International Journal for Housing Science and Its Applications
Publish August 5, 2025. Volume 46, Issue 3 Pages 3899-3908

https://doi.org/10.70517/ijhsa463331

Design of Crack Detection and Damage Analysis System for Civil Structures Based on Improved YOLO Algorithm

Linxuan Zhang[1] and Yu Deng[1,*]
[1] School of Civil Engineering and Architecture, Guangxi University of Science and Technology, Liuzhou, Guangxi, 545006, China
Corresponding authors: (e-mail: dengyu@gxust.edu.cn).

Abstract Aiming at the problems of low accuracy of crack monitoring and insufficient efficiency of damage assessment in health monitoring of civil structures, a crack detection and damage analysis system based on YOLO algorithm for civil structures is designed in the study. The crack feature extraction capability is enhanced by optimizing the YOLOX network structure, introducing lightweight convolution and coordinate attention mechanism. At the same time, a new MODSLayer module is designed, which enables the model to extract features in high dimensions of MODL-Head. The system integrates image acquisition and processing, multi-scale crack detection, geometric parameter quantification, and damage assessment modules, and realizes the automation of the whole process from detection to analysis. The MOD-YOLO algorithm F1 score is 0.521, which is 3.6% to 18.6% higher than the comparison algorithm, and the mAP reaches 58.491%, which is also much higher than the comparison algorithm. The results of this paper's model for crack length as well as width are basically consistent with the real results, with an average relative error of 0.96% and 1.21%, respectively. The system constructed in the study detected that the length of crack No. 7 has achieved the maximum value of 1938.7 mm, and the angle with the base surface reaches 75.7°, which may continue to grow longitudinally. In this paper, the system found that the nonlinear coefficient of the civil engineering construction and the length of the crack has a pattern of “surge-slowly increasing-decreasing”, which indicates that the model has a high sensitivity to recognize the damage degree of the cracks in the civil engineering structure.

Index Terms YOLO algorithm optimization, lightweight convolution, coordinate attention mechanism, crack detection, system design

I. Introduction

The structure of civil engineering is an important support in people's working, living and learning environments, and only by ensuring the structural safety of civil engineering can people's daily life be carried out smoothly [1], [2]. Once the structure of civil engineering is difficult to ensure, it will inevitably bring great threat to people's life and property safety. As the civil engineering structure is in the external environment for a long time, the occurrence of natural disasters such as earthquakes, typhoons, fires and so on will bring great influence to the civil engineering structure, the external environment of rain, snow, dust and so on will also make the civil engineering structure in the long-term under the action of the load, under the cumulative effect of the civil engineering structure is difficult to avoid the different degrees of damage [3]-[5]. Therefore, it is of great significance to accurately determine the degree and type of damage of civil engineering and evaluate the health status of civil engineering structures [6]. The necessary inspection and analysis of civil engineering structures through the application of inspection technology and damage identification methods is an important means to improve the use value of the project and ensure the safety of people's lives and properties, which makes it a key research topic at home and abroad [7], [8].

For a long time, the traditional human-based inspection method has been widely used in practical engineering as the main structural inspection means. The method is simple, flexible and suitable for the inspection of a small number of small and medium-sized structures [9]. With the rapid development of the level of construction technology, complex and large-scale structures are constantly being built and operated, and the shortcomings of manual inspection methods in inspecting large-scale large-scale engineering structures, such as inefficiency, high cost, and high risk, have been exposed [10]-[12]. Along with the development of science and technology, the civil engineering structural damage diagnosis is developing towards intelligence, and neural network technology has become the main tool for civil engineering structural damage detection and analysis [13], [14]. Neural network has the function of self-learning, and at the same time, this technology has strong fault tolerance and diffusion, with the help of neural network technology can be a good solution to the damage of the pattern and the influence of high noise [15]-[17]. The artificial neural network responds accordingly to different states, while the trained neural network can classify

patterns, reflecting the pattern of structural damage to a certain extent, making the identification of civil engineering structural damage diagnosis more convenient and accurate [18]-[20].

The study redesigns the classical depth-separable convolution based on the YOLOX algorithm to obtain more detailed information about the features through point-by-point convolution operation. Meanwhile, on top of the original coordinate attention mechanism, it is optimized by combining two pooling methods, average pooling and maximum pooling. And MODSNet feature extraction layer was built using MODSLayer to realize crack feature extraction without dimensionality reduction. Finally, a new lightweight decoupling head was designed to obtain the MOD-YOLO algorithm, and the optimized algorithm was used to form a crack detection and damage analysis system for civil structures. Different algorithms are selected to compare the effect and performance of crack detection, and the system is tested in the field to evaluate its practical effect.

II. Methodology

II. A. YOLO algorithm optimization

The YOLO family of algorithms suffers from limitations on the task of crack detection in civil engineering structures as well as insufficient feeling fields. For this reason, the YOLOX algorithm is chosen as the baseline model for its optimization in this paper.

