Deep Learning Based Steel Pipe Weld Defect Detection

Dingming Yang\textsuperscript{a}, Yanrong Cui\textsuperscript{a*}, Zeyu Yu\textsuperscript{b} and Hongqiang Yuan\textsuperscript{c}

\textsuperscript{a}School of Computer Science, Yangtze University, Jingzhou 434023, China;

\textsuperscript{b}School of Electronics & Information, Yangtze University, Jingzhou 434023, China;

\textsuperscript{c}School of Urban Construction, Yangtze University, Jingzhou 434000, China;
Deep Learning Based Steel Pipe Weld Defect Detection

Steel pipes are widely used in high-risk and high-pressure scenarios such as oil, chemical, natural gas, shale gas, etc. If there is some defect in steel pipes, it will lead to serious adverse consequences. Applying object detection in the field of deep learning to pipe weld defect detection and identification can effectively improve inspection efficiency and promote the development of industrial automation. Most predecessors used traditional computer vision methods applied to detect defects of steel pipe weld seams. However, traditional computer vision methods rely on prior knowledge and can only detect defects with a single feature, so it is difficult to complete the task of multi-defect classification, while deep learning is end-to-end. In this paper, the state-of-the-art single-stage object detection algorithm YOLOv5 is proposed to be applied to the field of steel pipe weld defect detection, and compared with the two-stage representative object detection algorithm Faster R-CNN. The experimental results show that applying YOLOv5 to steel pipe weld defect detection can greatly improve the accuracy, complete the multi-classification task, and meet the criteria of real-time detection.

Keywords: deep learning; object detection; YOLOv5; X-ray non-destructive testing; weld defect

Introduction

Steel pipes are widely used in high-risk and high-pressure scenarios such as oil, chemical, natural gas, shale gas, etc. If there is some defect in steel pipes, it will lead to serious adverse consequences. With the growing demand for steel pipe in China, more and more enterprises and even countries begin to pay attention to the quality and performance of steel pipe, and the defect detection and evaluation technology of steel pipe has become a research topic that researchers are keen on. At present, there are manual testing and X-ray testing. X-ray testing is one of the main methods for industrial non-destructive testing (NDT), and the test results have been used as an important basis for defect analysis and quality assessment of weld. X-ray detection can effectively detect the internal defects of steel pipe, but manual participation is still needed to
determine the type and location of weld defects of steel pipe (Yun et al. 2009). Therefore, applying object detection in the field of deep learning to the defect detection and identification of steel pipe welds can effectively improve the detection efficiency and promote the development of industrial automation.

With the wide application of artificial intelligence in the field of computer vision, machine learning and deep learning are widely used in object detection and image classification. Most predecessors used traditional computer vision methods to detect steel pipe weld defects (Yun et al. 2009; Wang et al. 2008; Malarvel et al. 2021; Mahmoudi et al. 2009). For example, Wang et al. (2008) used Multi-thresholds+SVM (Support Vector Machine) method to achieve an accuracy of 96.15% for X-ray image detection of steel pipe weld cracks; Malarvel et al. (2021) used OSTU + MSVM-rbf (Multi-class Support Vector Machine) method to achieve multi-class detection of weld defects in X-ray images and achieved an accuracy of 95.23%. Nowadays, object detection algorithms based on deep learning are constantly developing, the recognition accuracy and detection time have been greatly improved compared with traditional computer vision methods. Previous studies have achieved good results, but there are also some shortcomings. Such as:

- Accuracy rate needs to be further improved;
- Different types of defects make it difficult to do multiple classifications with traditional computer vision methods;
- Detection time is too long to achieve real-time detection, so it is difficult to apply to the industrial field;

In view of the above problems, this paper applies the state-of-the-art YOLOv5 to the defect detection task of steel pipe weld.
**Materials and Methods**

