J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 6373--6392

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

An Innovative Path for UAV Tilt Photography Image Processing Based on K-means Algorithm in Civil Engineering Disaster Management

Yurong Li[1], Shouwu Wang[2,3] ✉, Chunhua Han[4], Jingkai Meng[5], Bingqi Jiang[1]

[1] *College of Land and Resources Engineering, Kunming University of Science and Technology, Kunming, Yunnan, 650500, China*

[2] *City College, Kunming University of Science and Technology, Kunming, Yunnan, 650500, China*

[3] *Yunnan Provincial Department of Education Infrastructure Intelligent Operation and Maintenance Technology Innovation Team (Kunming University of Science and Technology), Kunming, Yunnan, 650500, China*

[4] *College of Transportation Engineering, Kunming University of Science and Technology, Kunming, Yunnan, 650500, China*

[5] *College of Architecture and Engineering, Kunming University of Science and Technology, Kunming, Yunnan, 650500, China*

ABSTRACT

Civil engineering disasters mostly occur in mountainous areas, and it is difficult to comprehensively monitor them using traditional technology, while this drawback can be avoided by utilizing UAV inclined photogrammetry technology. In this paper, with the support of the relevant experimental equipment, we obtain the images of civil engineering disasters with the help of this technology, and in order to avoid the influence of the interference factors in the images on the research results, we propose to use the K-means algorithm to pre-process the images. After completing the image processing, the improved YOLOV4 target detection algorithm is used to complete the design of the intelligent detection model of civil engineering disasters, and the processed images are input into the model for iterative training, so as to realize the intelligent management and early warning of civil engineering disasters. A region in Yunnan Province is taken as an example to explore and analyze the example. As of 2022, it is found that 180 landslides actually appeared in the region, while the model detected 172 landslides, resulting in the model's civil engineering disaster detection accuracy of 95.56%, which is within the permissible range, proving that the model has a good application efficiency, and can provide certain help and innovative guidance for the relevant units of civil engineering disaster management.

Keywords: K-means algorithm, UAV, YOLOV4, disaster detection model

✉ Corresponding author.
E-mail address: wangshouwu2024@163.com (S. Wang).

Received 20 February 2024; Revised 15 April 2024; Accepted 20 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-349

1. Introduction

China's mountainous areas are characterized by a wide range of topographic variations and complex geological conditions, and geologic disasters of different types and grades occur from time to time, threatening the personal and property safety of nearby residents [1]. With the construction of urban infrastructure and transportation networks, the development of mineral resources, etc., a large amount of mountainous land is excavated, especially the construction of highways adjacent to rivers, which makes the original stable rock and soil bodies gradually in an unstable or stable state, and is prone to disasters such as collapses, landslides, or mudslides under the conditions of heavy rainfall or earthquakes [2-5]. The continuous improvement of China's regional transportation network has promoted the rapid development of regional economy, but a large number of adjacent high and steep slopes have been formed during the construction of highways and railroads, which is also an important factor in inducing larger-scale landslides [6-8]. Therefore, how to quickly and effectively monitor all kinds of geologic hazards in civil engineering is an urgent problem to be solved at present.

In the past, the traditional way of statistics is to climb to the corresponding position to use the geological compass, tape measure and other tools for manual contact measurement, statistical efficiency is extremely low and dangerous, at the same time, for some of the personnel can not climb the high and steep slopes, artificial can not complete the measurement [9-11]. With the progress of science and technology, cross-penetration of various disciplines, a variety of non-contact measurement techniques and methods came into being, and achieved good results in practical engineering applications. Compared with the traditional contact measurement method, the non-contact measurement method has the obvious advantages of safety, efficiency, intelligence and accuracy [12-14]. And the common mainstream non-contact measurement methods include close-up photogrammetry, three-dimensional laser scanning technology and so on [15-16]. However, the digital close-up photogrammetry and 3D laser scanner equipment is expensive, the equipment itself has no positioning system, the measurement accuracy needs to be guaranteed through the additional deployment of control points, while the operation is relatively complex, there are certain limitations in the practical use [17-20].

For the above disadvantages and limitations of the information acquisition technology of rock structure surface, the continuous optimization of small UAVs and the continuous improvement of the supporting picture post-processing technology provide new ways and methods for the acquisition of rock structure surface information [21-24]. UAV tilt photogrammetry is a multi-angle and multi-directional photographic camera system carried on the UAV platform, i.e., this technology realizes the purpose of acquiring image data of the mapping area from multiple directions [25]. Compared with the initial vertical photography, tilt photography can show the height difference between the surveyed object and the area, making the captured image more authentic and intuitive [26-27]. However, due to the differences in the angle and other differences in acquiring data from different directions, it puts higher requirements on image fusion and data processing techniques [28-29].

In this paper, we select the experimental equipment of this research, use the UAV tilt photogrammetry technology to collect the images of civil engineering disaster related information, and process the collected images under the theoretical support of K-Means algorithm, so that the images reach the model standard requirements. Through the description of the application of UAV inclined photogrammetry technology in civil engineering disaster management, the civil engineering disaster intelligent detection model based on IYOLOV4 target detection algorithm is introduced, and the processed images are taken as inputs and put into the model for training and detection to achieve the purpose of intelligent detection of civil engineering disasters. A region in Yunnan Province is identified as the study area, relevant parameters and evaluation indexes are set, and the UAV tilt camera image processing and civil engineering disasters are analyzed by examples.

2. UAV tilt-shot image processing based on K-means algorithm

2.1. Experimental equipment and image acquisition

2.1.1. Experimental equipment. The image acquisition equipment is DJI UAV-Elf series-3Profes-sional, the weight of the vehicle is 1500g, the wind speed is 2m/s or less flight, the body of the vehicle is equipped with a GPS module and a visual positioning system, the visual positioning system altitude measurement range is from 50cm to 500cm, and the GPS module needs a good signal from the GPS satellite.

The camera is equipped with 1/2.3-inch CMOS image sensor, effective pixels 16.8 million, CMOS size ratio of 6:4, sensor width is 5.571mm, height is 4.167mm. The lens is FOV 87°, 42mm (40mm format equivalent), f/3.6, focusing at infinity. The maximum resolution of the photos is 8000×6000.

The image processing hardware platform was a laptop computer, model Dell N5110, with an Intel Core i5-2410M CPU@2.8GHz processor, 12.00 GB of operating memory, and Windows 10 as the operating system.The image processing and analysis software platform was MATLAB 2020b, and the software for manually segmenting the images was Photoshop 2016.

2.1.2. Image acquisition. UAVs can photograph civil engineering disasters at different angles and heights, capture the details of civil engineering disasters, better display the three-dimensional form of buildings, and provide more comprehensive data support for civil engineering disaster management. Therefore, compared with traditional measurement methods, UAV tilt photography technology can realize more accurate and efficient image acquisition. In the process of image acquisition, the following contents need to be paid attention to:

1) Take the civil engineering disaster as the center, and fly the photography at low speed and close distance under the premise of ensuring safety.

2) Adjust the attitude of the camera in real time during the process of UAV tilt photography to ensure that the acquired civil engineering disaster images contain different angles and positional information.

The UAV tilt photography process is affected by many factors, including equipment factors and environmental factors, etc., and the image degradation phenomenon is likely to occur during the above acquisition process, and the image clarity is low. For this reason, K-Means algorithm is used to process the collected images to ensure the rigor and effectiveness of the research results.

2.2. Image preprocessing based on K-means algorithm

2.2.1. K-means algorithm. The K-mean algorithm is a classical clustering algorithm, the algorithm starts with the number of clusters K set artificially by the algorithm writer or user, the final result of clustering is to determine the location of the K cluster centers of mass and form K clusters, the algorithm minimizes the similarity metric between the points (data) within each cluster, and the similarity metric between clusters is as large as possible [30]. After setting the value of K, the initial K center-of-mass positions are randomly generated, and appropriate distance measures such as Euclidean distance, Manhattan distance, Minkowski distance, etc. are traversed to calculate the distance from each point to each center-of-mass, and in accordance with the principle of distance minimization, each point is classified into the center-of-mass with which it is closest to, so that each point is assigned to a cluster. At this point, the center of mass may not be the optimal position, so it is necessary to constantly update the center of mass of each cluster in order to achieve the purpose of optimization. The method of updating the center of mass position is to find the mean value of each point in the cluster, and the position where the mean value is located is the new center of mass position of the cluster. After constantly seeking the mean value to continuously update the cluster position, until the center of mass of the cluster does not change the magnitude of change to meet the optimization criteria, the algorithm ends and the final clustering results are formed.

