Bridge Component Recognition by Object Detection Technology

University of Tsukuba, Student Member, ○ YU WEILEI
University of Tsukuba, Regular Member, Mayuko Nishio

1. INTRODUCTION

The object detection technology is a kind of computer vision technology, which combines deep learning, convolution neural network, sliding window and principles and methods. Using this technique, the computer can recognize the object in the image and mark the position of the object. The traditional maintenance work often uses visual inspection, but it depends on the technical skill of inspector, and requires a lot of manpower and financial resources. Computer vision inspection technology has the advantages of automation, standardization and low cost. And there have been a lot of studies on the application of computer vision technology in the field of civil engineering. Database on bridge components (bridge piers, T-girder, box girder, trusses) were collected and used to train on the deep learning detection models, the results show that the Yolo-v3 detection model, which combines image argumentation and learning rate-decay technology, can detect bridge components well, which set the foundation for the next step in the study of bridge damage evaluation.

2. APPLICATION OF OBJECT DITECTION TO BRIDGE COMPONENT

2.1 Object detection by YOLO Algorithm

At present, the mainstream detection model is R-CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, and so on. Among them, The YOLO (You Only Look Once) algorithm was originally proposed by Redmon and others in 2016 as a regression-based object detection method, and in 2018 it had developed to the third generation of Yolo-v3[(2)]. the Yolo series model can directly predict the category and location of different targets using only one stage, which is characterized by high speed and precision[(1)] Yolo-v3 network draws on the concept of residual neural network, imports multiple residual network modules and uses multi-scale prediction to improve performance in the identification of small targets. The model uses a number of well-performing 3×3 and 1×1 convolution layer, and some residual network structures for later multiscale predictions, which eventually have 53 convolution layers.

2.2 Application of objective detection to bridge inspection

According to previous research, it is found that bridge detection projects using object detection technology have high demands on the quality of data sets. Through research, it is found that dividing bridge data sets by object and distance can effectively improve detection accuracy. Therefore, the framework of the test is divided according to the detection distance. The Table1 shows the method of classification, and we divide the detection task into three main steps through the far and close range of image recognition.

2.3 Collection of bridge component image data

For the detection object in the step2, collected the bridge component image and train it using Yolo-v3 model. Get database by going to the scene of the existing bridge to photograph the components of the bridge. For an existing bridge, collect the data and images we need according to three different distances, Distant, Close, Ultra-close, as far as possible. mainly collect images of the bridge's main beams and piers taken at close range. Table 2 shows the details of the dataset set up. A total of 160 photos, including bridge piers, truss structures, box beams, T-beams, and according to the training set, validation set, test set to divide. The Yolo-v3 model was used to predict the type and main structure of the bridge, and 260 pictures divided the training sets according to different proportions, and it was found that 80% of the training sets were more accurate than the 60% training sets. Therefore, in this object will also set to 80% of the training set.

2.4 Bridge component detection results and discussions[西尾真由子1]

The model is trained with a GPU (Tesla V100) on a cloud computer (Jupiter Book), also python 3.7 and deep learning framework PaddlePaddle1.8.4 has been used. In order to deal with the disadvantages of the small sample number, Image argument had been set, and Learning Rate Decay strategy had been set to prevent overfit. Figure 1 shows the training loss value; the x-axis is the number of trainings and the y-axis is the loss value generated by the training. we can find that as the training progresses, the loss value obtained by the model becomes lower and lower, eventually tending to the minimum

Table 1 Detection Object and steps

Step1	Distant observation	identify bridge types and large components (cable tower, main arch, main beam
Step2	Close observation	Bridge pier, main beam structure, support, truss, cable, bar, etc.
Step3	Ultra-close observation	Identify damage (cracks, rust, fractures, spalling, holes)

Table 2 The dataset classification table

	Pier	Truss	Box girder	T girder	All
Train	56	42	35	35	112
validation	24	18	15	15	48

Keywords: bridge inspection, object detection,
Contact: Tennodai 1-1-1, Tsukuba, Japan　　Tel：080-9172-9666