II. A. 1) Lightweight convolutional introduction

In order to solve the problems of missing inter-channel correlation information and missing information caused by downscaling the extracted features that exists in the classical depth-separable convolution after deep convolution, this paper redesigns the depth-separable convolution by applying the idea of Resnet [21]. In this paper, the input feature layer X is added with the feature layer X_1 that lacks inter-channel correlation information after deep convolution operation, and the feature layer X_2 with inter-channel correlation information is obtained by supplementing the missing information with inter-channel correlation information possessed by the original input features. Then for X_2, a point-by-point convolution operation is performed to integrate the inter-channel information of the feature layer X_2 and perform feature extraction to obtain a feature layer with more information X_{out}:

$$X_1 = f_1(X) \tag{1}$$

$$X_2 = X_1 + X \tag{2}$$

$$X_{out} = f_2(X_2) \tag{3}$$

where f_1 denotes the deep convolution operation on the input feature layer, and f_2 denotes the point-by-point convolution operation on the reinforced feature X_2 after summing up the input feature layer X with the feature layer X_1 that is missing inter-channel correlation information after the deep convolution operation, and finally obtaining the feature layer with more information X_{out}.

II. A. 2) Integration of attention mechanisms

In this paper, we use a combination of two pooling methods, average pooling [22] and maximum pooling [23], to redesign a plug-and-play coordinate attention mechanism with salient and average information (DAF-CA). As shown in Figure 1, the DAF-CA designed in this paper pools the input feature layers along the height and width directions in two pooling modes, average pooling and maximum pooling, respectively, to generate two spatial dimensions of $(h,1)$ and two spatial dimensions of $(1,w)$, a total of four feature layers, h and w are the height and width of the input feature layers, which contain two feature layers with significant information that are not found in the original coordinate attention mechanism. In this paper, two identical spatial dimensions are stacked on the width and height dimensions respectively, and the stacked feature layers are changed into two new feature layers with spatial dimensions $(h,2)$ and $(2,w)$, and then a new feature layer x' with spatial dimensions $(h+w,2)$ or $(2,h+w)$ is generated through dimensional transformation and dimension stacking. Finally, the weights of each feature in the feature layer are calculated by the sigmoid function [24], and multiplied by the original feature layer respectively to obtain the feature layer X_{out} enhanced by features.

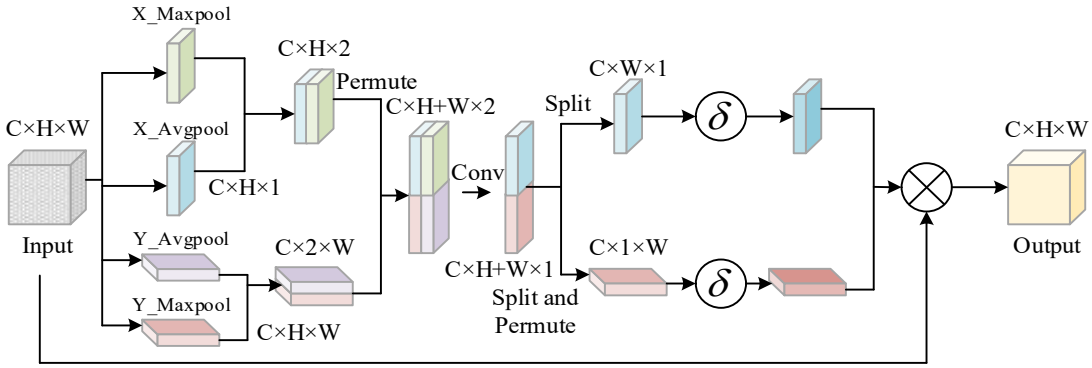

Figure 1: Schematic diagram of the DF-CA module

$$X_h^{'} = C(P_{ha}(X), P_{hm}(X)) \tag{4}$$

$$X_w^{'} = C(P_{wa}(X), P_{wm}(X)) \tag{5}$$

$$X^{'} = C(X_h^{'}, X_w^{'}) \tag{6}$$

$$X_{out} = X * \delta(f(f_1(X^{'})_{split})) \tag{7}$$

where P_{ha} and P_{hm} denote the average and maximum pooling along the h-dimensional direction, P_{wa} and P_{wm} denote the average and maximum pooling operations along the w -dimensional direction, respectively, and the c function denotes the stacking of the feature layers, f_1 represents the convolution operation by a decay rate of r, convolution kernel size of 1×2 or 2×1, Padding of (0, 0), the split function represents splitting the feature layer in half in the high or wide dimension, the f function represents the convolution feature extraction operation using a convolution kernel size of 1×1, the output channel dimensions of the original dimensions of the convolutional feature extraction module, the δ function denotes the use of a sigmoid activation function to assign weights to the feature layer with an interval size of (0, 1), and * denotes a dot product operation. The X_{out} is the enhanced feature layer after multiplying the original feature layer X with the weights.

