.09	94.88
✓	✓					96.97	92.54	95.93(+1.05)
✓		✓				96.44	92.88	95.93(+1.05)
✓			✓			96.69	93.30	96.18(+1.30)
✓				✓		96.29	93.87	96.56(+1.68)
✓					✓	96.65	93.41	96.29(+1.41)
✓	✓	✓				96.35	93.38	96.27(+1.39)
✓	✓		✓			97.43	94.29	96.99(+2.11)
✓	✓			✓		97.41	94.44	96.96(+2.08)
✓	✓				✓	**97.60**	**94.56**	**97.06(+2.18)**

Notes: O-DH denotes the original decoupled head [42]. The improved decoupled head contains A-DH, C-DH, and W-DH. The fusion scenario of CEM in A-DH, C-DH, and W-DH is Adaptive Fusion, Concatenation Fusion, and Weighted Fusion, respectively.
The data of bold entities represent the optimal results.

TABLE IV
RESULTS OF MODEL'S COMPARISON EXPERIMENTS

Model	Precision(%)	Recall(%)	*AP*(%)	Params(M)	FLOPs(G)	FPS
Faster RCNN	72.09	93.33	90.62	136.7	401.7	23
SSD	91.07	82.30	89.71	23.6	273.2	36
EfficientDet-D0	91.46	78.49	76.86	**3.8**	**7.4**	26
YOLOv5s	96.57	91.09	94.88	7.0	15.8	**100**
YOLOv7	93.92	78.90	78.64	36.5	103.2	80
YOLOX	97.03	93.20	95.55	54.1	155.7	38
LS-YOLO	**97.60**	**94.56**	**97.06**	22.6	42.5	62

The data of bold entities represent the optimal results.

adaptive fusion improved decoupled head to improve YOLOv5s achieves the *P*, *R,* and AP at 97.43%, 94.29%, and 96.99%, respectively. Using a combination of MSFE and concatenation fusion improved decoupled head results in the *P*, *R,* and AP at 97.41%, 94.44%, and 96.96%, respectively. The improvement of adding MSFE and weighted fusion improved decoupled head into YOLOv5s at the same time is the most significant, and all the metrics are the best in ablation experiments. Compared with YOLOv5s, the *P* is improved by 1.03%–97.60%. *R* is improved by 3.47%–94.56%, and AP is improved by 2.18%–97.06%. The ablation experiments have provided evidence of the mutually reinforcing effect between MSFE and the improved decoupled head, which demonstrates the effectiveness of the work presented in this article.

C. Model Comparison Experiment

To conduct an impartial performance assessment of the proposed LS-YOLO for landslide detection, we compared it with other six widely used and recently proposed models:

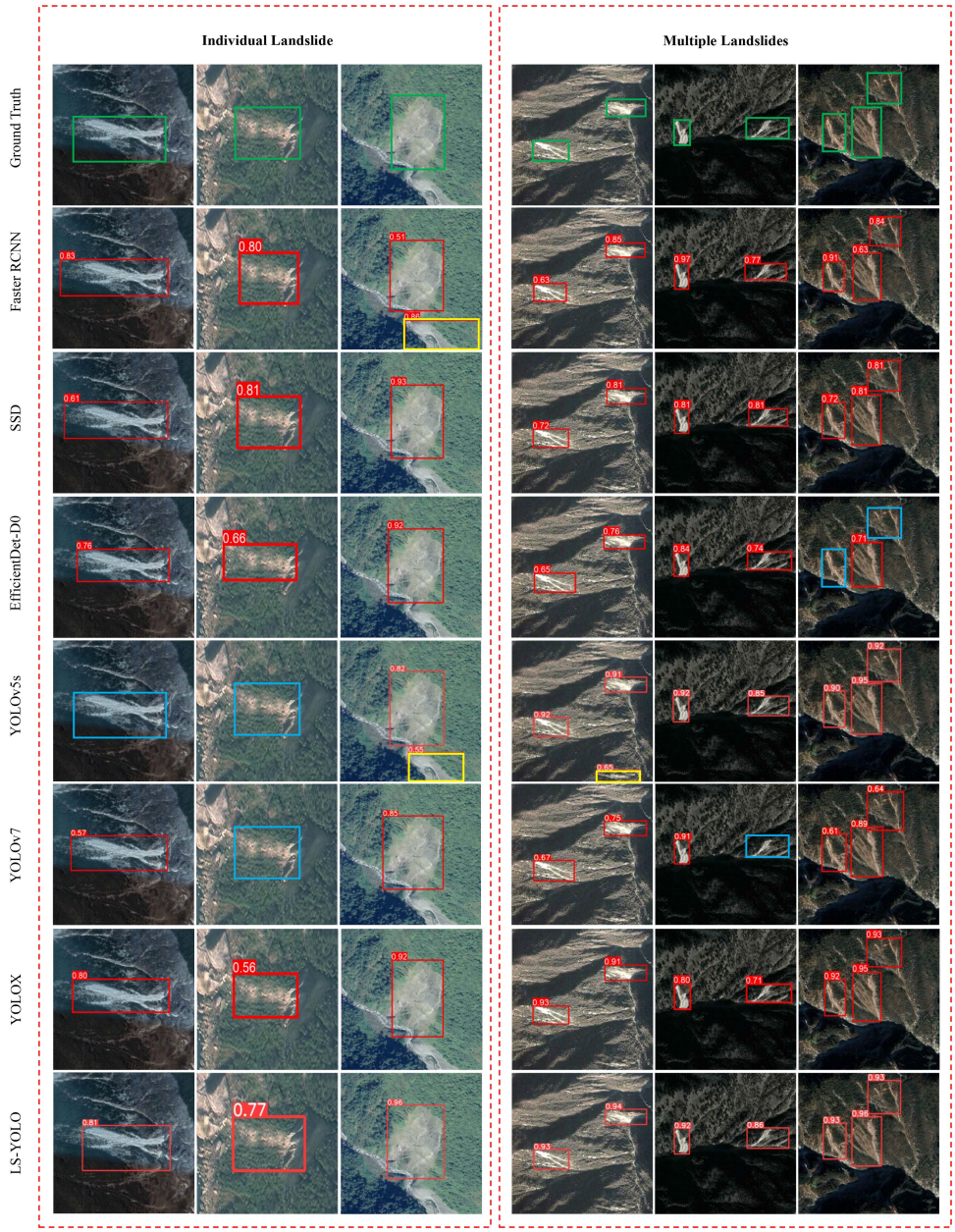

Fig. 16. Comparison of landslide detection results of different models.

faster RCNN, SSD, EfficientDet-D0, YOLOv5s, YOLOv7, and YOLOX. All seven models are trained for 300 epochs on the train set and tested on the test set of the MSLD. We used *P*, *R*, AP, Params, FLOPs, and FPS to evaluate the performance of models in landslide detection.

As shown in Fig. 14 and Table IV, the LS-YOLO achieves state-of-the-art performance in multiscale landslide detection. The *P* of LS-YOLO for landslide detection is 97.60%, which is 0.57% higher than the 97.03% of the second-place YOLOX. The recall of LS-YOLO in detecting landslides is 94.56%, which is 1.23% higher than the 93.33% of the second-place faster RCNN. The AP of LS-YOLO in detecting landslides is 97.06%, which is 1.51% higher than the 95.55% of the second-place YOLOX. Compared with the baseline model YOLOv5s, LS-YOLO improved the *P*, *R*, and AP in landslide detection by 1.03%, 3.47%, and 2.18%, respectively.

Due to the decoupled classification and regression tasks in the model, the LS-YOLO is more complex in structure and slower in running speed. As shown in Fig. 15 and Table IV, compared with the baseline model YOLOv5s, the parameters and FLOPs of LS-YOLO are increased from 7.0 M and 15.8 G to 22.6 M and 42.5 G, and the FPS of LS-YOLO is decreased from 100 to 62. However, LS-YOLO has a significant advantage in model size compared with faster RCNN, SSD, YOLOX, and YOLOv7. The running speed of LS-YOLO is just below YOLOv5s and YOLOv7, thus is more suitable for real-time landslide detection.

We noted some interesting results. The parameters and FLOPs of EfficientDet-D0 are 3.8 M and 7.4 G, respectively, which is

optimal among all the models. However, its running speed is the slowest with the FPS of 26. The main reason for this interesting phenomenon is that EfficientDet-D0 has high computational density and the need for frequent memory accesses. The model size of the two-stage model faster RCNN is large, with a much larger number of parameters and FLOPs than one-stage models. However, the precision of faster RCNN in landslide detection is just 72.09% and significantly lower than other models. Faster RCNN has lots of incorrect results in multiscale landslide detection. Thus, the detection accuracy and running speed of the model are not directly related to the model size.

Fig. 16 shows the comparison results of different models in landslide detection. The green boxes and red boxes denote the ground truth and the detection results of the models, respectively. Meanwhile, to achieve a more accurate analysis, we used blue boxes and yellow boxes to respond miss detected landslides and incorrectly detected landslides.