**Profile of YOLOv5**

Joseph Redmon et al. (2016a) published YOLOv1 in 2015, which pioneered the single-stage object detection algorithm. This algorithm divides images into 7*7 grids, and each grid is responsible for the classification of objects and coordinate regression at the same time. Joseph Redmon et al. (2016b) published YOLO9000 in 2016 to make up for the shortcoming of YOLOv1 with fewer detection categories and low accuracy, but the detection of small targets is still poor. Joseph Redmon et al. (2018) published YOLOv3 in 2018, which draws on the idea of FPN (Tsung-Yi Lin et al. 2017), and solves the detection problem of small objects. Alexey Bochkovskiy et al. (2020) improved their algorithm by absorbing the tricks of various fields on the basis of the network structure of YOLOv3 and released YOLOv4, which greatly improved the detection efficiency and AP. Two months later Ultralytics (a company) released YOLOv5 (Jocher et al. 2021).

According to the size of the model, YOLOv5 is divided into four versions: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. The larger the model is, the higher the accuracy will be, and the detection time of a single image will increase. Figure 1 shows the network structure of YOLOv5s. The technologies used in the Input of YOLOv5 include Mosaic data enhancement (Sangdoo et al. 2019), adaptive anchor calculation, and adaptive image scaling. The technology used in Backbone includes Focus structure and CSP structure. The techniques used in Neck include FPN+PAN structure; In Prediction, GIoU_Loss(Hamid Rezatofighi et al. 2019) is used to replace the ordinary IoU calculation method. YOLOv5 is slightly less capable than YOLOv4 in terms of performance, but much more flexible and faster than YOLOv4, so it has an advantage in model deployment.
Acquisition of dataset

The raw video images of X-ray are provided by the cooperating factories in RAW format. Through batch processing, the same width and height are cut out and exported as JPG images, and 3408 original images of weld defects of 8 types of steel pipe are obtained. Finally, Labelme (a software to mark object) was used to mark the defect area and defect category of steel pipe weld, which was then exported as the standard dataset format of YOLO or PASCAL VOC2007 (Ren et al. 2017). Figure 2 shows the types of steel pipe weld defects. The collected samples have a total of 8 types of defects, which are Blowhole, Undercut, Broken arc, Crack, Overlap, Slag inclusion, Lack of fusion, and Hollow bead. Table 1 shows the statistical table of steel pipe weld defects samples.
Figure 2. The example of steel pipe defects.

Table 1. Profile of sample images for 8 types of defects.

<table>
<thead>
<tr>
<th>Defect name</th>
<th>Number of original samples</th>
<th>Number of augmented samples</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blowhole</td>
<td>1339</td>
<td>12051</td>
<td>blow-hole</td>
</tr>
<tr>
<td>Undercut</td>
<td>35</td>
<td>315</td>
<td>undercut</td>
</tr>
<tr>
<td>Broken arc</td>
<td>531</td>
<td>4779</td>
<td>broken-arc</td>
</tr>
<tr>
<td>Crack</td>
<td>119</td>
<td>1071</td>
<td>crack</td>
</tr>
<tr>
<td>Overlap</td>
<td>219</td>
<td>1971</td>
<td>overlap</td>
</tr>
<tr>
<td>Slag inclusion</td>
<td>136</td>
<td>1224</td>
<td>slag-inclusion</td>
</tr>
<tr>
<td>Lack of fusion</td>
<td>416</td>
<td>3744</td>
<td>lack-of-fusion</td>
</tr>
<tr>
<td>Hollow bead</td>
<td>613</td>
<td>567</td>
<td>hollow-bead</td>
</tr>
<tr>
<td>Totals</td>
<td>3408</td>
<td>30672</td>
<td>———</td>
</tr>
</tbody>
</table>

**Data preprocessing**

**Raw dataset analysis**

First of all, the original data should be analyzed so as to serve as a reference when setting parameters for deep learning and to accelerate the training speed. It can be seen from observation that X-ray pictures are black and white pictures, which can be converted into single-channel grayscale images. In this way, 2/3 pixels data can be compressed and the training speed will be accelerated. Then use Matplotlib (a python lib to draw diagram) to draw the scatter plot of the center point position of the bounding box and the length & width of the bounding box in turn to see if there are any extreme aspect ratios and abnormal data. As shown in Figure 3, it can be concluded that most bounding boxes are wider than their height and that the bounding boxes for cracked defects are close to a square. Secondly, the displacement of most defects is in the horizontal direction, and the displacement of Overlap defects is from the bottom right to the top left. The distribution of scatter is more even, and there are not many abnormal
Motion deblurring

