Computers, Materials & Continua CMC, vol.63, no.1, pp.457-468, 2020

Method to Appraise Dangerous Class of Building Masonry Component Based on DC-YOLO Model

Hongrui Zhang[1], Wenxue Wei[1, *], Xinguang Xiao[1], Song Yang[1] and Wanlu Shao[1]

Abstract: This DC-YOLO Model was designed in order to improve the efficiency for appraising dangerous class of buildings and avoid manual intervention, thereby making the appraisal results more objective. It is an automated method designed based on deep learning and target detection algorithms to appraise the dangerous class of building masonry component. Specifically, it (1) adopted K-means clustering to obtain the quantity and size of the prior boxes; (2) expanded the grid size to improve identification to small targets; (3) introduced in deformable convolution to adapt to the irregular shape of the masonry component cracks. The experimental results show that, comparing with the conventional method, the DC-YOLO model has better recognition rates for various targets to different extents, and achieves good effects in precision, recall rate and F_1 value, which indicates the good performance in classifying dangerous classes of building masonry component.

Keywords: Deep learning, masonry component, appraisal of dangerous class, deformable convolution.

1 Introduction

The state and government have always been paying great attention on the appraisal and renovation of dilapidated buildings because the housing safety directly concerns the life and property safety of the people. In order to accurately and objectively appraise the dangerous classes of the houses and ensure the structural safety, the government has issued the *Standard on Dangerous Building Appraisal*, according to which, the dangerous class appraisal for masonry component is the first step for appraising the dangerous class of the house foundation and superstructure [Hu and Lin (2017); Chen, Wang, Yang et al. (2019)]. Up to now, there has neither public database been set up for the dangerous class appraisal of buildings like we do for face and handwriting identification, nor software been developed to support dangerous class appraisal of structure and houses. In this case, the appraisal is actually relied on eye observation and inspection to get conclusions according to the operator's experience and relative regulations. This could be quite subjective, and make relative experiments difficult. Therefore, proposing a method of appraising the dangerous class of the building structure, which is both objective and easy-

[1] College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, 266590, China.

[*] Corresponding Author: Wenxue Wei. Email: wwxjyh@163.com.

Received: 21 April 2019; Accepted: 18 July 2019.

to-operate, is of great significance to the master the house status, prevent damages caused by natural disasters, and renovate dilapidated buildings.

With the ever greater attention paid on livelihood issues of the people by the state, many scholars are committed to the appraisal of building dangerous classes. Yang et al. [Yang and Zhang (2005)] propose the method for dangerous building appraisal based on analytic hierarchy process and fuzzy theory, which classifies the building class by adopting weighted average mathematical model and asymmetric closeness method [Yang and Zhang (2005)]; Lu et al. [Lu (2016); Su and Wang (2004)] proposed the integrated multi-level fuzzy appraisal method based on the combined weights obtained from analytic hierarchy process and entropy method to comprehensively appraise the dangerous concrete structure buildings in fuzzy way, thereby making quantitative assessment for the reinforced concrete structure buildings, took use of the set pair analysis theory, to carry out identical discrepancy contrary (IDC) analysis on and determine the building conditions and classes [Su and Wang (2004)]. However, though all the above methods can be applied to appraise the dangerous classes of the building by mathematical means, they are not quite preferential for being operated in actual execution process, because they require the appraisers to have certain professional knowledge and experience [Wang (2018); Xu, Liang, Shi et al. (2016)]. Another type of method is the traditional ones that appraise according to experience, which is featured in great subjective factors. It needs to take use of relative instrument and equipment in order to acquire relative data by force bearing of certain structures. Too much requirement on manpower and material investment show that such method is not adaptable to the current situation of mass relocation and house renovation in China [Gao (2018); Wang (2018)].

Targeting at the problems of low operability, complicated process, and serious subjective factors in currently existing methods, a DC-YOLO model, which combines the target detection algorithm and the house detection together was designed in this paper. It adopts K-means clustering to acquire the best prior boxes, introduces in deformable convolution to adapt to irregular shapes of structure cracks, expands grid size to better identify small targets, and extracts key elements stipulated in Standard on Dangerous Building Appraisal to carry out masonry component risk assessment. Moreover, it further divides the risks into four classes, which makes the appraisal on building risks more contrapuntally.