1) Individual Landslide Detection Results: As shown in the left half of Fig. 16, when faced with the presence of individual landslides in remote sensing images, faster RCNN and YOLOv5s have many misdetected results and easily misdetecting rivers, valleys, etc., as landslides. YOLOv5s and YOLOv7 have some missed results and poor performance in detecting landslides, making it difficult to effectively detect landslides. SSD, EfficientDet-D0 YOLOX, and LS-YOLO have fewer misdetected and missed results; however, the accuracy of landslides localization of SSD, EfficientDet-D0, and YOLOX is lower than our proposed model LS-YOLO.

2) Multiple Landslides Detection Results: As shown in the right half of Fig. 16, when faced with the presence of multiple landslides in remote sensing images, EfficientDet-D0 and YOLOv7 have a high number of missed landslides, and YOLOv5s has some misdetected landslides. Although faster RCNN, SSD, YOLOX, and LS-YOLO have fewer misdetected and missed landslides, the LS-YOLO is more accurate in landslide location than the other models.

Following the experimental analysis, it is evident that the proposed model LS-YOLO displays considerable advantage in both detecting individual landslide and multiple landslides in remote sensing images when compared with the existing models. The LS-YOLO achieves state-of-the-art performance in landslide detection that could fulfill the necessities of applied industrial demands.

V. Discussion

The main problem in landslide detection with remote sensing images is that landslides vary greatly in shape and size. It is difficult for general object detection models to effectively capture multiscale landslide features. In this article, we proposed a novel and high-performance LS-YOLO for landslide detection based on YOLOv5s. The main strength of LS-YOLO lies in its capability of capturing multiscale features of landslides through spatial separable convolution and dilation convolution. The proposed MSFE enhances receptive fields by successive spatial separable convolutions and improves the multiscale landslide detection accuracy of the model. The improved decoupled head

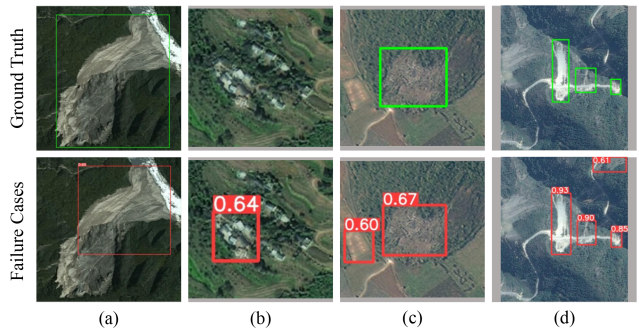

Fig. 17. Some failure cases using the LS-YOLO.

decouples the classification and regression tasks and captures multiscale context information by different dilated rate convolutions. The experimental results show that the proposed model LS-YOLO significantly improves the accuracy of landslide detection and location.

However, the current LS-YOLO still has some limitations. Compared with the baseline model YOLOv5s, the LS-YOLO is more complex in structure and slower in detection speed. The parameters and FPS of YOLOv5s are 7.0 M and 100, whereas LS-YOLO are 22.6 M and 62. Meanwhile, as shown in Fig. 17, LS-YOLO results in inadequate performance in detecting landslide under complex backgrounds. Especially with the interference of the river channels and houses, it produced some incorrect and missed results, and the accuracy of landslide localization is also low. In the future, we will focus on reducing the size and improving the robustness of our model in landslide detection.

Another limitation of the LS-YOLO is that it requires the creation of a fully annotated landslide dataset for the training process. Unfortunately, this is a costly and time-consuming endeavor that requires the expertise of specialists in remote sensing and landslide, and consumes considerable human and financial resources. To mitigate these costs, we plan to employ some semi-supervised learning techniques to our method in future work, which could help to transfer knowledge from limited labeled data to unlabeled data [44], [45].

VI. Conclusion

In this article, the MSLD was built using the Bijie landslide dataset together with 500 additional landslide images. Random seeds were used to promise each image randomly selected data augmentation techniques. The LS-YOLO was proposed to detect multiscale landslides with remote sensing images. The MSFE module based on ECA, average pooling, and spatial separable convolution is designed to improve the accuracy of the model in landslide detection. The decoupled head was improved using dilated convolution to enhance the precision of the model in landslide locations. Extensive experiments have demonstrated that our proposed model for detecting landslides performs better than the existing models. Compared with the baseline model YOLOv5s, due to the addition of the MSFE and the improved decoupled head, LS-YOLO detects and locates landslides more

accurately. In the future, we plan to propose a more efficient landslide detection model using lightweight modeling methods.