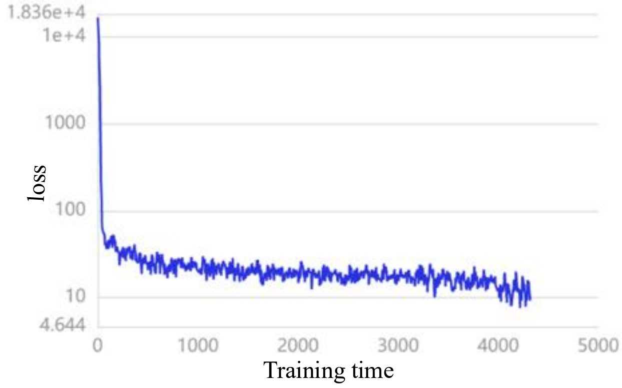

Fig 1 Training loss value

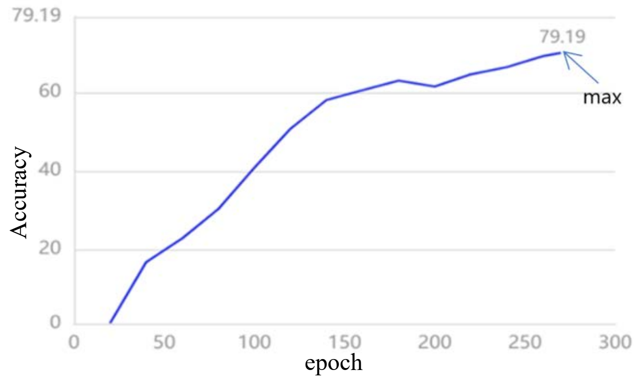

Fig 2 Accuracy of the detection

(1) (2) (3)

Fig3 Visualization results of bridge components inspection from different angles and scene

value of 4.644, It can be considered that the training process of the model is carried out smoothly, and the model is finally converged. Figure 2 shows the change in detection Precision. the x-axis is the number of training stages(epoch) and the y-axis is the precision by the training. Train 16 times in each epoch, and save the model after training in this epoch, and use it for accuracy testing, for a total of 270 epochs like this and the accuracy means the closeness between the prediction box and the truth box. The detection accuracy of the bounding box increases with the number of training epoch and ends up with a maximum value of approximately 79.19. Figure 3 show the results of the test, the object in the box is the detection target. In the frame, the type of prediction and the prediction precision of this type are obtained. The precision is calculated by the overlap size between the prediction box and the actual object box divided by the actual object frame size. Figures 3-(1) and (3) are the detection pictures of the lower part of the bridge, showing that a number of objects have been detected in the image, which were piers and T-girders. The detection accuracy of T-girders is about 0.6, while that of bridge pier is 0.68. Figures 3-(2) is the side image of the bridge, Multiple objects can also be detected, the average detection accuracy of the pier is 0.58, and the detection accuracy of the box-girder is low, about 0.27. From the results, the type of the main girder and the pier are well identified. and can detect multiple objects in a single image at once.

4. CONCLUSION

The model is sensitive to the data sets used for training, and the effects of different data sets greatly. Image argument technology is also applicable to bridge detection projects, which can effectively enhance the detection results. Then, there is good performance when using object detection technology to detect bridge components even though a small sample dataset is used. YOLO-V3 detection model has good applicability in bridge detection project especially for the detection of bridge piers. The model needs to be evaluated next, using indicators such as confusion matrices. Research is also needed on model building and data set collection in Step 3. Damage levels need to be defined with the help of standards. At the same time we need to define damage on accurate components, so need to get special data sets. In addition, with the progress of the YOLO series, YOLO v5 is already available. Therefore, the detection model can be updated. Considering the deployment of UAV, consider upgrading to video detection

REFERENCES

1) Redmon. Joseph, A. Farhadi."YOLOv3: An Incremental Improvement." ［C］ //IEEE Conference on Computer Vision and Pattern Recognition,2018: 1-6.

2) Redmon. Joseph, Divvalas. Girshick R et. al. : You only look once: unified real time object detection ［C］ //IEEE Conference on Computer Vision and Pattern recognition 2016: 779-788 ·