2 Theoretical basis

2.1 Target detection model-YOLO

YOLO (You only look once: unified, real-time object detection) [Redmon and Divvalay (2016)] is a kind of regression-based target recognition algorithm. Compared with previous generations, YOLOV3 recognizes faster for it only needs to do one time of forward pre-computing to achieve target detection, which greatly improves the recognition rate [Cai, Wu, Liu et al. (2018); Zhou, Jin and Dong (2017)]. YOLOV3 applies multi-scale prediction and residual network module, which help to improve its detection rate for small targets to large extent. The model adopts the 3×3 convolutional layer to extract features, and adopts the 1×1 convolutional layer to reduce the amount of computation [Shi, Chen and Yang (2019); Yang, Yang, Su et al. (2018)].

YOLO divides the input images into S×S cells. And the cells take use of the anchor boxes to predict the bounding box. YOLOV3 adopts 3 scales to detect targets for both COCO datasets and VOC datasets according to their characteristics of the dataset. Each scale has 3 anchor boxes. The information of the bounding box is represented by the quintuple T (x, y, w, h, c), of which, the x and y represent the horizontal and vertical coordinates of the bounding box, while the w and h represent the width and height of the bounding box, and c represents the confidence, which determine whether the current bounding box contains the estimated probability of the predicted target and its prediction accuracy. Relative calculation formula is as shown in formula (1).

$$c = P_o \times P_{IOU} \tag{1}$$

In which, PO indicates the probability that the bounding box contains the predicted target, and PIOU indicates the overlapping area of the bounding box and the real target area. When the bounding box contains the predicted target, P_O=1, otherwise P_O=0. If the cell contains multiple bounding boxes, then take the largest overlap area.

2.2 Deformable convolution network

A major difficult point existing in the traditional convolutional neural networks is how to make the network adapt to the geometric transformation of the target. For this problem, arXiv [Dai, Qi and Xiong (2017)] et al. proposed a deformable convolution, which introduced in the ability of learning spatial geometrical transformation, namely the deformable convolutional network [Gao, Li, Zhang et al. (2018); Xu, Zhang, Wu et al. (2018)].

The main idea of deformable convolution is to improve the receptive field in the traditional convolutional neural network, transforming from the original rectangle or square to the irregular receptive field, to better adapt to the actual situation. The introduced in deformable convolution kernel can extract characteristics according to the target to be detected. However, the contour of the extracting object is not standard rectangle, so that, the general contour of the target is taken as the bounding box of the convolution kernel, thus to reserve the characteristics of the target that is to be detected to the largest extent, not resulting in loss or damage of the object characteristic information due to bounding box shape.

3 DC-YOLO model

Cracks have solid inherent nature and irregular shape, stretching towards different direction due to different severity of wall damage, therefore, it can make the geometric transformation in the convolution inefficient if directly using the square convolution to extract the characteristics, thereby lowering the identification rate. As a matter of fact, the data set contains certain number of small targets with different crack sizes. This proposes high requirement on network accuracy.

DC-YOLO is a model proposed on the basis of YOLOV3. It finds the best prior box, expands the grid size, improves the detection for small targets, and introduces in deformable convolution, thereby making the shape of the convolution kernel adapt to the target contour shape which solves the problem that traditional models are difficult to adapt to the characteristics of the target.

3.1 DC-YOLO structure

The DC-YOLO model structure is as shown in Fig. 1. It takes the entire image as the input, and outputs the detected cracks and categories of the masonry components at the same time, including: ClassA, ClassB, ClassC, and ClassD. The input image is divided into S×S cells, each cell is responsible for the objects falling into the cell. The cells actually take use of the anchor boxes to predict the bounding box: perform inspection on three scales of 13×13, 26×26, and 52×52, and adopt doubled upsampling to make the characteristic graphs be transmitted on adjacent scales. The model contains 27 convolution layers, a fully connected layer and a Softmax classification layer. Meanwhile, a pooling layer is set to save the original picture information. The output scale of the fully connected layer is calculated as S×S×(2×5+C), in which, C represents the number of categories. And combining with the data set of this experiment, the value of C shall be 4, therefore it outputs a tensor of 1134 scales, and finally the classification function is implemented by softmax.