K The distance metric in the mean clustering algorithm can be chosen in various forms, such as Euclidean distance, Manhattan distance, Minkowski distance, etc., using different distance metrics, only in the calculation of the distance of that step of the calculation method is different, the rest of the algorithmic steps are the same [31]. Which distance method is used in practical applications should be selected according to the characteristics of the actual application.

The K-mean clustering algorithm to optimize the index of the center of mass position uses the intra-cluster error variance (SSE). For the specified k clusters, the smaller the distance between the intra-cluster points, the corresponding SSE is correspondingly smaller, and the clustering is more effective.

Let the data set be $A=\{a_1,a_2,\dots a_i\}$, where $a_1,a_2\dots a_i$ denotes each of the i data records in the data set. The initial center of mass is randomly generated for the first time, denoted by $K_1=\{K_1,K2_1,\cdots Kj_1\}$, where Kj denotes the j th center of mass and Kj_1 denotes the point where Kj is located for the first time. After the center of mass is generated, traverse from a_1, a_2 to a_i to calculate their respective distances to $K1_1,K2_1,\dots Kj_1$, and judge a_i which distance from Kj_1 is the closest to divide a_i into the cluster centered on that Kj_1. Calculate the SSE_1 of the first time, as in equation (1), which is the sum of the squares of the distances a_i from the center of mass Kj_1 that one is at:

$$SSE_1=\sum\nolimits_{m=1}^{j}\sum\nolimits_{a_{m1}\in Km_1}\left\|a_{ml}-Km_1\right\|^2 \tag{1}$$

Where SSE_1 denotes the first SSE, m denotes the variable representing the number of centers of mass at the time of programming, which takes integer values from 1 to j, a_{m1} denotes all the data belonging to the mth center of mass after the first center of mass is determined, Km_1 denotes the mth center of mass determined for the first time, and $a_{m1}\in Km_1$ denotes that a_m is classified into clusters centered on the center of Km_1.

The mean value of the points within each cluster is calculated and it is taken as the new center of mass K_2:

$$K_2=\{KI_2,KK_2,\dots Kj_2\} \tag{2}$$

$$Kj_2=mean\left(a_{m1}\right) \tag{3}$$

The $mean(\ )$ in Eq. (3) represents the mean value function by which the center of mass K_2 of the second time can be determined, and the data in A are reassigned to the new clusters according to the principle of closest proximity to calculate the SSE_2 of the second time:

$$SSE_2=\sum\nolimits_{m=1}^{j}\sum\nolimits_{a_{m2}\in Km_2}\left\|a_{m2}-Km_2\right\|^2 \tag{4}$$

Calculate the gap ΔSSE_2 between SSE_2 and SSE_1. ie:

$$\Delta SSE_2=\left|SSE_1-SSE_2\right| \tag{5}$$

To determine whether ΔSSE_2 meets the error requirement, set δ as a set positive number, which is the threshold of the error judgment and is determined according to the actual situation [32]. If ΔSSE_2 is less than δ, i.e:

$$\Delta SSE_2<\delta \tag{6}$$

Then the error of clustering is considered to satisfy the accuracy requirement and the algorithm stops. If it does not meet the requirements, then calculate the mean value of the points in each cluster with K_2 as the center of mass to form a new center of mass K_3, according to the principle of closest distance to the data in A re-assigned to the new clusters did not calculate the 3rd SSE_3, calculate ΔSSE_3 and determine whether to terminate or loop again until the end of the algorithm.

2.2.2. Original image and k-value determination. A high-precision UAV was used to photograph civil engineering disasters in Yunnan region, and the acquired images were used as the original data source, and the ground resolution of the aerial film was 0.04 m. From the viewpoint of the image data acquired by using the UAV, for the same batch of UAV tilt-photography image dataset, the flight altitude of the UAV is relatively fixed, and the area covered by a single image is limited in scope, and the number of categories contained in each image is is relatively small and basically equal, so only one k value is required to be calculated for processing the same batch of image data. Through a large number of tests to verify the UAV image dataset acquired this time, the corresponding Silhouette index value is calculated according to Eq. (7), and according to the relationship graph between the number of clusters and the Average Silhouette index value, as shown in Fig. 1, it is possible to determine an optimal number of clusters k of 8. Therefore, when using K-means for the image clustering, it will not fall into local optimization and improve the efficiency of subsequent processing. That is:

$$S_i = \frac{b_i - a_i}{\max\left(a_i, b_i\right)} \tag{7}$$

Where: a_i is the degree of difference between point i and the cluster to which it currently belongs, usually measured by the average Euclidean distance to each point. b_i is the minimum value of the average distance between point i and other points in different clusters. The value of S_i is in the range of $[-1,1]$. The larger the value of S_i (the closer it is to 1), the higher the degree of similarity between point i and points in the same cluster, and the higher the degree of dissimilarity with points in different clusters, i.e., the better the classification result.

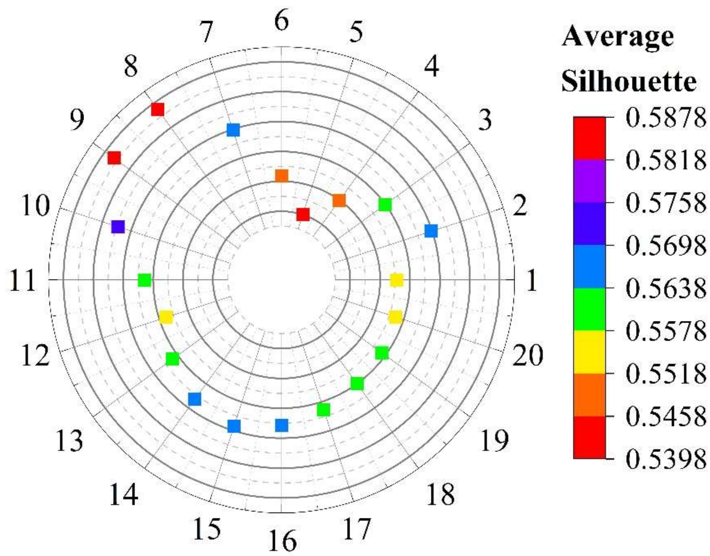


Fig. 1. Relationship between the number of clusters and the Average Silhouette value

2.2.3. K-means clustering initial segmentation. First, the original image is initially segmented using K-means clustering. According to the optimal number of clusters determined in section 2.2.2, the initial clustering value is set to 7. According to the visual judgment results of the combination, a new processing object can be obtained. In the image processing, the "keying" operation is used to eliminate the "interference information" in the new object to get the complete study area and complete the initial segmentation guided by K-means clustering.

2.2.4. Threshold Segmentation and Optimization. After processing to get the complete region, the image needs to be threshold segmented and further optimized to generate a binary image. Threshold segmentation is performed using the visual image threshold selection GUI tool, and the thresh_tool function automatically calculates an adaptive threshold first, which usually segments the image well. If you are not satisfied with the result, you can also move the threshold gray vertical line in the gray scale histogram left and right to adjust the threshold until the best segmentation results. As shown in Fig. 2, the threshold value for segmentation is 50. UAV tilt camera images have high resolution, and it is difficult to obtain accurate boundaries when using K-means clustering for the initial segmentation of two intermingled images, and the binary images obtained after threshold segmentation often contain background noise. Median filtering is a nonlinear operation, and median filtering of the binary image using the medfilt2 function filters out the peppercorn noise from the image while preserving the edges.

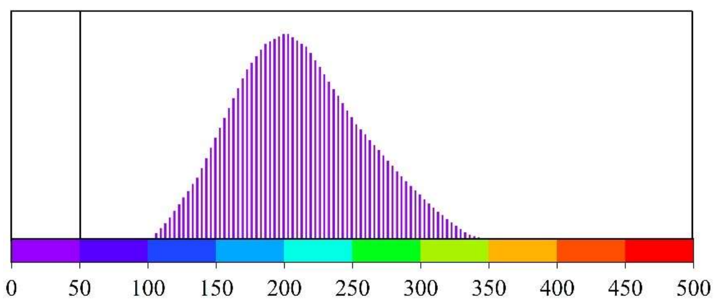


Fig. 2. Image threshold Select GUI tool

2.2.5. Marking and merging. The bwlabel function is used to label each separated part of the binary image, returning a labeling matrix L with the same size as the binary image, containing the category labels of each connected region in the binary image. After the K-means initial segmentation and threshold segmentation of the object contains only one kind of features, so there is only one kind of category label in the label matrix L, and its corresponding label value is 1. Using the label2rgb function, the label matrix L can be converted into a pseudo-color image, which completes the identification and extraction of labels. Repeat the above steps, respectively, to complete the processing of the remaining K-means clustering results, and finally identify seven categories of features from the original image. The 7 types of feature labels are merged to get the color label matrix, and the original remote sensing image is superimposed with the color label matrix as the final image classification result.