When the cylindrical steel pipe rotates in the assembly line, there will have relative movement between the X-ray camera used to film the weld defects of the steel pipe and the steel pipe in the direction of the weld. Moreover, the exposure time of the camera to shoot a single frame of weld defects is too long, so the motion blur will be generated. According to the research of Kupyn et al. (2018), motion blur will have an impact on the accuracy of object detection algorithm of YOLO series, so it is necessary to remove motion blur in some images. The process of motion deblurring is shown in Figure 4.

First of all, we use the Hough Transform to detect the straight line at the weld edge. The direction of motion of the steel pipe can be estimated from the angle of the straight line (that is, the angle of image blur), and the distance of motion blur can be obtained from the frame rate of the camera and the speed of the steel pipe rotation. Then we used the estimated blurry kernel to deconvolution the original blurry image to get the result in Figure 4c.
Figure 4. The process of blind motion deblurring.

Data enhancement

Convolutional neural network (CNN) usually requires a large number of training samples to effectively extract image features and classify them. In order to effectively improve data quality and increase data feature diversity, the original data was enhanced to 9 times the original data by using light change, random rotation, random cut out, Gaussian noise addition, horizontal flipping, random adjustment of saturation, contrast and sharpness, random resize and random clipping. Thus, the over-fitting in the training stage is effectively reduced and the generalization ability of the network is improved. Figure 5 shows an example of what happens after data enhancement.

Figure 5. The example after data augmentation.
Experiments

Experimental environment

Table 2 and Table 3 are the hardware environment and software environment of the experiment in this paper.

Table 2. The environment of hardware.

<table>
<thead>
<tr>
<th>Phase</th>
<th>CPU</th>
<th>GPU</th>
<th>RAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>Intel(R) Xeon(R) CPU E5-2623 v4 @ 2.60GHz</td>
<td>Quadro P5000</td>
<td>30GB</td>
</tr>
<tr>
<td>Test</td>
<td>Intel(R) Core(TM) i7-4710MQ CPU @ 2.50GHz</td>
<td>GTX950M</td>
<td>8GB</td>
</tr>
</tbody>
</table>

Table 3. The environment of software.

<table>
<thead>
<tr>
<th>Phase</th>
<th>OS</th>
<th>Python</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>Linux-5.4.0-65-generic-x86_64-with-glibc2.10</td>
<td>3.8.5</td>
<td>Official YOLOv5x</td>
</tr>
<tr>
<td>Test</td>
<td>Windows 10 professional edition</td>
<td>3.8.0</td>
<td>Official YOLOv5x</td>
</tr>
</tbody>
</table>

Experimental process

In this paper, the state-of-the-art deep learning algorithm YOLOv5 is used to train the detection model of steel pipe weld defects. After manually annotating the original image, the dataset is obtained through data enhancement, and then the dataset is converted into a single channel grayscale image. Because the dataset is relatively small, it is divided into training set and validation set in a ratio of 8:2. An experimental process designed in this paper is shown in Figure 6. After several epochs of YOLOv5 training, the training set and validation set obtained a model containing weight and bias parameters. In this paper, Recall, Precision, F1 score, mAP(mean of Average Precision) and detection time of single image were used as evaluation indexes.
The calculation method of Precision is shown in Formula (1). TP is the sample identified as true positive. In this paper, it is the identification of correct weld defects of steel pipe. FP is the sample identified as false positive. In this paper, it is the weld defect of steel pipe identified wrongly. The formula describes the proportion of true positive in the identified pictures of steel pipe weld defects. The calculation method of Recall is shown in Formula (2). FN is the sample identified as false negative, in this paper is the background of error identification; The formula describes the ratio of the number of correctly identified steel pipe weld defects to the number of all steel pipe weld defects in the dataset. The calculation method of F1 score is shown in Formula (3). When Precision and Recall are required to be high, F1 score can be used as an evaluation index. The calculation method of AP is shown in Formula (4). AP is introduced to solve the problem of limitation of Precision, Recall and F1 score single point value. In order to obtain an indicator that can reflect the global performance, In this paper, using the Interpolated average precision.