REFERENCES

[1] N. Casagli, E. Intrieri, V. Tofani, G. Gigli, and F. Raspini, "Landslide detection, monitoring and prediction with remote-sensing techniques," *Nature Rev. Earth Environ.*, vol. 4, no. 1, pp. 51–64, 2023.

[2] K. B. Sim, M. L. Lee, and S. Y. Wong, "A review of landslide acceptable risk and tolerable risk," *Geoenviron. Disasters*, vol. 9, no. 1, 2022, Art. no. 3.

[3] F. S. Tehrani, M. Calvello, Z. Liu, L. Zhang, and S. Lacasse, "Machine learning and landslide studies: Recent advances and applications," *Natural Hazards*, vol. 114, no. 2, pp. 1197–1245, 2022.

[4] 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, 2021, Art. no. e3998.

[5] Z. Lv, F. Wang, W. Sun, Z. You, N. Falco, and J. A. Benediktsson, "Landslide inventory mapping on VHR images via adaptive region shape similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Sep. 2022, Art. no. 5630211.

[6] Y. Ju et al., "Loess landslide detection using object detection algorithms in northwest China," *Remote Sens.*, vol. 14, no. 5, 2022, Art. no. 1182.

[7] J. E. Nichol, A. Shaker, and M.-S. Wong, "Application of high-resolution stereo satellite images to detailed landslide hazard assessment," *Geomorphology*, vol. 76, no. 1/2, pp. 68–75, 2006.

[8] Y. Chen et al., "Susceptibility-guided landslide detection using fully convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 998–1018, 2023.

[9] M. Galli, F. Ardizzone, M. Cardinali, F. Guzzetti, and P. Reichenbach, "Comparing landslide inventory maps," *Geomorphology*, vol. 94, no. 3/4, pp. 268–289, 2008.

[10] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal, "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection," *Remote Sens.*, vol. 11, no. 2, 2019, Art. no. 196.

[11] C. Niu, O. Gao, W. Lu, W. Liu, and T. Lai, "Reg-SA–UNet++: A lightweight landslide detection network based on single-temporal images captured postlandslide," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 9746–9759, Nov. 2022.

[12] W. Li, Y. Fu, S. Fan, M. Xin, and H. Bai, "DCI-PGCN: Dual-channel interaction portable graph convolutional network for landslide detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, May 2023, Art. no. 4403616.

[13] S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza, "Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking," *Nature Commun.*, vol. 11, no. 1, 2020, Art. no. 3952.

[14] M. A. Łukosz, R. Hejmanowski, and W. T. Witkowski, "Evaluation of ICEYE microsatellites sensor for surface motion detection—Jakobshavn glacier case study," *Energies*, vol. 14, no. 12, 2021, Art. no. 3424.

[15] S. O. Y. Amankwah et al., "Landslide detection from bitemporal satellite imagery using attention-based deep neural networks," *Landslides*, vol. 19, no. 10, pp. 2459–2471, 2022.

[16] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, 2021.

[17] 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, 2020.

[18] F. Catani, "Landslide detection by deep learning of non-nadiral and crowdsourced optical images," *Landslides*, vol. 18, no. 3, pp. 1025–1044, 2020.

[19] Y. Li, P. Wang, Q. Feng, X. Ji, D. Jin, and J. Gong, "Landslide detection based on shipborne images and deep learning models: A case study in the three gorges reservoir area in China," *Landslides*, vol. 20, no. 3, pp. 547–558, 2022.

[20] R. Fu et al., "Fast seismic landslide detection based on improved mask R-CNN," *Remote Sens.*, vol. 14, no. 16, 2022, Art. no. 3928.

[21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.

[22] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1440–1448.

[23] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. 28th Int. Conf. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[24] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2961–2969.

[25] W. Liu et al., "SSD: Single shot multibox detector," in *Proc. 14th Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.

[26] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 10778–10787.

[27] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[28] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.

[29] 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.*, 2023, pp. 7464–7475.

[30] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," 2021, *arXiv:2107.08430*.

[31] O. Ghorbanzadeh, H. Shahabi, A. Crivellari, S. Homayouni, T. Blaschke, and P. Ghamisi, "Landslide detection using deep learning and object-based image analysis," *Landslides*, vol. 19, no. 4, pp. 929–939, 2022.

[32] S. L. Ullo et al., "A new mask R-CNN-based method for improved landslide detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3799–3810, Mar. 2021.