3. Civil engineering disaster management inquiry

3.1. Application of drone tilt-shift photography in disaster management

UAV tilt-shift photography technology is now widely used in the inspection field, and its application in civil engineering disaster detection has many advantages. In this subsection, the application of UAV tilt photography in disaster management will be illustrated from three aspects: bridge, slope, and underground engineering disasters.

3.1.1. Application in bridge inspection. Bridges belong to large civil engineering structures and are critical nodes in lifeline projects. The parameters for bridge structural health testing can be roughly categorized into three types: vehicle load information (traffic density, real-time location distribution of vehicles, load, etc., environmental load information (wind load, temperature load, seismic load, etc.), and structural information (frequency, period, damping ratio, etc.). Specific technical route: firstly, analyze the structural characteristics of the bridge, select and arrange the location of the measurement points according to the structural weak points, construction difficulty, component importance and other factors, then acquire the bridge image with the help of UAV tilt photography technology, and segment the image by K-Means algorithm, and finally identify the processed image with the help of deep learning algorithm. The potential safety hazards appearing in the bridge can be detected at an early stage, and repairs can be carried out to reduce unnecessary losses.

3.1.2. Application in slope stability testing. Slope is one of the most common forms of engineering in engineering construction. Landslides, slides, subsidence, mudslides, rockfalls, etc. produced by destabilizing damage of slopes, which on the surface seem to be different manifestations of slope rock movement, may bring serious damage or even disaster at any time. UAV tilt camera technology has many applications in slope stability: using the technology to detect the lateral deformation of soft ground embankment, distributed fiber grating test technology to measure the displacement of piles, detecting the stress condition and slope displacement of soil nails and anti-slip piles, and highway slopes for stability analysis. Specific technical route: firstly, obtain the slope image, secondly, pre-process the interference information in the image, and finally summarize to the recognition model for data analysis.

3.1.3. Application in underground engineering disaster detection. In the construction process of civil engineering, underground projects are difficult to construct due to their complex hydrogeological conditions, and are prone to high-risk geologic disasters such as sudden water surges and landslides, which cause serious casualties and economic losses. The technology and the deep learning algorithm recognize the underground engineering disaster warning information so that the construction personnel have enough time to prepare for the possible geological disasters. This technology will be one of the important development directions of civil engineering disaster detection, and has a broad application prospect. The development of UAV inclined photography technology in civil engineering disaster detection technology has great scientific significance and social benefits.

3.2. Intelligent Detection Model for Civil Engineering Hazards

On the basis of the theoretical analysis of image preprocessing and UAV tilt photography application in Chapter 2, the task of constructing an intelligent detection model for civil engineering disasters is realized with the help of the improved YOLOV4 target detection algorithm in the deep learning algorithm. The detailed design process is as follows:

3.2.1. YOLOV4 target detection algorithm. The fourth generation algorithm YOLOv4 released in 2020 further balanced the detection speed and detection accuracy, the authors based on a large number of deep learning designs and techniques developed by the previous generation of YOLOv3 conducted multiple experiments on the YOLOv3 will have gained improvements retained after the integration and innovation of the YOLOv4 algorithm proposed [33]. Taking the input processed UAV tilt camera image size 608*608 pixels as an example, the network structure is shown in Figure 3, and the specific improvements in the network structure are as follows:

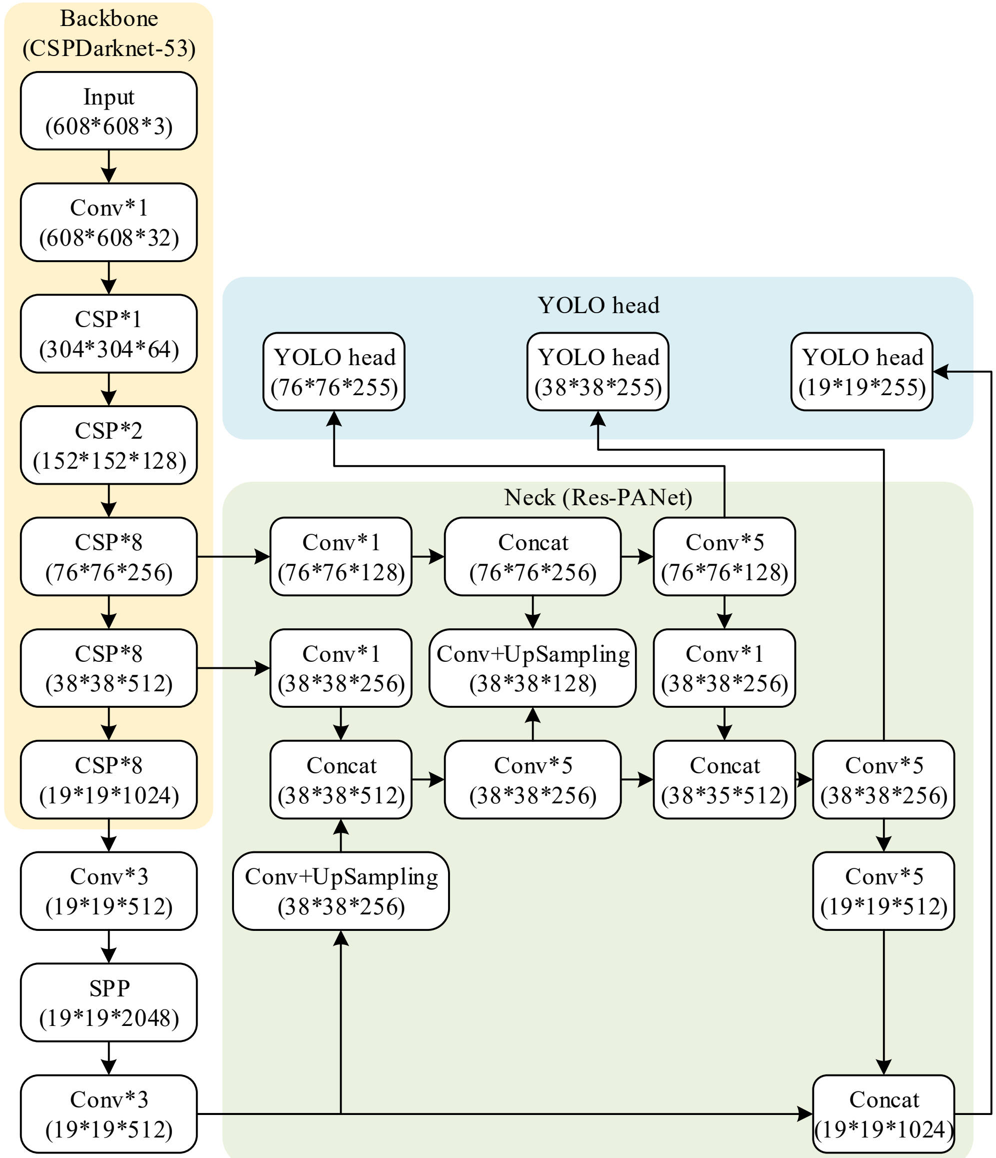


Fig. 3. YOLOv4 network structure

1) CSPDarknet-53. Drawing on the idea of CSPNet (framework segmental partial network), the CSP-Res block is designed on the basis of the residual block, which is divided into two paths after the first convolution, respectively, through the residual unit and the convolutional layer and then do the stacking operation. One of the reasons for the large amount of computation of the unified convolutional neural network is that there is a large amount of repeated gradient information in the convolutional layer in the process of extracting features, and the design of the cross-stage part can save the gradient changes within the module to be stacked into a feature map in order to reduce the amount of computation and to improve the ability of the network to fit. In addition, the authors introduced SPP (Spatial Pyramid Pooling) as an additional module of CSPDarknet-53, after the 19*19*512 effective feature layer is input to the SPP module by three convolution operations, the SPP module is stacked after the maximum overlap pooling of the feature maps of four different scales, respectively.The SPP module can convert feature maps of arbitrary size into feature vectors of set size. Experiments show that the SPP module is conducive to the convergence of the network and mitigation of overfitting.