\[
Precision = \frac{TP}{TP + FP} \quad (1)
\]

\[
Recall = \frac{TP}{TP + FN} \quad (2)
\]

\[
F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)
\]
\[ P_{\text{interp}}(k) = \max_{k \geq k} P(\hat{k}) \]
\[ AP = \sum_{k=1}^{N} P_{\text{interp}(k)} \Delta r(k) \]

**Analysis of experimental results**

**Identify results and data analysis**

The detection result for 8 types of defects are shown in Figure 7. On the whole, both the position of defects and the confidence of classification are relatively good. Undercut's good performance in the case of a relatively small number of samples could not be attributed to the 8 data enhancement methods used in the data preprocessing stage of this paper and Mosaic data enhancement by YOLOv5. The Broken can still be identified as the same defect and obtain good confidence even if they are very different in appearance. Among them, the Slag inclusion defects are not obvious to distinguish from the background in the naked eye, and they are similar to the Undercut defects in appearance. Benefiting from repeated training, good results can also be achieved.

![Figure 7. The result of detection.](image)

As shown in Table 4, four evaluation indexes of each defect category in the last epoch are presented. On the whole, except for Blowhole defect, the accuracy of all other defects can be maintained between 0.962 and 1.00, the recall rate between 0.99 and
1.00, and the F1 score between 0.998 and 1.00. Blowhole defect due to its small defect target, a single steel pipe sometimes has dense pores, so the accuracy is lower than other types of defects. In the 218th epoch, the mAP of the model reached 99.02%, but after 633 epochs of training, the mAP decreased to 98.71%, showing some degree of overfitting. The best training model saved in this paper can be used in the actual steel pipe weld defect detection and applied in the industrial production environment.

Table 4. Some statistical parameters of confusion matrix

<table>
<thead>
<tr>
<th>Type</th>
<th>blowhole</th>
<th>undercut</th>
<th>broken-arc</th>
<th>crack</th>
<th>overlap</th>
<th>slag-inclusion</th>
<th>lack-of-fusion</th>
<th>hollow-bead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>0.505</td>
<td>1.00</td>
<td>0.962</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Recall</td>
<td>0.96</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>F1 score</td>
<td>0.661</td>
<td>1.00</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.994</td>
</tr>
<tr>
<td>AP</td>
<td>0.951</td>
<td>0.995</td>
<td>0.992</td>
<td>0.995</td>
<td>0.995</td>
<td>0.995</td>
<td>0.978</td>
<td>0.995</td>
</tr>
<tr>
<td>mAP@0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.987</td>
<td></td>
</tr>
</tbody>
</table>

*Performance comparison of weld defect detection algorithm for steel pipe*

As shown in Figure 8, we used the same dataset to conduct experiments respectively in Faster R-CNN (Ren et al. 2017; Bubbling 2020) and YOLOv5 (Jocher et al. 2021), then compared the precision data and total loss data generated during the experiment. As shown in Figure 8a, Faster R-CNN calculates the precision mean after each epoch of training, and has a tendency of descending and then slowly climbing, with unstable values in the second half. YOLOv5, on the other hand, started off with a shaky precision, then slowly climbed up and settled down. As shown in Figure 8b, the total loss of Faster R-CNN tended to be stable between 50-100 epoch, and then had two relatively large wave peaks. Since Faster R-CNN uses the Adam (Diederik Kingma et al. 2014) optimizer, it can converge faster than SGD (Stochastic Gradient Descent). The initial total loss of YOLOv5 was relatively small and tended to be stable between 100-
150 epoch, with a small peak around 160 epoch. YOLOv5 also uses the optimizer Adam, and the initial value of Momentum is 0.999. In general, compared with the Faster R-CNN, YOLOv5 has better convergence speed in precision & total loss and stability after convergence than the Faster R-CNN.

