FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1140-1145 | e-ISSN: 2347-470X

Research on Steel Surface Defect Detection Algorithm Based on Improved Deep Learning

Fei Ren[1], GuangRong Wang[2], ZhiQi Hu[3], MinNing Wu[4] and Madhavi Devaraj[5]
[1]*School of Information Technology, Mapua University, Manila, Philippines, renfeivision@outlook.com*
[2]*Sinoma (Nanjing) Mining Research Institute Co., Ltd, Nanjing, China, 2058087205@qq.com*
[3]*Sinoma Mining Construction Co., Ltd., Nanjing, China, rf1314@vip.qq.com*
[4]*School of Information Engineering, Yulin University, Yulin, Shaanxi, cvvision@foxmail.com*
[5]*School of Information Technology, Mapua University, Manila, Philippines, madhavidevaraj@gmail.com*

***Correspondence:** Madhavi Devaraj; madhavidevaraj@gmail.com

ABSTRACT- With the development of industrial economy, more and more enterprises use machine vision and artificial intelligence to replace manual detection. Therefore, the research of steel surface defect detection based on artificial intelligence is of great significance to promote the rapid development of intelligent factory and intelligent manufacturing system. In this paper, Yolov5 deep learning algorithm is used to build a classification model of steel surface defects to realize the classification and detection of steel surface defects. At the same time, on the basis of Yolov5, combined with the attention mechanism, the backbone network is improved to further improve the classification model of steel surface defects. The experiment shows that the Recall and mAP of improved Yolov5 perform better on the steel surface defect data set. Compared with Yolov5, the number of C3CA-Yolov5 parameters decreased by 13.02%, and the size of *pt* files decreased by 12.72%; the number of C3ECA-Yolov5 parameters decreased by 13.36%, and the size of *pt* files decreased by 13.22%.

General Terms: Artificial intelligence, Machine learning, Algorithms *et. al.*
Keywords: Machine Vision, Artificial Intelligence, Deep learning, Steel surface defects, Attention mechanism.

ARTICLE INFORMATION
Author(s): Fei Ren, GuangRong Wang, ZhiQi Hu, MinNing Wu and Madhavi Devaraj;
Received: 17/08/2022; **Accepted**: 06/12/2022; **Published**: 20/12/2022;
e-ISSN: 2347-470X;
Paper Id: IJEER 1708-22;
Citation: 10.37391/IJEER.100461
Webpage-link:
www.ijeer.forexjournal.co.in/archive/volume-10/ijeer-100461.html

crossref member
CROSSREF.ORG
THE CITATION LINKING BACKBONE

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

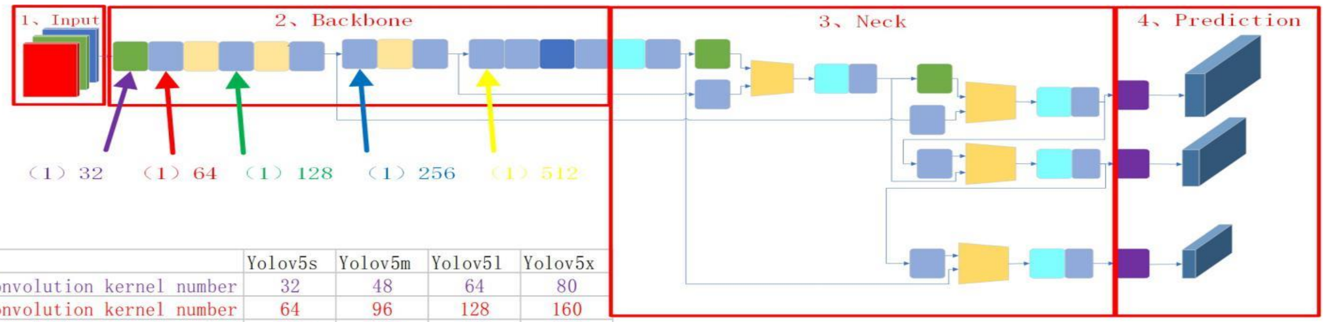
With the development of advanced technology, machine vision technology is more and more widely used in industry. In actual production, using machine vision technology to replace workers has become the choice of more and more enterprises. Using artificial intelligence and machine vision technology to detect steel surface defects has become the focus of many scholars.
Many scholars have carried out in-depth research on surface defects based on machine vision and artificial intelligence technology to achieve industrial intelligence.