2) PANet (Path Aggregation Network). PANet carries out more complex feature extraction compared to FPN, and the most important feature is the repeated extraction and fusion of different feature layers. From the output, the 19*19 enhanced feature layer comes from the 19*19 effective feature layer after

two up-sampling and stacking operations, then two down-sampling and stacking and finally stacked with itself, similarly the other two enhanced feature layers also come from multiple up- and down-sampling and feature fusion.PANet strengthens the feature extraction capability by increasing the information transfer between the deeper and shallower network layers.

3.2.2. Improvements based on the attention mechanism module. The Attention Mechanism Module (CBAM) consists of two parts, the Channel Attention Module (CAM) and the Spatial Attention Module (SAM), and experiments have shown that channel-weighting followed by spatial weighting of the input feature maps improves the model more than other combinations.
1) Channel Attention Module. The channel attention module weights the channel dimensions of the feature map, and automatically optimizes the weights of each channel in the feature map through the back-propagation algorithm, which greatly improves the sensitivity to the effective features and suppresses the irrelevant information under the condition of increasing the number of very small parameters, which has an overall gain effect on the model. Take the input feature map $F \in R^{C \times H \times W}$ as an example, the channel attention does the following to the feature map:

(1) The input feature maps are subjected to global maximum pooling and global average pooling to obtain two $1*1*C$ feature maps $Max\ Pool(F)$ and $Avg\ Pool(F)$, respectively.

(2) The feature maps $Max\ Pool(F)$ and $Avg\ Pool(F)$ are subjected to multilayer perceptron (MLP) to obtain two $1*1*C$ -channel attention feature maps $MLP\left(Max\ Pool(F)\right)$ and $MLP\left(Avg\ Pool(F)\right)$, respectively, and the number of neurons in the hidden layer is C / r in order to further reduce the number of parameters, and r is the compression ratio, which is a hyperparameter.

(3) The two channel attention feature maps are subjected to a summing operation, which is performed by the Sigmoid activation function to generate $1*1*C$ channel attention weight $M_c(F)$.

(4) Generate F' by multiplying $M_c(F)$ with each element of the input feature map F, i.e., F' is the result of weighting each channel in the original feature map F.

2) Spatial attention module. The design idea of the spatial attention module is to weight each spatial location in the feature map to enhance the sensitivity of the model to the location of the target in the image. Taking the output of channel attention $F' \in R^{C \times H \times W}$ as an example, spatial attention does the following to the feature map:

(1) The feature map F' is first subjected to global maximum pooling and global average pooling in the channel direction to obtain two $H \times W$ -dimensional feature maps $Max\ Pool\left(F'\right)$ and $Avg\ Pool\left(F'\right)$.

(2) After stacking the two channels to generate a feature map of size $H \times W \times 2$, a 7*7 convolution operation is performed to change the original number of channels to 1, thus obtaining a 2D spatial attention feature map $M_s(F)$, and the value of each position in the map is the weight of the feature at that position in the image.

(3) Finally, multiply F' with the corresponding position in $M_s(F)$, and the output is the feature map processed by the attention mechanism module.

In the clustered UAV inclined photographic images, weathering spalls and rolling stones are often located in the lower part of the image and have regular shapes, and landslides and rainfall erosion are often located in the center of the image and occupy a large area, so for the unevenly distributed small targets, the spatial attention module can enhance the model's ability to learn the location of the target. From the above characteristics, the introduction of CBAM is theoretically compatible with the characteristics of the detected targets, therefore, this paper designs to insert CBAM into each of the three effective feature layers after the YOLOv4 backbone feature extraction network to improve the detection effect of the model.

3.2.3. Improvement based on feature map residuals. Since shallow neurons have smaller receptive fields, the model's ability to detect small targets can theoretically be improved by enhancing the influence of shallow neurons on the results. Therefore, in this paper, we design new residual edges for shallow feature compensation based on the original PANet. This is done by multiplying coefficient ω_1 of the 76*76*256 feature layer by the output of the three CBAMs described in the previous section and doing a sum operation with the output of the feature fusion, multiplying coefficient ω_2 of the 38*38*512 feature layer by the output of the first feature fusion, and then multiplying coefficient 3 of the 19*19*512 feature layer by the 19*19*1024 firstly by the 19*1 convolution expansion channel, and then multiplying coefficient ω_3 by the output of the feature fusion, where the residual edges are designed based on the original PANet. The output after fusion is summed, where 01, w2, u3 are the learning parameters. To distinguish the original PANet, the improved feature fusion network is named as Learning Res-PANet.

3.2.4. Improvement of the detection header coding method. The input of the detection head part is the three reinforced feature layers reinforced by the feature pyramid, and the detection head in YOLOv4 follows the design of the previous version, i.e., the three sets of reinforced feature maps are convolved once with 3*3 convolution and once with 1*1 convolution and the final output is the tensor which contains the prediction and regression information. Therefore, in this paper, the original YOLO head is modified as follows, taking 19*19*512 feature layer as an example:

1) Firstly, integrate the number of channels to 256 by 1*1 convolution layer, i.e. output 19*19*256 feature map.

2) The output of the upper layer is divided into two groups in parallel by two 3*3 convolution operations to further integrate the features.

3) The first group of feature maps then performs 1*1 operation for channel integration to output all the category information, and the second group of feature maps again in parallel by 1*1 convolution to do channel integration to output location information and confidence level.

In order to differentiate from the original YOLO head, YOLO4 head is used in the following paper to denote the improved detection head part.

3.2.5. Overall design of the detection modeling framework. Based on the improvements described above, this paper proposes an improved target detection algorithm based on YOLOv4, which is named IYOLOv4 (Improved YOLOv4) for easy differentiation, and the network structure of IYOLOv4 is shown in Fig. 4.