Figure 8. Compare with faster R-CNN

As shown in Table 5, a comparison is made between Multi-thresholds+SVM, OSTU+SVM, Faster R-CNN+ResNet50 and YOLOv5. On the whole, the defect detection algorithm based on deep learning is better than the defect detection algorithm based on traditional computer vision in both performance and detection time of a single image. Among them, Multi-thresholds+SVM algorithm takes the longest time. Although the accuracy of the Multi-threshold+SVM algorithm is better than that of Faster R-CNN, the detection time of single image is not as good as Faster R-CNN. YOLOv5 is superior to Faster R-CNN in both accuracy and detection time of a single image. The detection time of a single image satisfies the engineering work of the model in the later stage of this paper. YOLOv5's detection speed is to be expected because it’s one-stage. Another kind of object detection algorithms is two-stage. For example, the Faster R-CNN algorithm forms region proposals (which may contain areas of the object) first and then classifies each region proposal (also corrects the position at mean
time). This type of algorithm is relatively slow because it requires multiple runs of the detection and classification process.

Table 5. Performance comparison of steel pipe defection algorithms

<table>
<thead>
<tr>
<th>Object detection model</th>
<th>Accuracy or Precision/%</th>
<th>Detection time per picture/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-thresholds+SVM (Wang et al. 2008)</td>
<td>96.15 acc</td>
<td>180 (Mahmoudi et al. 2009)</td>
</tr>
<tr>
<td>OSTU+MSVM-rbf (Malarvel et al. 2021)</td>
<td>95.23 acc</td>
<td>—</td>
</tr>
<tr>
<td>Faster R-CNN+ResNet50 (Ren et al. 2017)</td>
<td>95.5 acc (mAP@0.5=78.1)</td>
<td>0.437</td>
</tr>
<tr>
<td>YOLOv5x (Jocher et al. 2021)</td>
<td>97.8 pre (mAP@0.5=98.7)</td>
<td>0.120</td>
</tr>
</tbody>
</table>

**Conclusion**

In the field of steel pipe weld defect detection, deep learning method has more advantages than traditional computer vision method. Convolutional neural network does not need to extract image features manually, and can realize end-to-end input detection and classified output. The research of this paper has the following three contributions:

- Applying the state-of-the-art object detection algorithm YOLOv5 to the field of steel pipe weld defects detection, The detection accuracy of steel pipe weld defects and the detection time of a single image are pushed to a new height level, with the accuracy reaching 97.8% (mAP@0.5=98.7%). Under the YOLOv5x model testing, the detection time of a single picture is 0.12s (GPU=GTX950M), which meets the real-time detection on the steel pipe production line;
- Did a lot of work in the data preprocessing stage, combining the traditional data enhancement method with the Mosaic data enhancement method of YOLOv5, which not only greatly increased the size of the dataset, but also effectively reduced the over-fitting of the training;
The results of YOLOv5 were compared with previous defect detection algorithms, and the advantages of YOLOv5 in model deployment and model engineering were demonstrated on the premise of comprehensive indicators.

This study can provide methods and ideas for real-time automatic detection of weld defects of steel pipe in industrial production environment, and lay a foundation for industrial automation. Although this paper uses state-of-the-art deep learning algorithm and convolutional neural network model for real-time detection of steel pipe weld defects in industrial production scenarios, its performance and performance are also relatively good. However, in the case of limited dataset, other defects which not in the dataset cannot be correctly identified. In this case, we can use traditional computer vision or mathematical methods to build an expert system to identify other defects that do not appear in the dataset. It is also possible to design an automatic updating model system in combination with few-shot learning in engineering, which is used to manually label the type and bounding box coordinate information by the quality inspector when the defect cannot be identified, so that the system can automatically learn and update the model. These deficiencies point out the direction and provide ideas for the follow-up research.

References


Author Contributions

Conceptualization, D.Y., Y.C., and Z.Y.; Software, D.M., Resources, Z.Y., H.Y.; Supervision, Y.C., Z.Y.; Writing—original draft, D.Y.; Writing—review and editing, D.Y., Y.C., and Z.Y. All authors have read and agreed to the published version of the manuscript.