II. A. 3) Feature extraction layer optimization

In this paper, we use MODSConv to design a MODSLayer module for feature extraction without dimensionality reduction, and rebuild the Backbone and Neck layers of the network with MODSLayer and combine them into MODSNet.The input features are extracted by the feature extraction layer composed of n MODSConvs. The feature layer X_1 is extracted by n MODSConv, and the feature layer X_1 is then summed with the original input features to prevent the gradient from disappearing and enhance the information of the feature layer by adopting the same idea of Resnet. Finally, the feature layer X_{out} without dimensionality reduction is obtained by channel integration and feature extraction using a normal convolutional layer with a convolutional kernel of 1 × 1:

$$X_1 = f_n(f_{n-1}(\cdots f_2(f_1(X)))) \tag{8}$$

$$X_2 = (X_1 + X) \tag{9}$$

$$X_{out} = f(X_2) \tag{10}$$

where $f_1 \cdots f_n$ is the feature layer undergoing feature extraction operation through MODSConv, f denotes the convolution kernel of 1×l for the feature layer X_1 obtained by summing up the input feature layer X with the feature layer X_2 obtained after n MODSConv feature extraction layers. Ordinary convolution operation, and finally get the new feature layer X_{out} after MODSLayer feature extraction module.

II. A. 4) Prediction Layer Optimization

MODL-Head on the input feature layer and YOLOX series decoupling head first through a convolutional layer is unified channel dimensionality reduction and then processed in a different way, this paper designed MODL-Head in

accordance with the classification and regression of the two tasks are processed separately in accordance with the idea of the input features are divided into two routes, and directly in accordance with the original scale respectively through the MODSConv feature extraction, and then completed separately classification and regression tasks respectively:

$$X_{Cls} = f(f_1(X)) \tag{11}$$

$$X_{\mathrm{Re}g} = f(f_1(X)) \tag{12}$$

$$X_{IoU} = f(f_1(X)) \tag{13}$$

where $X_{Cls}, X_{Reg}, X_{IoU}$ are the prediction results of classification, prediction frame regression, and confidence, respectively, and f_1 denotes the use of the MODSConv operation, and f denotes the use of the ordinary convolution operation with a convolution kernel size of 1×1. The final prediction results are judged by $X_{Cls}, X_{Reg}, X_{IoU}$ fused together.

II. B. MOD-YOLO based crack detection model for civil structures

In order to improve the performance of crack detection based on YOLO algorithm, this paper proposes a method named DAF-CA, which adopts average pooling and maximum pooling to obtain the average information and significant information and fuses them to assign the attention weights more reasonably.In order to reduce the number of parameters and computation, the method of extracting features by dimensionality reduction is adopted, which is based on MODSConv and DAF-CA. MODSLayer is designed based on MODSConv and DAF-CA, which well protects the rich information between channels with a strategy of extracting features without dimensionality reduction.

In addition, it is found in this paper that the SPP as well as the SPPF modules in the original YOLO architecture are not able to characterize the information of the global sensing field range well. Therefore, in order to further improve the detection performance of the network, this paper proposes a method to expand the sensory field of SPPF to the global by using global average pooling and global maximum pooling, so as to enhance its ability to represent the global information, and names it GRF-SPPF.Finally, considering that the Head part of the original YOLO architecture adopts the strategy of unifying the dimensions first and then extracting the features, which may lead to a large amount of information loss between channels, therefore, in order to solve the above possible problems this paper proposes a MODL-Head based on MODSConv, which extracts features directly without dimensionality reduction.

In this paper, the original algorithm is improved by using the above four improvement strategies, and the MOD-YOLO detection algorithm, which significantly improves the overall detection performance of cracks in civil engineering structures, is finally constructed by relying on the above improvement scheme.

II. C. Design of crack detection and damage analysis system for civil engineering structures

Combined with the optimized YOLO algorithm above, the design of the crack defect detection and damage analysis system for civil engineering structures mainly contains three processes, and Fig. 2 shows the overall flow of the system design.

(1) Image data acquisition and processing of civil engineering structures:

Before designing the system, it is necessary to carry out relevant preparatory work, in order to obtain the civil structure crack images required by the system, first of all, it is necessary to collect the civil structure images and screen out the civil structure crack defect images that meet the requirements of the system, and then the civil structure crack defect images will be pre-processed using the data enhancement to expand the sample library of civil structure crack defects. Finally, the processed civil structure crack defect pictures are denoised using the designed denoising model to construct a civil structure crack data set that meets the system requirements.

(2) Detection algorithm design and model construction:

Load the constructed civil structure crack defect data set into the detection algorithm model, train and adjust the detection algorithm by choosing different optimizers and learning rates, save the detection weights when the detection effect is the best, and use it in the civil structure crack defect detection system.

(3) Civil structure crack defect detection system construction:

Firstly, QT-Designer in PyCharm is used to create the basic interface of the crack defect detection system for civil engineering structures, i.e. the login interface and the crack detection interface. Then, use the various control subfunctions in the toolbox of PyQt5 to design the functions of displaying system name, account registration, account login and password retrieval in the login interface. Similarly, design the functions of importing pictures,

displaying pictures to be detected, displaying pictures of detection results and outputting detection results in the crack detection interface. Finally, the PyQt5 function library is used to convert the interface UI format file of the designed civil structure crack defect detection system into Python language. After obtaining the required python language file, the python code is used to import the best detection model weights obtained in step 2 into the detection system, and the system can complete the large-scale screening of civil structure crack defects in the database by calling the obtained detection model weights.