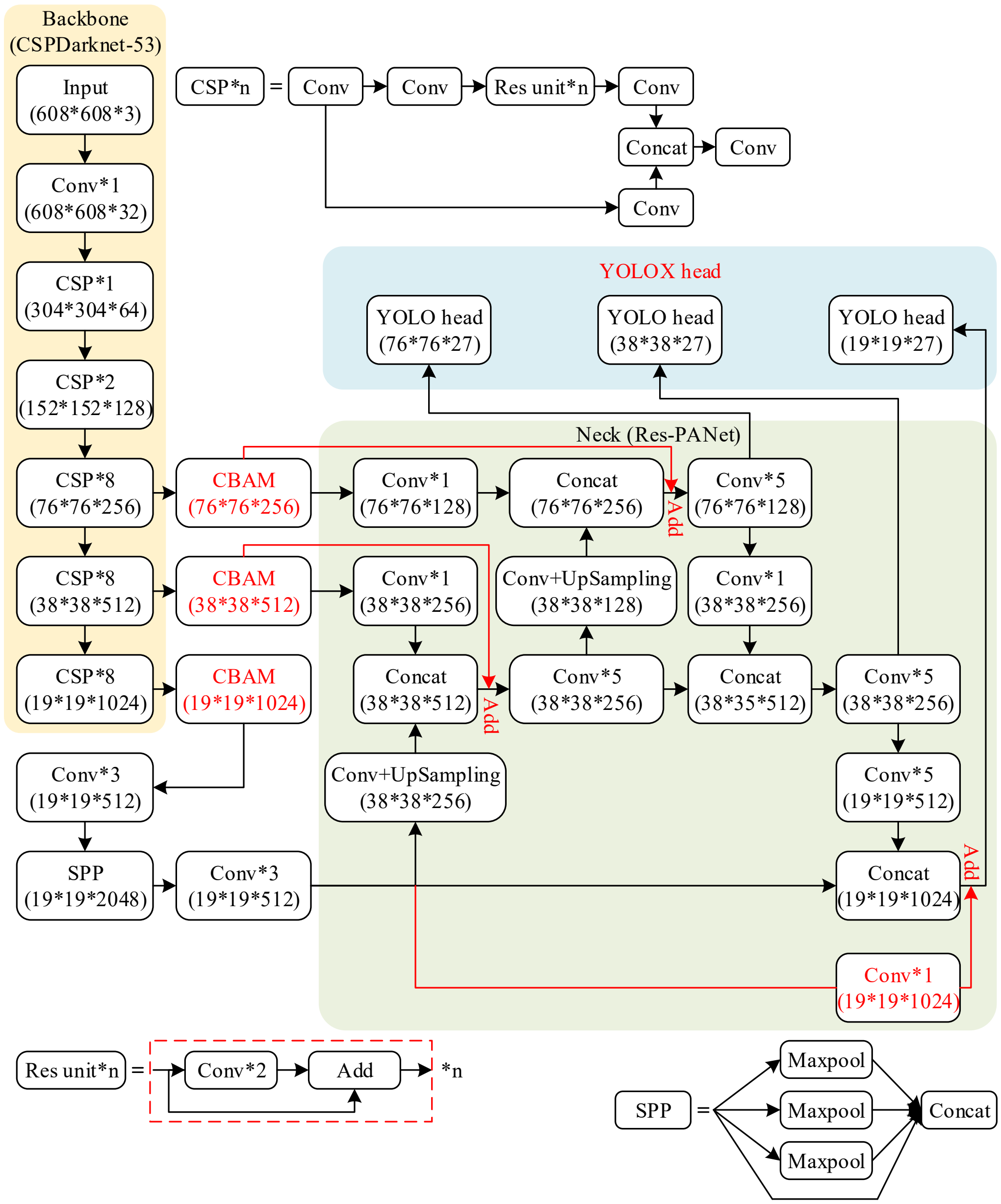


Fig. 4. Network structure of IYOLOv4

4. Image pre-processing and disaster management research and analysis

4.1. Image Preprocessing Research Analysis

4.1.1. Evaluation indicators. Evaluation measures are mathematical indicators of the performance of an algorithm. For this purpose, matching rate, misclassification rate and several metrics indicating the accuracy of image segmentation are established.

1) Matching rate and misclassification rate. The matching rate and misclassification rate are calculated by the formula:

$$\text{District allocation rate} = \frac{P_i - |P_i - Q_i|}{|Q_i|} \times 100 \tag{8}$$

$$\text{Error score rate} = \frac{|P_i - Q_i|}{|Q_i|} \times 100\% \tag{9}$$

Where, P_i is the target pixel value obtained by the segmentation algorithm, and Q_i is the actual target pixel value, where, Q_i is obtained by Photoshop manual standard segmentation.

2) Accuracy index. Accuracy refers to the degree of consistency between the algorithm segmentation results and the real segmentation results, which can be measured based on region (area or volume). Assuming that an image contains only one target region and the rest is the background, and let the set S represent the collection of pixels in the target region obtained by the segmentation algorithm, the set T represent the collection of pixels in the real segmentation target region, and I represent the collection of pixels in the whole image, then the following four measures are defined:

$$TP = S \cap T \tag{10}$$

$$FP = S - T \tag{11}$$

$$FN = T - S \tag{12}$$

$$TN = 1 - T - S \tag{13}$$

Where: FP is the set of pixels obtained by different segmentation algorithms for segmentation, FN is the set of pixels obtained by the PS manual labeling segmentation algorithm, TP is the set of pixels produced by the intersection of the 2 sets of FP and FN, and TN is the set of pixels obtained by the whole image I after removing FP and FN. On this basis the following 4 ratios are defined:

$$TPF = \frac{|TP|}{|T|} \tag{14}$$

$$FNF = \frac{|FN|}{|T|} \tag{15}$$

$$FPF = \frac{|FP|}{|1-T|} \tag{16}$$

$$TNF = \frac{|TN|}{|1-T|} \tag{17}$$

Where TPF is the ratio of the number of true positive target pixels to all correct pixels in the target region, FNF is the ratio of the number of miss-segmented target pixels to all correct pixels in the target region. FPF is the ratio of the number of incorrectly segmented background pixels into target pixels to the correct pixels in the background, and TNF is the ratio of segmented background pixels to the correct number of background pixels.

4.1.2. Data analysis:

1) Image color distribution. With the continuous improvement of computer processing technology, people are paying more and more attention to the segmentation of UAV tilt camera images, making full use of the color information of the image for image segmentation, which can be achieved by using fuzzy class methods, clustering, and edge detection to achieve the segmentation of colorful graphics. This section focuses on the automatic segmentation of images using K-Means to complete the clustering of color graphics. In order to reduce the computational complexity, the image can be processed using the RGB color space. The color values are analyzed by means of sampling points, and the distribution of sampling points in the image is observed, and the color values of sampling points are shown in Figure 5. The data performance in the figure shows that the image color mean values are 127.06, 91.62, and 76.51, which indicates that the target color value and the background color value have a large difference in the R channel, and the difference is not obvious in the G channel and B channel.

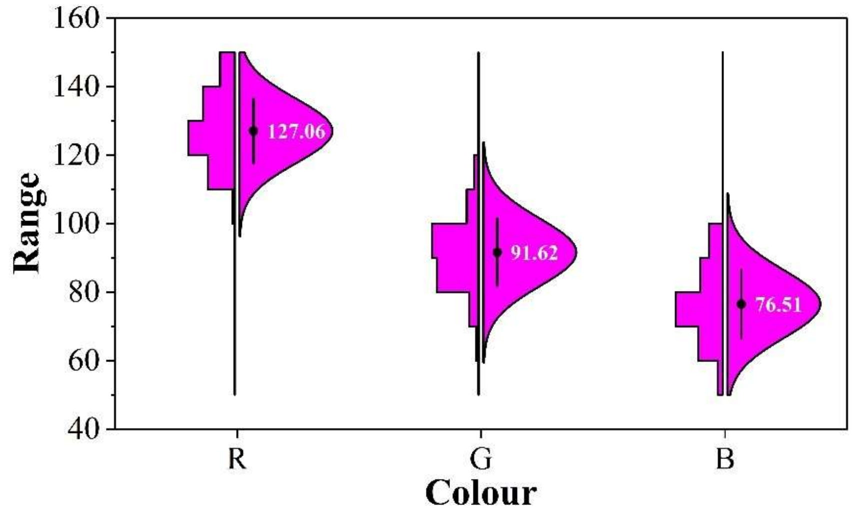


Fig. 5. Color value of the sampling point

2) Comparative analysis of iteration number and segmentation time. The following segmentation experiments of 7 groups of images (civil engineering disasters: flood X1, inundation X2, mudslide X3, bridge collapse X4, tunnel collapse X5, river dam damage X6, landslide on the top of the mountain X7) are conducted to validate the effectiveness of the algorithm proposed in this paper, and the comparative analysis of the number of iterations and the segmentation time is shown in Fig. 6 to Fig. 7. For example, in the image (civil engineering disaster: flood X1), the number of iterations of this paper's algorithm is 5, and the segmentation time is 16.36s, compared with the FCM algorithm, the remaining image has the same law, which indicates that this paper's algorithm performs more outstandingly and excellently above the image segmentation processing.