[33] M. I. Sameen and B. Pradhan, "Landslide detection using residual networks and the fusion of spectral and topographic information," *IEEE Access*, vol. 7, pp. 114363–114373, 2019.

[34] L. Cheng, J. Li, P. Duan, and M. Wang, "A small attentional YOLO model for landslide detection from satellite remote sensing images," *Landslides*, vol. 18, no. 8, pp. 2751–2765, 2021.

[35] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, 2019, Art. no. 60.

[36] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao, and F. Shen, "Image data augmentation for deep learning: A survey," 2022, *arXiv:2204.08610*.

[37] 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.*, 2016, pp. 779–788.

[38] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of YOLO algorithm developments," *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, 2022.

[39] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 11534–11542.

[40] G. Song, Y. Liu, and X. Wang, "Revisiting the sibling head in object detector," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 11563–11572.

[41] Y. Wu et al., "Rethinking classification and localization for object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 10186–10195.

[42] C. Li et al., "YOLOv6: A single-stage object detection framework for industrial applications," 2022, *arXiv:2209.02976*.

[43] J. Xiao et al., "Tiny object detection with context enhancement and feature purification," *Expert Syst. Appl.*, vol. 211, 2023, Art. no. 118665.

[44] E. Protopapadakis, A. Doulamis, N. Doulamis, and E. Maltezos, "Stacked autoencoders driven by semi-supervised learning for building extraction from near infrared remote sensing imagery," *Remote Sens.*, vol. 13, no. 3, 2021, Art. no. 371.

[45] D. Hong, N. Yokoya, G.-S. Xia, J. Chanussot, and X. X. Zhu, "X-ModalNet: A semi-supervised deep cross-modal network for classification of remote sensing data," *ISPRS J. Photogramm. Remote Sens.*, vol. 167, pp. 12–23, 2020.

Wenjie Zhang received the B.E. degree in information security from Hainan University, Haikou, China, in 2020. He is currently working toward the Ph.D. degree in instruments science and technology with the School of Aerospace Science and Technology, Xidian University, Xi'an, China.

His current research interests include computer vision, object detection, and remote sensing.

Zhiheng Liu (Member, IEEE) received the Ph.D. degree in photogrammetry and remote sensing from the School of Geological Engineering and Geomatics, Chang'an University, Xi'an, China, in 2021.

He was a jointly educated student with Satellite Surveying and Mapping Application Center, China, from 2015 to 2016, and has been a visiting Ph.D. student with the University of Plymouth, Plymouth, U.K., since 2019. He has been a Postdoctor with the School of Aerospace Science and Technology, Xidian University, Xi'an, China, since 2021. His current research interests include remote sensing, computer vision, and deep learning.

Suiping Zhou received the B.E., M.E., and Ph.D. degrees in electrical engineering from Beihang University, Beijing, China, in 1989, 1992, and 1995, respectively.

He is currently a Full Professor with the School of Aerospace Science and Technology, Xidian University, Xi'an, China. Previously, he was a full Professor with the Department of Computer Science, Middlesex University, U.K., and an Assistant Professor with the School of Computer Engineering, Nanyang Technological University, Singapore. His current research interests include large-scale distributed interactive applications, parallel/distributed systems, and big data in space science and technology.

Wenjuan Qi received the B.S. degree in software engineering in 2021 from Xidian University, Xi'an, China, where she is currently working toward the M.S. degree in electronic information with the School of Aerospace Science and Technology.

Her research interests include time-series data forecasting and remote sensing.

Xinjun Wu received the B.E. degree in electronic science and technology from Yanshan University, Qinhuangdao, China, in 2022. He is currently working toward the master's degree in electronic information with the School of Space Science and Technology, Xidian University, Xi'an, China.

His research interests include object detection and image segmentation.

Tianyu Zhang received the Ph.D. degree in environmental geology from the School of Water and Environment, Chang'an University, Xi'an, China, in 2020.

He has been a Postdoctor with the School of Geological Engineering and Geomatics, Chang'an University, since 2020. His current research interests include neotectonics, tectonic geomorphology, and risk assessment of loess landslide.

Ling Han received the B.S. and M.S. degrees from Wuhan University, Wuhan, China, in 1991 and 1994, respectively, and the Ph.D. degree from Northwest University, Xi'an, China, in 2005, all in remote sensing.

Her research interests include the deep learning remote sensing image classification, registration, artificial intelligence, cloud detection in snow-covered images, and landslide interpretation in loess plateau areas.