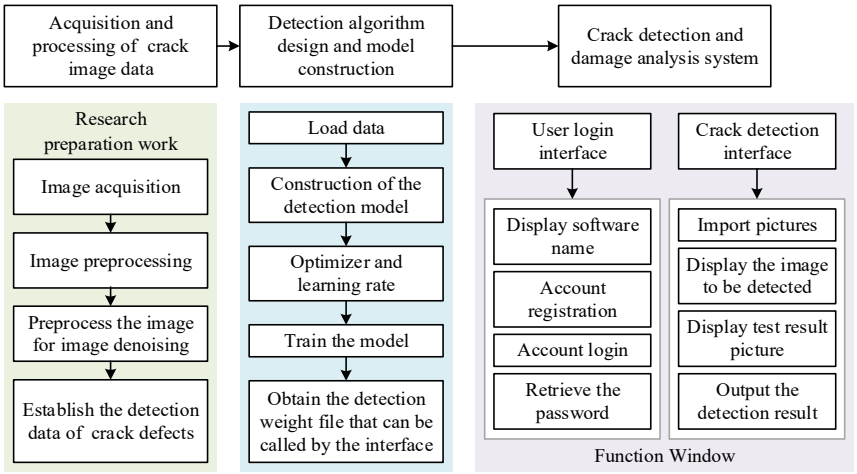

Figure 2: Implementation process of the wellbore crack defect detection system

II. D. Experimental setup

II. D. 1) Data acquisition

The crack datasets used in this study include the publicly available crack dataset RDD2022 as well as a self-made crack dataset. Among them, RDD2022 is a road crack dataset based on the 2022 Crowd-Receptive Road Damage Detection Challenge (CRDDC), which is based on the RDD2020 dataset with the addition of new road damage images collected from Norway, the United States, and China, with a total of 47,420 images There are a total of more than 40,000 images of pavement cracks, of which 28% are from Japan, 20% are from India , 22% from Norway, 13% from the US, 7% from the Czech Republic, and 10% from China. and labeled with labelImg labeling software divided into eight categories, which are longitudinal crack D00, transverse crack D10, tortoise crack D20, pothole crack D40, longitudinal splice D01, transverse splice D11, intersection fuzzy D43, and roadway white line fuzzy D44, and the label saving format is XML.In this study, we used the cracks from three countries, namely China, Japan, and India. images for training and analysis, and because pavement distresses mainly consist of transverse cracks, longitudinal cracks, cracked cracks and potholes, the eight types of pavement cracks are reordered using sorting software and the remaining four types of labels that are not necessary to be used in training are frozen in order to form a new dataset.

The self-collected civil structure cracks will be re-labeled and made into a new dataset, which includes more than 5000 samples in total, and the labels are divided into five categories: transverse cracks HCrack, longitudinal cracks LCrack, cracks in tortoise cracks HoneyCrack, 45°Crack, and 30°Crack, and the labels are saved in the format of TXT.

The collected crack dataset will be randomly assigned by an algorithm to form a training set and a test set. Specifically, the RDD2022 public dataset is allocated according to the ratio of 7:3, while the homemade crack dataset is divided into training set, test set and validation set according to the ratio of 7:2:1. The resolution of the images in the two datasets is different, 300 dpi for RDD2022 and 96 dpi for the homebrew dataset.For the convenience of training, all the images will be uniformly adjusted to 640×640 pixels by the size modification software.

II. D. 2) Data set labeling and transformation

The homemade crack dataset needs to be re-labeled using LabelImg software, and the labeling steps are to establish the folder where the dataset is stored first, and two subfolders are established under the folder datasets, which are images and labels, and the images and labels folders include three folders, which are train, test, and val, where images is used to store the crack images to be used, and labels stores the corresponding labeling files and index files classeses, val, where images is used to store the crack images to be used, and labels stores the corresponding labeling files and index files classes. Then the LabelImg annotation software was opened in Pytorch-

GPU environment, and the crack photos were labeled by drawing rectangular boxes, and the save format was set to TXT.Since the save format of the public crack dataset RDD2022 was XML, and the research adopted the YOLO series algorithm, which required the use of TXT format datasets, the training format from XML format to TXT.

II. D. 3) Equipment and software

The specifications of the computer hardware equipment used to realize the functions of building the initial improved YOLO network algorithm, processing the dataset algorithm, training the model, evaluating the model and predicting the results are as follows:

GPU: NVIDIA GeForce RTX 2080 Ti
CPU: Inter(R) Xeon(R) Platinum 8260
CUDA: 12.0
Cudnn: 8.2.2

The experiments were conducted using the Python language with the deep learning framework Pytorch, configuring the environment required for training in Anaconda and writing Python programs using the VSCode editor to train the network in the Anaconda software. The software versions and languages used in this time are as follows:

Anaconda: 2019.10
VSCode: 2017
Python: 3.8.5
Torch: 1.8.0

III. Results and analysis

III. A. Comparison of crack detection effect and performance of different algorithms

In order to validate the detection performance of the YOLO-CD algorithm in the YOLO series, it was trained and tested in comparison with Faster RCNN, CenterNet, RetinaNet, YOLOv5n, and YOLOX, on the datasets described above. Important academic evaluation metrics were used, including F1, mean accuracy (mAP), frames per second (FPS) of the screen, number of model parameters (Params), number of floating-point operations per billion operations per second (GFLOPs), and weight file size. F1 integrates the accuracy and recall to reflect the overall performance of the network in a more comprehensive way. The mAP is a composite evaluation metric that combines the precision and recall of different categories. The calculation involves the average precision (AP) of each category, and the final result is the average of the APs of all categories. FPS denotes the number of frames detected per second, Params denotes the size of the model, and GFLOPs is used to measure the complexity of the algorithm. F1 and mAP calculated based on accuracy and recall are used as the detection effectiveness metrics. FPS, Params, GFLOPs, and weight file size are used to measure the performance of the algorithm in this paper.

Table 1 shows the results of crack detection and performance comparison of different algorithms. From the table, it can be seen that compared with these algorithms Faster RCNN, CenterNet, Reti-naNet, YOLOv5n, and YOLOX, MOD-YOLO improves the F1 by 18.6, 9.5, 13.7, 4.8, and 3.6 percentage points, and the mAP improves the mAP by 23.99, respectively, while guaranteeing the running speed, 17.14, 16.04, 9.04, and 6.52 percentage points. The experimental results show that compared with other target detection algorithms, MOD-YOLO algorithm achieves extremely high detection accuracy with a small model size, and has the advantages of high detection efficiency, low cost of use, and simpler deployment.

Table 1: Crack detection effect and performance comparison results

Algorithm	F1	mAP(%)	FPS	Params(105)	GFLOPs	Weight file(MB)
Faster RCNN	0.335	34.500	23.762	1345.423	387.550	101.952
CenterNet	0.426	41.347	41.829	328.660	75.422	133.920
RetinaNet	0.384	42.456	29.429	360.715	158.363	148.313
YOLOv5n	0.473	49.451	166.768	18.216	13.797	3.810
YOLOX	0.485	51.969	112.065	64.788	3.182	10.604
MOD-YOLO	0.521	58.491	132.888	28.652	30.292	4.356

Ablation experiments are conducted to verify the performance enhancement of the YOLO algorithm using lightweight convolution, DAF-CA attention mechanism, MODSNet, and MODL-Head module, and the results of the model structure ablation experiments are shown in Table 2. From the data analysis, it can be concluded that the lightweight convolution reduces the number of model parameters by 8.60%, the computation amount by 4.98%, and the FPS is improved by 19.933. Combined with the DAF-CA attention mechanism, the average detection accuracy is improved by 7.194%. The MODSNet and MODL-Head modules are also added to obtain the optimized model of

this paper, which improves the F1 score of the model by 0.321 and the average accuracy by 30.902%. The proposed algorithm enhances the multi-dimensional feature fusion extraction ability of the model and realizes the improvement of model detection performance and detection accuracy.

Table 2: Model structure ablation experiment

Light mass convolution	×	√	√	√	√
DAF-CA	×	×	√	√	√
MODSNet	×	×	×	√	√
MODL-Head	×	×	×	×	√
F1	0.224	0.337	0.443	0.498	0.545
mAP(%)	26.613	40.183	47.377	52.703	57.515
FPS	69.899	89.832	114.549	123.718	130.816
Params(105)	42.976	39.282	35.179	31.502	28.691
GFLOPs	42.195	40.295	35.929	31.746	30.993

In order to further validate the crack inspection and damage analysis model proposed in this paper, the study conducted a crack measurement validation experiment on a homemade dataset. The actual length and width of the cracks were measured, and the crack length and maximum width were calculated using the model in this paper for the captured crack images, and the results of the crack length and width calculations are shown in Fig. 3. The results show that the model proposed in this paper measures the crack length similarly to the actual length, with an overall average relative error of 0.96%, and measures the crack width similarly to the actual maximum width, with an average relative error of 1.21%. One of the reasons for the errors is that the crack skeleton map extracted by semantic segmentation is not exactly in line with the original crack morphology, which leads to the gap between the measurement results and the actual length and width, and on the other hand, due to the deviation of the scale factor in the calibrated pixel scaling relationship, but the overall average relative error is controlled within 5%.