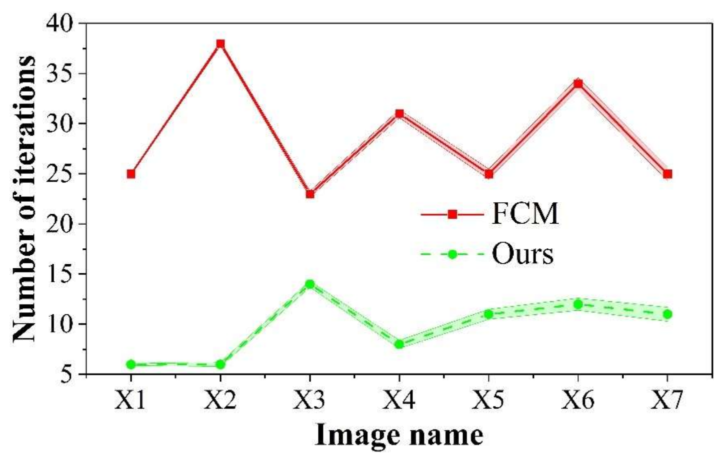


Fig. 6. Comparison of iterations

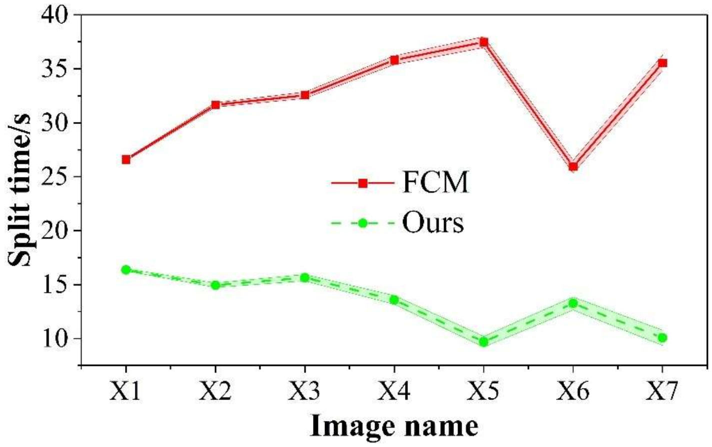
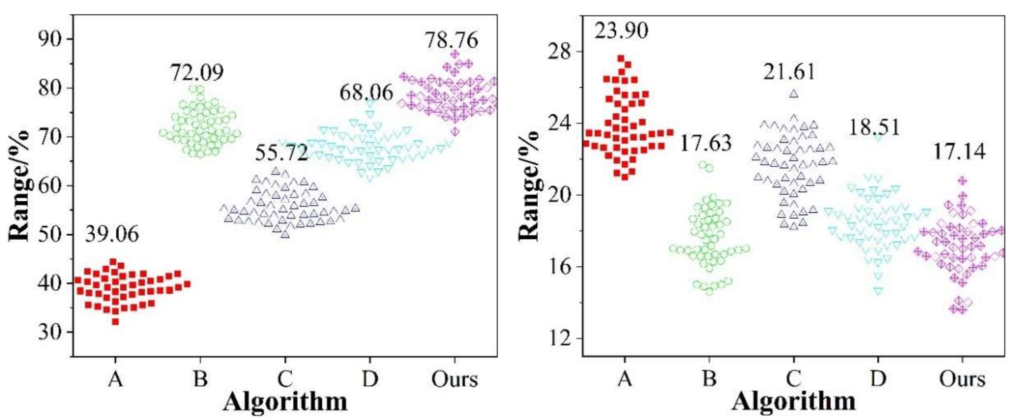


Fig. 7. Split time comparative analysis

3) Numerical analysis of evaluation indexes. Under the MATLAB R2014b toolbox platform, the segmentation algorithm of the image is programmed and realized. In order to better illustrate the effectiveness of the algorithm, the segmentation effect based on Algorithm A, Algorithm B, Algorithm C, Algorithm D and this paper's algorithm is used for comparison, and the numerical analysis of evaluation indexes is shown in Figure 8, where (a) ~ (f) are the matching rate, the misclassification rate, the TPF, the FNF, the FPF, and the TNF, respectively. From the experimental data in the figure, the matching rate of this paper's algorithm is 78.76% and the false segmentation rate is 17.14%, which has obvious advantages over the other four algorithms with the highest matching rate and the lowest false segmentation rate.FNF refers to the ratio of the number of missed segmentation target pixels to the number of all correct pixels in the target area, and FPF refers to the number of background pixels incorrectly segmented to target pixels to the proportion of correct pixels in the background, and the lower the ratio of the two indicates that the segmentation The lower the ratio of the two, the better the segmentation effect. TPF refers to the ratio of the intersection of the pixels in the target region obtained by the segmentation algorithm and the pixels in the target region obtained by manual segmentation of the original image to the number of correct pixels in the target region, and TNF refers to the ratio of the number of segmented background pixels to the number of correct background pixels, and the higher the ratio of the two indicates that the segmentation effect is better. Overall, compared to several other segmentation algorithms, the algorithm in this paper has the best image segmentation effect, which makes its image meet the criteria of intelligent detection model for civil engineering disasters.



(a)Matching rate (b)Error score rate

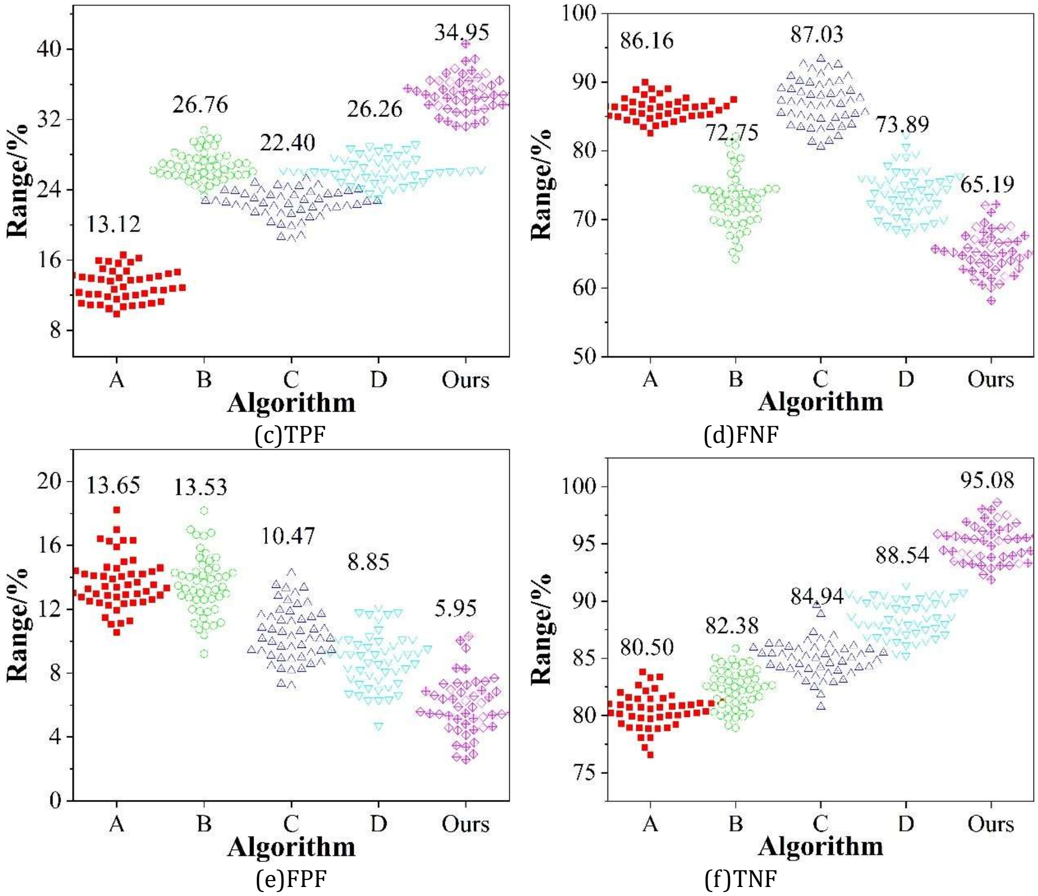


Fig. 8. Numerical analysis of evaluation index

4.2. Disaster Management Research Analysis

4.2.1. Experimental setup. The network models used in the study are implemented on Pytorch, a deep learning development platform, and all experiments use the same data partitioning criterion to divide the processed image dataset into a training set and a validation set, where the ratio of training set to validation set is 8:2. The ratio of the training set to the validation set is 8:2, and all experiments use the same experimental equipment to ensure that the hardware facilities are the same, so as to provide a basis for comparative analysis of the experiments in the future. Each time the Batch Size fed into the model for training is 32. The input image size is 640×640pixel for a total of 200 iterations.

4.2.2. Setting of indicators:

1) Average Precision. Average precision (AP) is one of the most important evaluation indexes in the target detection algorithm, which is calculated by finding the area under the curve of accuracy and recall, and the larger the AP value is, the better the model performance is. Among them, the accuracy rate refers to the ratio of the number of areas predicted to be disasters that are really disasters to the total number of areas predicted to be disasters, and the recall rate is the ratio of the number of correctly predicted disaster areas to the number of real disasters on the ground.The formula for calculating AP is as follows:

$$AP = \frac{1}{r_{number}} \sum_{r \in \{0,x,y,\cdots,1\}} P_{interp}(r) \tag{18}$$

$$P_{interp}(r) = \max_{\tilde{r}, \tilde{r} \geq r} P(\tilde{r}) \tag{19}$$

where $P_{\text{int}erp}$ denotes the maximum accuracy that meets a specific condition, and $P\left(\tilde{r}\right)$ denotes the accuracy when the recall is $\tilde{r}$.