Shen Xizhong [1] and others proposed a surface defect detection model for aluminum profiles based on yolo-v5. This model not only expands the feature extraction and detection scale of the original model, but also improves the calculation method of the adaptive anchor frame, which greatly improves the detection performance of the network for small targets. However, the AP is relatively low in the detection of scratch defects, and the feature extraction needs to be further strengthened; Lou Xudong [2] proposed a method for detecting the defects of the installation holes of expansion valves based on machine vision. The defects of the installation holes are extracted by traditional feature matching and other methods, which can meet the requirements of automatic detection of the defects of the installation holes of expansion valves in actual production. Moreover, the data sample size is not high. However, when the defect area is large, there may be false labeling of defects, and the generalization of the algorithm needs to be verified. Li Ke[3] and others introduced mobile net and spatial attention module based on the u-net network model, and achieved good detection results for the weld bubble defect in the chip X-ray image. However, they did not further study the reasoning time of the model. Erozan Ahmet Turan[4] proposed a method to detect defective transistors in printed circuits through optical detection. 95% of defective transistors can be correctly detected by manually extracting image features and using supervised learning classification, and the execution time is considered. However, it needs to be further expanded to enhance the stability of the data set verification algorithm; Kento Nakashima [5] realized the defect detection of wrap FLM product through the CNN network named sssNet. While developing the corresponding upper computer interface, the accuracy and recall of the model were 0.999 and 0.952, respectively. However, the template matching was greatly affected by the pixel value. The target film area extracted by this method as the input image may affect the detection accuracy in actual industrial applications. Aiming at the problem of steel surface defect detection, this paper realizes target detection based on Yolov5 deep learning model, and improves Yolov5 by combining CoordAttention and ECA attention mechanism. Finally, the performance of steel surface defect detection network can be effectively improved.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1140-1145 | e-ISSN: 2347-470X

2. DEEP LEARNING YOLOV5 ALGORITHM

Yolo series algorithm is an end-to-end deep learning target detection algorithm. Because of its advantages such as fast inference speed, high detection accuracy and balanced performance, this series of algorithms have been widely concerned by the industry since they came out. Yolov5 is continuously improved on the basis of previous generations, with faster detection speed, higher detection accuracy and smaller model volume [6-7]. Yolov5 is divided into four types: Yolov5s, Yolov5m, Yolov5l and Yolov5x. The depth and width of the network backbone increase in turn. In general, the detection time and effect also increase. Based on the yolo5m deep learning model with balanced accuracy and time, this paper improves the attention mechanism and establishes a steel surface defect detection algorithm. The network structure of Yolov5 is shown in *figure 1*.



	Yolov5s	Yolov5m	Yolov5l	Yolov5x
(1) Convolution kernel number	32	48	64	80
(2) Convolution kernel number	64	96	128	160
(3) Convolution kernel number	128	192	256	320
(4) Convolution kernel number	256	384	512	640
(5) Convolution kernel number	512	768	1024	1280

Figure 1: Four network structures of Yolov5 and the number of corresponding convolution cores

Yolov5 model includes the following four parts: input, backbone, neck and head [8]. The input part is enhanced by mosaic data, and four images are selected for random splicing each time to generate a new target data picture, which greatly enriches the diversity of neural network learning data; Backbone realizes the extraction of image features, which is mainly realized by focus, C3, spp and other modules, and extracts the network through convolution and other operations. Focus obtains a double down sampling feature map by slicing and convolution, and realizes speed increase without information loss; The C3 module improves the convolution layer on the basis of the BottleneckCSP module to accelerate the reasoning speed; The SPP module is called the spatial pyramid pooling layer, which can convert any size feature map into a fixed size feature vector, and achieve the consistency of input features through maximum pooling; Neck is mainly used to generate pyramid features, which enhances the fusion effect of different scales of the same object; Head cuts the channel and outputs the vector data predicted by the network including the category of prediction box, confidence and coordinate position. As shown in the figure, the number of convolution kernels of Yolov5s, Yolov5m, Yolov5l and Yolov5x structures in Yolov5 is different at different stages. Take the second convolution as an example, Yolov5s uses 64 convolution kernels, so the corresponding Eigen graph is 152 * 152 * 64. While Yolov5m uses 96 feature maps, and the size of the feature map is 152 * 152 * 96. The more the number of convolution kernels, the wider the thickness, that is, the width of the feature map, the stronger the learning ability of the network to extract features, but the longer the corresponding detection time[9]. In this paper, Yolov5m with balanced detection effect and efficiency is used to detect the surface defects of steel.