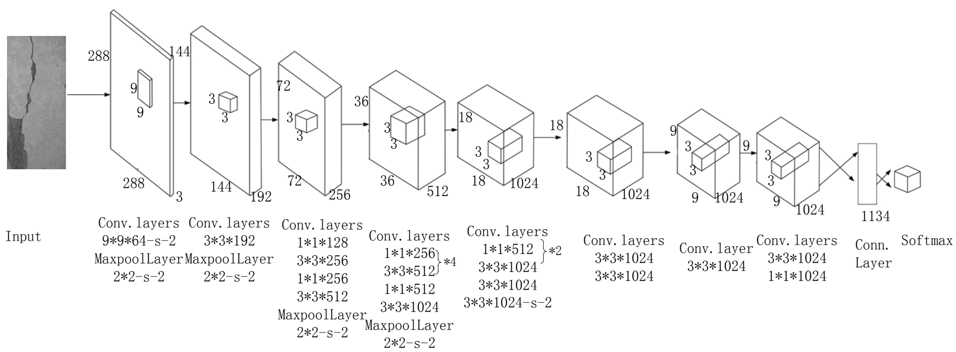

Figure 1: Model structure diagram

And ass for the 3×3 convolution kernel, please see Fig. 2.

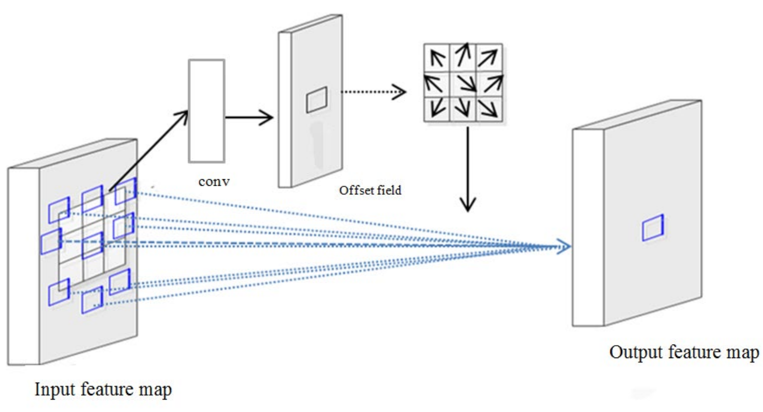

Figure 2: Deformable convolution

3×3 convolution kernel is expressed as:

R={(-1, -1), (-1, 0), ……. , (0, 1), (1, 1)}

As for Point p0, whose convolution kernel weight is w, output characteristic graph is x, and output characteristic graph is y, the mathematical representation of traditional convolution is as shown in Formula (2):

$$\mathrm{y}(\mathrm{p}_0)=\sum_{p_n\in R} w(p_n)\bullet x(p_0+p_n) \tag{2}$$

where pn represents any position in Area R.

As for the 3×3 convolution kernel adopted by DC-YOLO, a position offset △pn that can learn characteristics is introduced in based on traditional convolution, please see Formula (3):

$$\mathrm{y}(\mathrm{p}_0)=\sum_{p_n\in R} w(p_n)\bullet x(p_0+p_{n+\Delta}p_n) \tag{3}$$

where △pn can be expressed as: {pn |n=1,...,N},N∈R

Since the sampling is performed in an irregular area, the positions of the off-layer characteristic points are not continuous. Therefore, the bilinear interpolation is adopted to transform the outputs of any positions re converted into continuous. The Formula (3) is implemented by using bilinear interpolation as in Formula (4).

$$\mathrm{x}(\mathrm{p})=\sum_{q} G(q,p)\bullet x(q) \tag{4}$$

where p represents the discrete sampling points, q is the continuous points set after bilinear interpolation.