2) FPS. FPS, i.e., detection rate per second. In addition to detection accuracy, another important evaluation index of the target detection algorithm is the speed in detect/second (f/s). In practice, when the FPS is greater than 30f/s, the determination model can provide the ability to process images in real time. Of course, the time required to process an image can also be used to evaluate the detection speed, the shorter the time, the faster the speed.

3) Number of parameters. A model that is to meet practical needs should not only consider how well it performs in terms of accuracy, but also its robustness, scalability, and dependence on resources. Parameter count refers to the number of parameters contained in the model, which directly determines the size of the model file, and also affects the amount of memory occupied by the model when inferring. The unit: mb, the smaller the number of parameters the better it is for deployment on mobile or edge devices.

4.2.3. Model validation analysis:

1) Impact of Attention Mechanisms on Models. Different attention mechanisms pay different attention to the feature map. In the experiments, eight other advanced attention mechanisms are compared: the SE module, the scSE module, three versions of Non-Local module, the GC module, the Split module, and the CABM module, and the effects of using different attention mechanism modules on the model performance are shown in Table 1. From the table, it can be seen that although the CABM attention mechanism used in this paper has a higher number of layers than the other attention mechanisms, it has the smallest number of parameters, and at the same time has a higher detection accuracy.

Table 1. The effect of different attentional mechanism modules on model performance

Description	Layers	Params	AP	FPS
YOLOV4-CABM	274	1.466	0.945	46
YOLOV4-SE	236	2.523	0.859	58
YOLOV4-scSE	238	2.821	0.756	61
YOLOV4-non-local-concatenation	261	3.144	0.829	46
YOLOV4-non-local-dot_product	261	3.144	0.905	50
YOLOV4-non-local-embedded_gaussian	261	3.144	0.788	45
YOLOV4-GC	258	2.531	0.865	50
YOLOV4-Split	238	1.842	0.856	66

2) Effect of the number of detection heads on model performance. The number of detection heads affects the detection accuracy of the model. In this study, one detection head was added to each of the above eight attention mechanism experiments for detecting finer-grained disaster samples: four sizes of disaster feature maps: large, medium, average, and small. The experimental results of the effect of adding a detection head on model performance are shown in Table 2, the model with the smallest number of parameters is still the model that uses the CABM attention mechanism, but the model has lower detection accuracy. The model with the highest detection accuracy has the most parameters and the lowest speed. For the rest of the models, the detection accuracy is below 90%. The model with the smallest number of parameters is still the model that uses the CABM attention mechanism, but this model has a low detection accuracy. The model with the highest detection accuracy has the most parameters and the lowest speed. For the remaining models, the detection accuracy is below 90%. If the number of detection heads continues to increase, the model performance will continue to decrease, and it is solidly not recommended to add extra detection heads.

Table 2. Increasing the influence of the detection head on the model performance

Description	Layers	Params	AP	FPS
YOLOV4-CABM-head	304	1.506	0.906	35
YOLOV4-SE-head	272	2.898	0.876	36
YOLOV4-scSE-head	278	3.156	0.784	34
YOLOV4-non-local-concatenation-head	295	3.489	0.909	32
YOLOV4-non-local-dot_product-head	295	3.489	0.803	32
YOLOV4-non-local-embedded_gaussian-head	295	3.489	0.827	32
YOLOV4-GC-head	288	2.616	0.869	44
YOLOV4-Split-head	278	2.917	0.827	36

3) Comparison of different detection models. To determine the performance of IYOLOV4, comparisons were made with 11 state-of-the-art detection models, including: the FasterRCNN, 3 Efficientdet, 2 Centernet, SSD-efficient, YOLOv4, YOLOv4-tiny, Mobilenetv2-YOLOv4, and Mobilenetv3-YOLOv4. Older versions of these classical models are not listed due to their poor performance. The results of the comparison of different detection models are shown in Table 3, compared to the Efficientdet-D0 model which has a small number of parameters, the number of parameters of IYOLOV4 is only 1.436 mb. The detection accuracy increases by 6.30% compared to the Centernet-hourglass model. The detection speed is up to 36 f/s. Although, many models reduce the parameter redundancy, it leads to lower detection accuracy.

Table 3. Comparison results of different detection models

Model	Params	AP	FPS
IYOLOV4	1.436	0.945	36
Faster RCNN	36.721	0.676	2
Efficientdet-D0	3.883	0.757	9
Efficientdet-D1	6.573	0.787	7
Efficientdet-D2	8.097	0.878	6
SSD-efficient	97.086	0.865	28
Centernet-resnet50	32.737	0.765	28
Centernet-hourglass	191.065	0.889	4
YOLOv4	63.717	0.648	5
YOLOv4-tiny	5.812	0.688	9
Mobilenetv2-YOLOv4	46.244	0.746	10
Mobilenetv3-YOLOv4	48.255	0.706	9

4) Detection application analysis. In order to verify the application performance of the IYOLOV4 model, the trained model is applied to landslide detection in unknown areas. An autonomous state is located in the western region of Yunnan Province, China, and landslides are frequent in its subordinate areas. 180 landslide areas were found to exist in the region by geological survey, mostly concentrated on the sides of large mountains and roads. The geological survey found that as of 2022, there are actually 180 landslide areas in the region. The suspected landslide areas are shown in Figure 9, and the model detects a total of 172 suspected landslide areas from the area, with a model detection accuracy of 95.56% and an error rate of 4.44%. When facing complex terrain, even if there are two or more landslide piles at the same time, the model can also detect them more accurately. For example, the use of UAV oblique photography technology can truly reflect the overall condition of civil engineering disasters, and can intuitively reflect the location, size, scope and surrounding environment, as well as the actual situation, so as to realize comprehensive monitoring. Then with the help of the model to assess and demonstrate the impact range, intensity, etc. of the disaster. In addition, in the civil engineering disaster site, the model in this paper can be used to simulate the terrain, which can effectively prevent the occurrence of accidents, greatly improve the efficiency of the work, and provide guiding value for the innovation

of civil engineering disaster management.

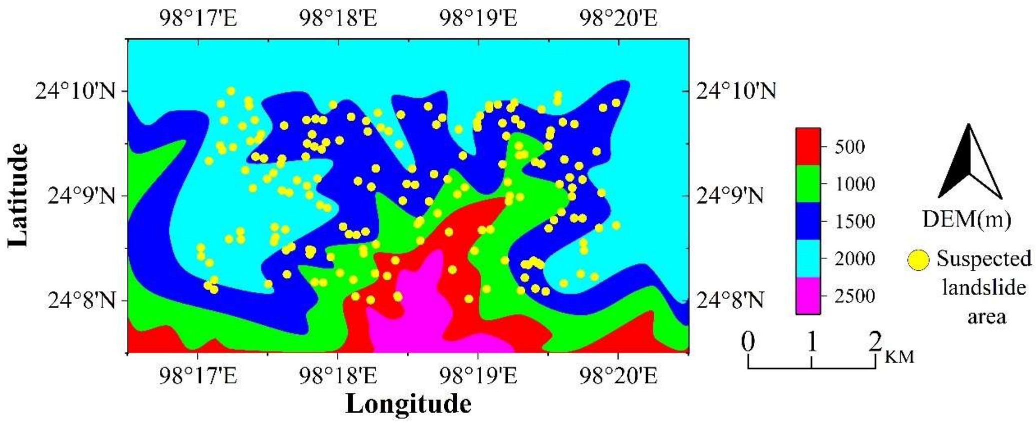


Fig. 9. Suspected landslide area

5. Conclusion

In this paper, civil engineering disaster images are acquired using unmanned tilt photography technology, and it is found that there is still a lot of interfering information in the images, which are image processed with the help of K-Means algorithm, so that the images can be used in the subsequent model. In the context of deep learning algorithms, the improved YOLOV4 target detection algorithm is utilized to construct an intelligent detection model for civil engineering disasters, and the model is validated and evaluated and analyzed.

1) Compared with other algorithms, this paper's algorithm performs well in terms of matching rate, misclassification rate, TPF, FNF, FPF, and TNF, especially in terms of matching rate (78.76%) and misclassification rate (17.14%), which indicates that after the image processing of the K-Means algorithm, it makes its image satisfy the requirements of the civil engineering disaster detection model.