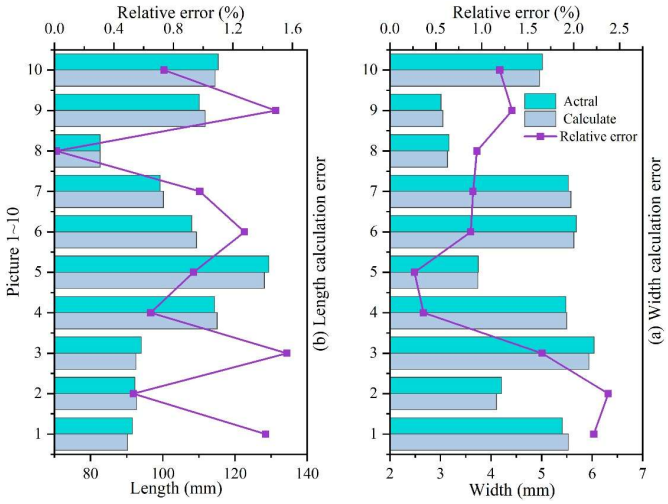

Figure 3: Crack length and width calculation result

III. B. Analysis of actual engineering feasibility of crack detection

In order to verify the feasibility of the system proposed in this study for practical engineering applications, structural cracks in a civil engineering project at a site are analyzed in situ. The images captured by a wall-climbing robot equipped with the system in this paper may truncate the complete crack into several parts. Therefore, four neighboring images of a central image containing a crack are determined, and if the neighboring image contains a crack, it is spliced with the central image. All the images involved have been distortion corrected. The process is repeated and then the corresponding binarization results of the spliced images are used to separate the cracks using a single connected domain algorithm to obtain independent crack images. Finally, the geometric measurement of the complete crack is performed, for the crack length can be realized by counting the number of skeleton points. For the crack direction, it is calculated by the second order moments of all crack pixels in the single connected domain, which is calculated as follows:

$$\theta = \arctan\left[\frac{\left|\mu_{xx}-\mu_{yy}+\sqrt{(\mu_{xx}-\mu_{yy})^2-4\mu_{xy}^2}\right|}{2\mu_{xy}}\right] \tag{14}$$

$$\begin{cases} \mu_{xx} = \sum_{i}^{N}\frac{(x_i-\bar{x})^2}{N}+\frac{1}{12} \\ \mu_{yy} = \sum_{i}^{N}\frac{(y_i-\bar{y})^2}{N}+\frac{1}{12} \\ \mu_{xy} = \sum_{i}^{N}\frac{x_i y_i}{N} \end{cases} \tag{15}$$

where: x_i, y_i are the image coordinates of the crack skeleton pixels, and $\bar{x}$, $\bar{y}$ are the average values of all coordinates. N is the total number of crack skeleton pixels, and μ_{xx}, μ_{yy} are the second-order moments of the crack skeleton pixels in the x direction and y direction, which are used to measure the degree of dispersion of the crack skeleton pixels in a single direction. μ_{xy} is the mixed second-order moments of the slit skeleton pixels in the x and y directions, and is used to measure the correlation of the slit skeleton pixels in both directions.

For elongated cracks with no branches and no significant angular changes, based on the above method using a larger image block to completely wrap the crack, the approximate direction of the crack in the plane can be solved by the above equation and combined with the naked eye to make a preliminary judgment of the growth direction of the crack.

Table 3 lists the geometric information of the 10 cracks, and the crack dimensions are quantified with sub-millimeter accuracy. Due to the fact that more tests have been carried out with this structure as the test object, the average widths of all cracks exceeded 0.2 mm, and the average widths of most of the cracks exceeded 0.3 mm. The maximum and average widths of crack No. 3 obtained the maximum values among all cracks, which were 0.75 and 0.53 mm, respectively, and the maximum length of crack No. 7 obtained the maximum value of 1,938.7 mm. According to the calculation results of the angle between the crack direction and the datum line, the length of crack No. 7 obtained the maximum value of 1,938.7 mm. angle with the baseline, the angle between crack No. 7 and the basal plane is 75.7°, which may continue to grow along the longitudinal direction of the structure in the subsequent development. Therefore, the growth of crack No. 7 needs to be paid attention to in time, while the other cracks are transverse cracks, which have less impact on the structure. Based on the structural crack conditions shown in the table, comprehensive information support can be provided to civil engineering operation and management units to improve the science of management and maintenance decision making.

Table 3: 10 geometric information of cracks

Crack number	Maximum width		Mean width		Length		Angle(°)
	Pixels	mm	Pixels	mm	Pixels	mm	
1	1.73	0.38	1.68	0.37	2479.18	545.4	93.1
2	2.32	0.51	1.27	0.28	3309.45	728.08	101.3
3	3.41	0.75	2.41	0.53	6343.73	1395.62	90.51
4	2.45	0.54	2.18	0.48	2871.82	631.8	95.2
5	1.14	0.25	1.95	0.43	3194.68	702.83	81.7
6	2.14	0.47	1.95	0.43	2729.45	600.48	98.7
7	1.91	0.42	1.77	0.39	8812.27	1938.7	75.7
8	1.68	0.37	1.18	0.26	5243.73	1153.62	85.8
9	1.91	0.42	1.00	0.22	5644.73	1241.84	93.9
10	2.50	0.55	2.14	0.47	5836.82	1284.1	94.2
1 Pixels=0.22mm							

III. C. Analysis of crack damage identification in civil engineering structures

The damage development of two specimens of the same size, material and maintenance conditions was realized using three-point bending loading respectively. Eleven damage stages were set for each specimen, i.e., damage degree of 0 (healthy state), 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0, respectively.