The cracks of masonry components have the following characteristics: It contains certain number of small targets, while being dominated by targets of the middle size. It needs to detect irregular cracks of different sizes stretching towards different directions, which has strict requirement on the network accuracy. It was found through the experiment that, the size of the grid has a great influence on the detection rate of the small targets. The DC-YOLO model adds an 18×18 grid after the multi-layer convolution, and the expanded grid size improves the identification to small targets and enhances the accuracy of the whole network, detecting more small targets under the same conditions. After that, two 9×9 networks were set to further extract the characteristics and reduce the characteristic space and computing complexity with a 1×1 convolution layer.

3.2 Best a priori box

According to the characteristics of the wall cracks, and combining with the relationship between the number of prior boxes and the average Intersection-over-Union (IOU), as shown in Fig. 3, the K-means clustering [Guo and Lin (2017); Zheng, Liu, Pan et al. (2019)] method is adopted to obtain the best prior box. The user-defined distance metric formula of K-means clustering is as shown in Formula (5), and the IOU value between box and cluster center box is selected as the distance index.

$$d(box,centroid)=1-IOU(box,centroid) \tag{5}$$

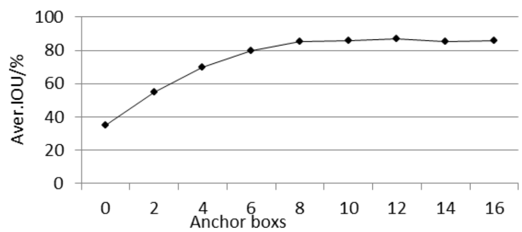

Figure 3: Relationship between the number of prior boxes and the average intersection-over-union (IOU)

It can be seen from the figure that, when the number of boxes is 9, the average IOU does not rise anymore. So 9 prior boxes were taken, which were respectively (10×13), (13×15), (16×25), (30×61), (50×64), (65×110), (115×87), (155×189), (367×345), and considering that the characteristics of middle sized in data sets are in great amount, great attention was paid on size of the prior box of middle scale at the time of setting. 3 scales were taken for inspection. And as for each cell of each scale, 3 prior boxes were taken use of to predict 3 bounding boxes.

3.3 Definition to loss function

DC-YOLO adopts the mean square error as the loss function which consists of three parts: coordinate error, IOU error and classification error, please see Formula (6).

$$\text{loss} = coordError + iouError + classError \tag{6}$$

The $coordError$ represents coordinate error, and the specific calculation method is as shown in the Formula (7).

$$coordError = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} \left[(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right] \tag{7}$$

where S represents the grid size of the characteristic graph map, B represents the number of prediction boxes per cell, λ_{coord} represents the weight coefficient of the coordinate error, and xi, yi, wi, and hi respectively represent the horizontal and vertical coordinates of the central point of the predicted object, and the width and height of the predicted box; $\hat{x}_i$, $\hat{y}_i$, $\hat{w}_i$, and $\hat{h}_i$ respectively represent the horizontal and vertical coordinates of the central point of the real sliding window of the cell i and I_{ij}^{obj} represents that the target falls into No.j predicted box of Cell i.

$iouError$ represents IOU error, and the specific calculation method is as shown in Formula (8)

$$iouError = \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{obj} (c_i - \hat{c}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{noobj} (c_i - \hat{c}_i)^2 \tag{8}$$

where, λ_{noobj} represents the weight of the IOU error, ci and $\hat{c}_i$ respectively represent the confidence value of the cell i of the predicted box and the actual confidence value, and I_{ij}^{noobj} indicates that the target does not fall into No. j predicted box of Cell i.

$classError$ represents the classification error, and its specific calculation method is as shown in Formula (9).

$$classError = \sum_{i=0}^{S^2} I_i^{obj} \sum_{c \in classes} [p_i(c) - \hat{p}_i(c)]^2 \tag{9}$$

where, $p_i(c)$ and $\hat{p}_i(c)$ represent the conditional probability that Cell i contains the Class C target in prediction and actual sliding window.

4 Experiments and results

Since there is no public database for building testing, the data applied in experiments of this paper was actually collected by the research team, please see Fig. 4.

Figure 4: Data set display

4.1 Data annotation

Several characteristics that commonly exist in the dangerous masonry structure of buildings were taken as the basis for experiment annotation, which are respectively: load-bearing wall or column has vertical crack with over 1.0 mm width and length of over 1/2 of floor height; the load-bearing wall or column has multiple vertical cracks with length over 1/3 of the floor height; the wall or column has horizontal crack due to eccentric compression; the connecting part of two adjacent components breaks and forms a penetrated crack.