2) Within the definition of civil engineering disasters, the model application analysis is carried out by taking the hilltop landslide as an example, 180 landslide areas existed in a certain area in 2020, and the model detected 172 suspected landslide areas, which results in a model detection accuracy of 95.56%, while the error rate is 4.44%, which is completely within the acceptable range, i.e., the model is able to satisfy the relevant units and enterprises' civil engineering disaster management needs of civil engineering.

Funding

This work was supported by Research on BIM Modeling Methods for Highway Engineering throughout the Entire Lifecycle.

References

[1] Juang, C. H., Dijkstra, T., Wasowski, J., & Meng, X. (2019). Loess geohazards research in China: Advances and challenges for mega engineering projects. Engineering geology, 251, 1-10.

[2] Xie, L., & Qu, Z. (2018). On civil engineering disasters and their mitigation. Earthquake Engineering and Engineering Vibration, 17, 1-10.

[3] Jaisingpure, J., Raut, P. R., Kale, R., Ganjare, P., Chaudhari, H., Misal, R., ... & Kale, P. (2022). Construction techniques to prevent landslides. Int. J. of Aquatic Science, 13(1), 143-150.

[4] Lu, X., Liao, W., Fang, D., Lin, K., Tian, Y., Zhang, C., ... & Zhao, P. (2020). Quantification of disaster resilience

in civil engineering: A review. Journal of safety science and resilience, 1(1), 19

[5] Firoozi, A. A., & Firoozi, A. A. (2024). Disaster Management in Civil Engineering. In Neuromorphic Computing: Transforming Disaster Management and Resilience in Civil Engineering (pp. 23-32). Cham: Springer Nature Switzerland.

[6] Gong, W., Juang, C. H., & Wasowski, J. (2021). Geohazards and human settlements: Lessons learned from multiple relocation events in Badong, China–Engineering geologist's perspective. Engineering Geology, 285, 106051.

[7] Wang, Z. F., Cheng, W. C., & Wang, Y. Q. (2018). Investigation into geohazards during urbanization process of Xi'an, China. Natural Hazards, 92, 1937-1953.

[8] Yin, Z. Y., Jin, Y. F., & Zhang, X. (2021). Large deformation analysis in geohazards and geotechnics. Journal of Zhejiang University-SCIENCE A, 22(11), 851-855.

[9] Sergey, K., & Gennadiy, N. (2022). Methodology for the risk monitoring of geological hazards for buildings and structures. International Journal of Disaster Risk Management, 4(1), 41-49.

[10] Zhu, H. H., Shi, B., & Zhang, C. C. (2017). FBG-based monitoring of geohazards: current status and trends. Sensors, 17(3), 452.

[11] Yang, F., Jiang, Z., Ren, J., & Lv, J. (2022). Monitoring, prediction, and evaluation of mountain geological hazards based on InSAR technology. Scientific Programming, 2022(1), 2227049.

[12] He, L., Tan, J., Hu, Q., He, S., Cai, Q., Fu, Y., & Tang, S. (2018). Non-contact measurement of the surface displacement of a slope based on a smart binocular vision system. Sensors, 18(9), 2890.

[13] Li, Y., Li, J., Pei, J., Chen, W., Luo, G., & Xia, G. (2024). Application of non-contact video quantitative measurement method in reservoir bank landslide monitoring. Frontiers in Earth Science, 12, 1378046.

[14] Chen, K., & Jiang, Q. (2023). A non-contact measurement method for rock mass discontinuity orientations by smartphone. Journal of Rock Mechanics and Geotechnical Engineering, 15(11), 2892-2900.

[15] Ding, Q., Wang, F., Chen, J., Wang, M., & Zhang, X. (2022). Research on generalized RQD of rock mass based on 3D slope model established by digital close-range photogrammetry. Remote Sensing, 14(9), 2275.

[16] Jie, D. A. N. G., Ji, D. O. N. G., Songbiao, H. E., & Xuanmei, F. A. N. (2022). Application of airborne LiDAR and ground 3D laser scanning in geological hazard risk investigation of Dujiazhai collapse in Shuicheng, Guizhou. The Chinese Journal of Geological Hazard and Control, 33(4), 106-113.

[17] Di Stefano, F., Cabrelles, M., García-Asenjo, L., Lerma, J. L., Malinverni, E. S., Baselga, S., ... & Pierdicca, R. (2020). Evaluation of long-range Mobile Mapping System (MMS) and close-range photogrammetry for deformation monitoring. A case study of Cortes de Pallás in Valencia (Spain). Applied Sciences, 10(19), 6831.

[18] Fanos, A. M., & Pradhan, B. (2019). A novel rockfall hazard assessment using laser scanning data and 3D modelling in GIS. Catena, 172, 435-450.

[19] Guo, Y., Li, X., Ju, S., Lyu, Q., & Liu, T. (2022). Utilization of 3D Laser Scanning for Stability Evaluation and Deformation Monitoring of Landslides. Journal of Environmental and Public Health, 2022(1), 8225322.

[20] Manzoor, S., Liaghat, S., Gustafson, A., Johansson, D., & Schunnesson, H. (2020). Establishing relationships between structural data from close-range terrestrial digital photogrammetry and measurement while drilling data. Engineering Geology, 267, 105480.

[21] Wang, J., Zhu, S., Luo, X., Chen, G., Xu, Z., Liu, X., & Li, Y. (2020). Refined micro-scale geological disaster susceptibility evaluation based on UAV tilt photography data and weighted certainty factor method in Mountainous Area. Ecotoxicology and Environmental Safety, 189, 110005.

[22] Hujun, J. I. A., Lijuan, W. A. N. G., & Dongli, F. A. N. (2021). The application of UAV LiDAR and tilt photography in the early identification of geo-hazards. The Chinese Journal of Geological Hazard and Control, 32(2), 60-65.

[23] Wu, B., Li, B., Li, S., Li, H., & Zhou, R. (2021, April). Research on the application of uav tilt photography technology in engineering project. In Journal of Physics: Conference Series (Vol. 1885, No. 2, p. 022037). IOP Publishing.

[24] Zhao, Z. H., Sun, H., Zhang, N. X., Xing, T. H., Cui, G. H., Lai, J. X., ... & He, H. J. (2024). Application of unmanned aerial vehicle tilt photography technology in geological hazard investigation in China. Natural Hazards, 120(13), 11547-11578.

[25] Zhou, T., Lv, L., Liu, J., & Wan, J. (2021). Application of UAV oblique photography in real scene 3d modeling. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43, 413-418.

[26] Tang, M., Mei, X., Li, Y., Chen, C., Liu, X., & Lu, H. (2024). Application of oblique photogrammetry technique in geological hazard identification and decision management. Earthquake Research Advances, 4(3), 100269.

[27] Hu, X., & Li, D. (2020). Research on a single-tree point cloud segmentation method based on UAV tilt photography and deep learning algorithm. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 4111-4120.

[28] Hao, J., Zhang, X., Wang, C., Wang, H., & Wang, H. (2023). Application of UAV digital photogrammetry in geological investigation and stability evaluation of high-steep mine rock slope. Drones, 7(3), 198.

[29] Ruban, I., Khudov, H., Makoveichuk, O., Khudov, V., Kalimulin, T., Glukhov, S., ... & Shamrai, N. (2022). METHODS OF UAVS IMAGES SEGMENTATION BASED ON K-MEANS AND A GENETIC ALGORITHM. Eastern-European Journal of Enterprise Technologies, 118(9).

[30] Shen Xia. (2024). Retraction Note: E-commerce User Recommendation Algorithm Based on Social Relationship Characteristics and Improved K-Means Algorithm. International Journal of Computational Intelligence Systems,17(1),

[31] Bengang Liu,Wenjiang Wu,Yunpeng Zhang,Zeguang Dong & Baode Li. (2024). Tool wear image segmentation algorithm based on K-means clustering. Journal of Physics: Conference Series,2787(1),

[32] Et-taleby Abdelilah,Chaibi Yassine,Boussetta Mohammed,Allouhi Amine & Benslimane Mohamed. (2022). A novel fault detection technique for PV systems based on the K-means algorithm, coded wireless Orthogonal Frequency Division Multiplexing and thermal image processing techniques. Solar Energy,237,365-376.

[33] Yunxia Chen & Yan Wu. (2025). Detection of Welding Defects Tracked by YOLOv4 Algorithm. Applied Sciences,15(4),2026-2026.