The relationship between the degree of cracking and the nonlinear coefficient was obtained as shown in Fig. 4. From the figure, it can be seen that the change rule between the nonlinear coefficient and the crack length of specimen 1 and specimen 2 is basically the same. In the healthy state, the nonlinear coefficient is close to 0. In the early stage of microcrack development (0-30% damage degree) the slope is higher, which is more sensitive to damage development. While in the middle stage of microcrack growth (20% to 80% damage degree), the damage coefficient grows more slowly. After the specimen is completely damaged, the nonlinear coefficient falls back sharply. This variation rule of the nonlinear coefficient is due to the difference in the distribution of the actual crack morphology in the crack length direction. The results of the damage recognition test show that the system in this paper can have good recognition sensitivity for civil structures with damage degree ranging from 0% to 80%. Therefore, the damage recognition of early microcracks in civil structures can be well realized by using the technique of this paper, and a good sensitivity can be guaranteed.

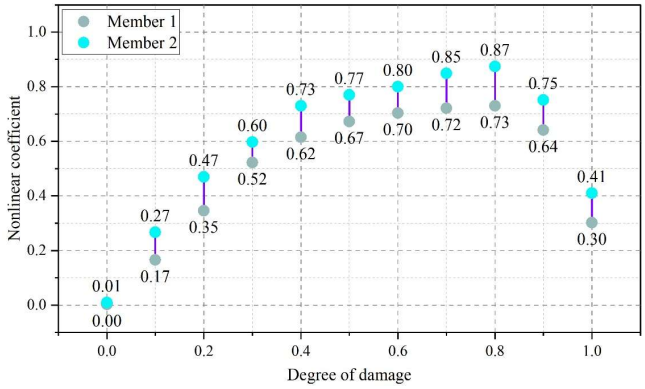

Figure 4: The relationship between crack degree and nonlinear coefficient

In order to verify the damage analysis function of the system, ABAQUS software is used in this paper for numerical simulation of crack damage in civil structures. In this study, Embedded is used to simulate the constraint relationship between reinforcement and concrete, i.e., bond-slip is not taken into account.The two ends of the finite element model of the RC simply-supported beam are constrained by hinges, and the displacement load is applied at the three-quarter point position of the model.

The load-displacement curves of simulation and test are compared in Fig. 5, and the overall patterns of the two curves are in good agreement. The yield and ultimate loads of the members are 75 kN and 94 kN, respectively, and the corresponding displacements are 10.51 mm and 47.2 mm. At the same displacement, the simulated load capacities are 74.6 kN and 92.4 kN, respectively, and the errors with the tests are in the range of 0.68%~2.48%. The results verify the reliability of the system in this paper for damage analysis of civil structures.

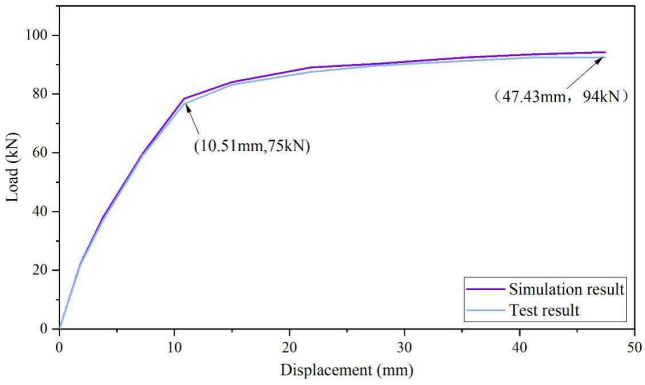

Figure 5: Load - displacement curve of simulation and experiment

IV. Conclusion

In this paper, the YOLOX algorithm is used as a baseline model for optimization. The MOD-YOLO model is obtained by introducing four optimization efforts: lightweight convolution, fusion attention mechanism, feature extraction layer optimization, and prediction layer optimization. The collected public dataset of civil structure cracks as well as the homemade dataset are inputted into the model, and the appropriate optimizer and learned are selected for training

and parameter tuning, which are loaded into the crack detection and damage analysis system. Comparative experiments, field tests and damage identification analysis are combined to verify the efficient performance of the model and the system in crack detection and damage analysis of civil structures. The optimized model in this paper has a better crack detection effect, with F1 scores improved by 3.6-18.6 percentage points and mAP improved by 6.52-23.99 percentage points compared with the comparison algorithm. In addition, the average relative errors of the model on crack length and width measurements are 0.96% and 1.21%, respectively, which are much lower than the set error ranges. In the field test, the system in this paper detected that the length of crack No. 7 was the largest and the angle with the base surface was 75.7°, and there was the possibility of continued extension. The system shows good damage recognition sensitivity for civil structures with damage levels ranging from 0% to 80%. The experiment proves that the system in this paper can realize the intelligent detection of cracks in civil engineering structures, which has significant engineering application value.

References

[1] Zar, A., Hussain, Z., Akbar, M., Rabczuk, T., Lin, Z., Li, S., & Ahmed, B. (2024). Towards vibration-based damage detection of civil engineering structures: overview, challenges, and future prospects. International Journal of Mechanics and Materials in Design, 20(3), 591-662.