Harmful cracks refer to those that can affect the performance, function and durability of building structure if they get developed continuously. There's no clear requirement on the harmful crack width of masonry structure, so that the description on the wall crack width in the Standard on Building Appraisal was taken for reference and "1 mm" was taken as the boundary of the classification. In this case, the cracks with over 1 mm width shall be deemed as harmful cracks. In addition, according to the field investigation, a "5 mm" classification boundary was also added in the experiment.

Four classes are defined according to the national requirements on building components and the requirements of this experiment on classification:

1. Class A: no dangerous point

2. Class B: have dangerous points, usually singles horizontal cracks with less than 1mm width or vertical crack with length shorter than 1/3 of the layer height. These dangerous points will not affect the safety of the component structure.

3. Class C: partially dangerous: have any one of the four above mentioned characteristics, but the crack width is less than 5 mm;

4. Class D: overall dangerous: have any two or more of the four above mentioned characteristics or have crack with over 5 mm width. This indicates that the components are unable to satisfy the requirement on use safety and have risk of collapse.

The experimental data set includes 2040 training sets and 400 test sets. The open source tool labelimg was applied to mark the data with characteristics of four classes as mentioned above.

4.2 Appraisal standard

The precision, recall rate and F1 value were selected as the standard for evaluating the network performance, specifically, please see the followings:

Precision indicates the proportion of a certain type of pictures occupying the total pictures that are divided as this type. This value reflects the ability to judge the classification correctness of a certain type.

The recall rate indicates the ratio of a certain correctly detected type to a certain type existing in the test set. This value reflects the ability of the network for correctly detecting a certain type.

The F1 value is an indicator used in statistics to measure the accuracy of the classification model. It considers both the precision and the recall rate at the same time, and can be deemed as a weighted average of the two. The calculation is as shown in Formula (10).

$$F_1 = \frac{2 \cdot \mathrm{Pr}\,ecision \cdot \mathrm{Re}\,call}{\mathrm{Pr}\,ecision + \mathrm{Re}\,call} \tag{10}$$

4.3 Experimental results

The experimental configuration is as shown in Tab. 1.

Table 1: Configuration for experiments

Name	Configuration
OS	CentOS 7.5
CPU/GHz	Intel(R) Xeon(R) E5-2690 v4 @ 2.60 GHz
Memory/GB	500 G
GPU	NVIDIA-SMI 390.77
DeepLearning Framework	Darknet

The parameters set for the network training are as follows: set batch as 64, set max_batches as 15000, adopt the optimized method of variable learning rate-Adam, set the initial learning rate of the first 13000 times as 0.001, and set the learning rate of the later 2000 times as 0.0001, in order to acquire the best model more accurately, it was set to perform one iteration for every 1000 times to save the weight files during the training process.

Tab. 2 to Tab. 5 show the classification results of YOLO, YOLOV2 [Wei, Quan and Hou (2017)], YOLOV3 [Zhang, Yang and Li (2019)], SSD512 [Chen, Wang, Yang et al. (2019); Xie (2019)] and DC-YOLO model of this paper, which have all been trained and tested based on the data set of this experiment. It can be seen from the tables that, the DC-YOLO model shows over 0.9 classification accuracy to all categories, shows over 0.87 recall rate, and around 0.9 F1 value. Compared with other methods, it obtains improved results to different extents in all the three aspects.