3. STEEL SURFACE DEFECT DETECTION ALGORITHM BASED ON IMPROVED YOLOV5

3.1 C3 Module

Yolov5 has been continuously improved since its inception. The original BottleeckCSP module in the backbone has been replaced by the C3 module. The structure and function of the C3 module are basically the same as that of the BottleneckCSP architecture, except for the selection of the correction unit. The C3 structure is divided into two branches. One uses three standard convolution layers and multiple Bottleeck modules, the other only uses basic convolution, and finally uses concat operation to merge the two branches.

3.2 Deep Learning Network Based on C3CA-Yolov5

The research on lightweight networks shows that the channel attention mechanism can significantly improve the performance of the model, but the channel attention usually ignores the extremely important location information. Therefore, Qibin Hou[10] of the National University of Singapore and others proposed a new attention mechanism that embeds location information into channel attention, called coordinated attention (abbreviated as coordinated attention, also known as CA). The traditional channel attention is to transform the feature tensor into a single feature by using 2-dimensional global pooling. CoordAncement embeds location information into the attention channel and decomposes the attention channel into aggregated features in different directions. The one-dimensional features in two different directions represent long-range dependence and accurate position information respectively. Finally, they are fused in a weighted manner on the channel to achieve enhanced

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1140-1145 | e-ISSN: 2347-470X

expression of the target of interest and improve the expression ability of the features [11]. The characteristic diagram formula with both horizontal and vertical directions is shown in formula (1):

$$f = \delta(F(|z^h, z^w|)) \qquad (1)$$

Where, F is the 1x1 convolution function, δ represents the nonlinear activation function, z^h represents the horizontal direction characteristic map, and z^w represents the vertical direction characteristic map.

CoordAncement can not only capture cross channel information, but also obtain direction and position information. Due to its flexible and lightweight characteristics, it can achieve better integration with classic deep learning modules such as mobilenetv2. Its specific structure is shown in *figure 2*. The function of enhancing features is realized through coordinate information embedding and coordinate attention generation [12].

The CoordAttention attention mechanism added to the C3 structure of YOLOv5 algorithm in this paper is the C3CA module. In this paper, all C3 structures in the original algorithm are replaced by the newly written C3CA module in YOLOv5's backbone, so that the entire network structure is lighter and the improved depth learning steel surface defect detection based on YOLOv5 is realized.

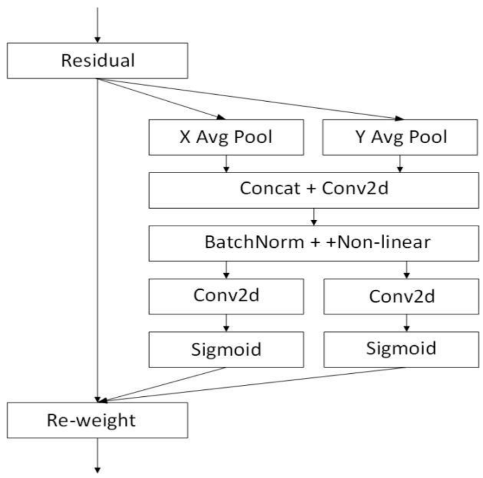


Figure 2: CoordAttention network structure

3.3 Deep Learning Network Based on C3ECA-Yolov5

Efficient channel attention module (ECA for short). In 2020, Qi Long Wang [13] and other scholars proposed an efficient local cross-channel interaction strategy without dimension reduction, which uses one-dimensional convolution to efficiently achieve local cross-channel interaction. ECA can effectively avoid the impact of data dimensionality reduction on the attention learning effect of the channel. At the same time, ECA involves fewer parameters and can effectively overcome the contradiction between performance and complexity.