[2] Frigui, F., Faye, J. P., Martin, C., Dalverny, O., Pérès, F., & Judenherc, S. (2018). Global methodology for damage detection and localization in civil engineering structures. Engineering structures, 171, 686-695.

[3] Makhloof, D. A., & Ibrahim, A. R. (2021). Damage Assessment of Reinforced Concrete Structures through Damage Indices: A State-of-the-Art Review. Cmes-Computer Modeling In Engineering & Sciences, 128(3).

[4] Zhang, C., Gholipour, G., & Mousavi, A. A. (2020). Blast loads induced responses of RC structural members: State-of-the-art review. Composites Part B: Engineering, 195, 108066.

[5] Bado, M. F., & Casas, J. R. (2021). A review of recent distributed optical fiber sensors applications for civil engineering structural health monitoring. Sensors, 21(5), 1818.

[6] Bao, Y., Chen, Z., Wei, S., Xu, Y., Tang, Z., & Li, H. (2019). The state of the art of data science and engineering in structural health monitoring. Engineering, 5(2), 234-242.

[7] Gui, G., Pan, H., Lin, Z., Li, Y., & Yuan, Z. (2017). Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection. KSCE Journal of Civil Engineering, 21, 523-534.

[8] Hung, C. C., El-Tawil, S., & Chao, S. H. (2021). A review of developments and challenges for UHPC in structural engineering: Behavior, analysis, and design. Journal of Structural Engineering, 147(9), 03121001.

[9] Gres, S., Andersen, P., Johansen, R. J., Ulriksen, M. D., & Damkilde, L. (2018). A comparison of damage detection methods applied to civil engineering structures. In Experimental Vibration Analysis for Civil Structures: Testing, Sensing, Monitoring, and Control 7 (pp. 306-316). Springer International Publishing.

[10] Gres, S., Ulriksen, M. D., Döhler, M., Johansen, R. J., Andersen, P., Damkilde, L., & Nielsen, S. A. (2017). Statistical methods for damage detection applied to civil structures. Procedia engineering, 199, 1919-1924.

[11] Hou, R., & Xia, Y. (2021). Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. Journal of Sound and Vibration, 491, 115741.

[12] Zhang, C., Mousavi, A. A., Masri, S. F., Gholipour, G., Yan, K., & Li, X. (2022). Vibration feature extraction using signal processing techniques for structural health monitoring: A review. Mechanical Systems and Signal Processing, 177, 109175.

[13] Gulgec, N. S., Takáč, M., & Pakzad, S. N. (2019). Convolutional neural network approach for robust structural damage detection and localization. Journal of computing in civil engineering, 33(3), 04019005.

[14] Feng, C., Zhang, H., Wang, S., Li, Y., Wang, H., & Yan, F. (2019). Structural damage detection using deep convolutional neural network and transfer learning. KSCE Journal of Civil Engineering, 23(10), 4493-4502.

[15] Zhang, Y., Huang, J., & Cai, F. (2020). On bridge surface crack detection based on an improved YOLO v3 algorithm. IFAC-PapersOnLine, 53(2), 8205-8210.

[16] Neves, A. C., Gonzalez, I., Leander, J., & Karoumi, R. (2017). Structural health monitoring of bridges: a model-free ANN-based approach to damage detection. Journal of Civil Structural Health Monitoring, 7, 689-702.

[17] Raushan, R., Singhal, V., & Jha, R. K. (2025). Damage detection in concrete structures with multi-feature backgrounds using the YOLO network family. Automation in Construction, 170, 105887.

[18] Ma, J., Yan, W., Liu, G., Xing, S., Niu, S., & Wei, T. (2022). Complex texture contour feature extraction of cracks in timber structures of ancient architecture based on YOLO algorithm. Advances in Civil Engineering, 2022(1), 7879302.

[19] Liang, Y., & Xu, Z. (2025). Intelligent inspection of appearance quality for precast concrete components based on improved YOLO model and multi-source data. Engineering, Construction and Architectural Management, 32(3), 1691-1714.

[20] Du, F. J., & Jiao, S. J. (2022). Improvement of lightweight convolutional neural network model based on YOLO algorithm and its research in pavement defect detection. Sensors, 22(9), 3537.

[21] Wei Liu,Xiaohui Ye & Wendi Yan. (2025). Power quality disturbance classification based on dual-parallel 1D2D fusion of improved ResNet and attention mechanism. Measurement,252,117358-117358.

[22] Zhaohui Chen,Elyas Asadi Shamsabadi,Sheng Jiang,Luming Shen & Daniel Dias da Costa. (2024). An average pooling designed Transformer for robust crack segmentation. Automation in Construction,162,105367-.

[23] Michael Kohler & Sophie Langer. (2025). Statistical theory for image classification using deep convolutional neural network with cross-entropy loss under the hierarchical max-pooling model. Journal of Statistical Planning and Inference,234,106188-106188.

[24] Andreas Skiadopoulos & Maria Knikou. (2025). Optimal sigmoid function models for analysis of transspinal evoked potential recruitment curves recorded from different muscles.. PloS one,20(1),e0317218.