Table 2: Experimental results of different methods for Class A

Methods	Precision	Recall	F_1
YOLO	0.7746	0.8077	0.7908
YOLOV2	0.8379	0.8476	0.8427
YOLOV3	0.8932	0.9003	0.8967
SSD512	0.8756	0.8914	0.8834
DC-YOLO	**0.9333**	**0.9998**	**0.9655**

Table 3: Experimental results of different methods for Class B

Methods	Precision	Recall	F_1
YOLO	0.7698	0.7046	0.7357
YOLOV2	0.7987	0.7790	0.7887
YOLOV3	0.8535	0.8058	0.8290
SSD512	0.8657	0.8341	0.8496
DC-YOLO	**0.9062**	**0.8788**	**0.8923**

Table 4: Experimental results of different methods for Class C

Methods	Precision	Recall	F_1
YOLO	0.8502	0.7905	0.8193
YOLOV2	0.8997	0.8018	0.8479
YOLOV3	0.9482	0.8299	0.8851
SSD512	0.9579	0.8309	0.8899
DC-YOLO	**0.9714**	**0.8717**	**0.9189**

Table 5: Experimental results of different methods for Class D

Methods	Precision	Recall	F_1
YOLO	0.9062	0.9137	0.9099
YOLOV2	0.9267	0.9384	0.9325
YOLOV3	0.9589	0.9692	0.9640
SSD512	0.9409	0.9735	0.9569
DC-YOLO	**0.9713**	**0.9853**	**0.9782**

With the increase of the number of iterations, the DC-YOLO model gradually becomes close to 1, the total loss value drops to about 0.026941, and the average loss value drops to about 0.033607 avg. The results are shown in Fig. 5.

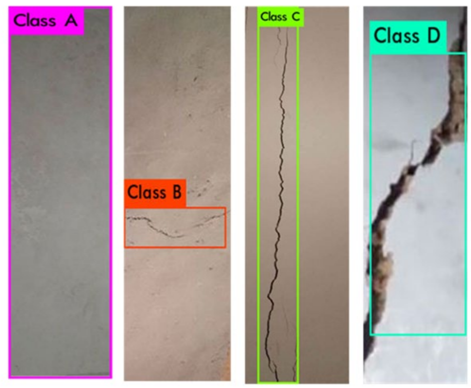

Figure 5: DC-YOLO experiment results

At the same time, DC-YOLO classification accuracy is higher than others under a same threshold. Different problems like missed detection and repeated detection happened when respectively detecting by YOLO, YOLOV2, YOLOV3, and SSD512. DC-YOLO improves the recognition rate of the target, and reduces the occurrence of missed detection and repeated detection. Part of the excerpt is as shown in Figs. 6 and 7.

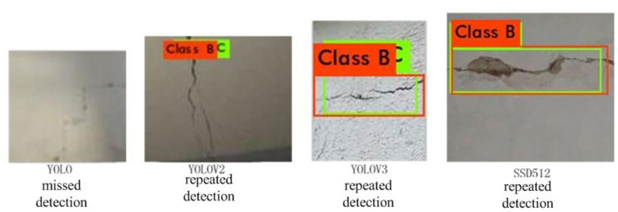

Figure 6: Problems found in traditional method

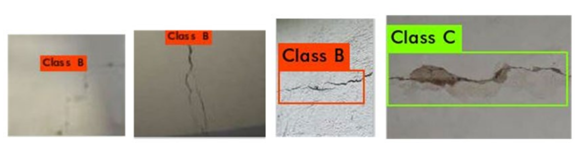

Figure 7: DC-YOLO test results

5 Conclusions

The DC-YOLO model was proposed in this paper to realize the automated identification of the dangerous classes of building masonry components. Three strategies were adopted in the model of this paper, including: expanding the grid size to detect small targets more efficiently, using K-means clustering to obtain the best prior box, and introducing in the deformable convolution to learn the deformation information of the component cracks. Compared with traditional single models, this model is featured in high accurate rate of classification. This consolidates the foundation to certain extent and is of great practical significance for the realizing the appraisal of overall risk of dilapidated buildings in the future.

Acknowledgement: The work is supported by National key research and development plan of China (2016YFC0801408), the Graduate Science and Technology Innovation Project of Shandong University of Science and Technology (SDKDYC180344).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Cai, C.; Wu, K.; Liu, Q.; Cheng, H.; Ma, Q. (2018): Panoramic multi-target real-time detection based on improved YOLO algorithm. *Computer Engineering and Design*, vol. 39, no. 10, pp. 3259-3264, 3271.

Chen, P. (2002): Discussion on several issues relating to standard on dangerous building appraisal (JGJ125-99). *Building Structure*, vol. 32, no. 11, pp. 71-73.

Chen, H.; Wang, Q.; Yang, G.; Hang, J.; Yin, C. et al. (2019): SSD target detection algorithm based on multi-scale convolution characteristics fusion. *Computer Science and Exploration*, pp. 1-15.

Dai, J.; Qi, H.; Xiong, Y. (2017): Deformable convolutional networks. https://blog.csdn.net/Quincuntial/article/details/78767668.

Gao, X. (2018): *Detection and Classification of Concrete Bridge Cracks Based on Deep Learning (Ph.D. Thesis)*. Shanxi Normal University.

Gao, X.; Li, H.; Zhang, Y.; Yan, M. (2018): A vehicle detection method for remote sensing image dense area based on deformable convolution neural network. *Journal of Electronics & Information Technology*, vol. 40, no. 12, pp. 2812-2819.

Guo, Z.; Lin, T. (2017): Research on k-means algorithm for mass-data-oriented rapid clustering. *Computer Applications and Software*, vol. 34, no. 5, pp. 43-7,53.

Guo, Q. (2015): Precautions during identification of building risks. *Building Structure*, vol. 45, no. 9, pp. 11-13.

Joesph, Redmon.; Santosh, Divvalay. (2016): YOLO: real-Time object detection, available. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE Computer Society.*

Lu, B. (2016): Research on integrated multi-level fuzzy appraisal for dangerous RC structural buildings based on combined weights of AHP and entropy methods editorial department of engineering mechanics. *Chinese Society of Theoretical and Applied Mechanics.*

Shi, H.; Chen, X.; Yang, Y. (2019): Improved helmet wearing detection method of YOLO v3. *Computer Engineering and Application*.

http://kns.cnki.net/KCMS/detail/11.2127.TP.20190311.1538.012.html.

Su, W.; Wang, J. (2004): Comprehensive attribute hierarchy optimization method for industrial building plans. *Industrial Construction*, vol. 34, no. 8, pp. 14-16, 18.

Wang, T. (2018): *Reliability Identification and Analysis of Existing Masonry Structures (Master's Thesis)*. Anhui University of Science and Technology.

Wang, L. (2018): *Detection of Concrete Pavement Cracks Based on Deep Learning (Ph. D. Thesis)*. Hebei University of Engineering.

Wei, Y.; Quan, J.; Hou, Y. (2017): Research on aerial image positioning of UAV based on YOLO v2. *Laser & Optoelectronics Progress*, vol. 54, no. 11, pp. 101-110.

Xie, Z. (2019): Research on gesture recognition based on improved convolutional neural network. *Computer Applications and Software*, vol. 36, no. 3, pp. 192-195+281.

Xu, J.; Liang, J.; Shi, L.; Cao, X. (2016): Discussion on some issues related to the development of masonry structure in China. *Building Structure*, vol. 46, no. 15, pp. 91-97.

Xu, G.; Zhang, W.; Wu, X.; Su, J.; Guo, Z. (2018): Lane line semantic segmentation algorithms based on convolutional neural network. *Journal of Electronic Measurement and Instrument*, vol. 32, no. 7, pp. 89-94.

Yang, J.; Zhang, Y. (2005): Application of fuzzy analytic hierarchy process and fuzzy theory in the identification of dangerous buildings. *Journal of Hebei University of Technology*, vol. 34, no. 6, pp. 91-95.

Yang, G.; Yang, J.; Su, Z.; Chen, Z. (2018): Improved YOLO characteristics extraction algorithm and its application in privacy situation detection of service robot. *ACTA Automatica Sinica*, vol. 44, no. 12, pp. 2238-2249.

Zhang, F.; Yang, F.; Li, C. (2019): Rapid vehicle detection method based on improved YOLOv3. *Computer Engineering and Applications*, vol. 55, no. 2, pp. 12-20.

Zheng, Z.; Liu, Y.; Pan, C.; Li, G. (2019): Improved YOLOV3 remote sensing image and application in aircraft identification. *Electro-optical and Control*, pp. 1-6.

Zhou, F.; Jin, L.; Dong, J. (2017): Research on convolutional neural networks. *Chinese Journal of Computers*, vol. 40, no. 6, pp. 1229-1251.