As shown in *figure 3*, after the global average pooling layer, the 1x1 convolution layer is used and the full connection layer is removed to achieve the capture of local cross channel interaction information. The convolution kernel size k of 1x1 convolution layer represents the coverage of local cross channel interaction [14-15]. In different network models, the convolution kernel size k can be adjusted to the best by manual testing, but manual adjustment requires a lot of computing power and development time. There is a mapping relationship between the convolution kernel size k and the channel dimension C, as shown in formula (2). The convolution kernel size k can be calculated adaptively by using the mapping relationship, as shown in formula (3):

$$C = \phi(k) = 2^{\gamma \times k - b} \qquad (2)$$

$$k = \psi(C) = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{odd} \qquad (3)$$

Where, $|t|_{odd}$ represents the nearest odd neighbors of t, γ sets to 2 and b sets to 1. The convolution kernel size of one-dimensional convolution can be adaptively calculated by *equation (3)*.

The ECA attention mechanism added to the C3 structure of YOLOv5 algorithm in this paper is the C3ECA module. In this paper, all C3 structures in the original algorithm are replaced by the newly written C3ECA module in YOLOv5's backbone, so that the entire network structure is lighter and the improved depth learning steel surface defect detection based on YOLOv5 is realized.

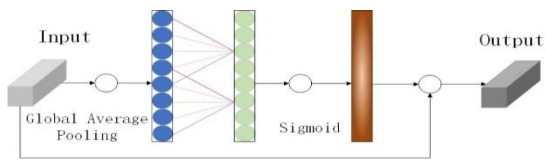


Figure 3: ECA network structure

4. EXPERIMENTAL VERIFICATION AND COMPARISON

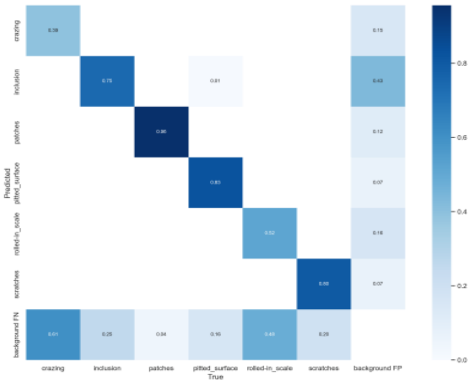
4.1 Experimental Environment and Data Set

The experimental environment of this paper is based on Windows 10 operating system, with 32GB memory, using pytorch1.8 as the deep learning framework, python version 3.8, CUDA version 11.3, and Intel CPU i7-11700@2.50GHz.The GPU is NVIDIA GeForce RTX 3080ti. The data set used in the experiment is the data set of steel surface defects produced by song kechen's team of Northeast University of China, including 6 types of defects: cracking, inclusion, patches, pitted surface, rolled scale and scratches, there are 1800 pictures in total. In this paper, the training set, verification set and test set are divided into 6:2:2. The data set is 200x200 gray scale map.

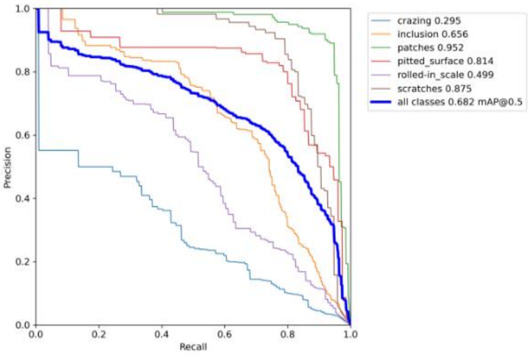
4.2 Experimental Results

4.2.1 Yolov5 Deep Learning Model

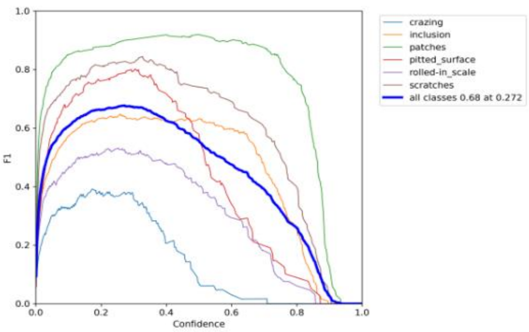
The best weight is obtained after 200 times of training. The results of the test set are as follows: mAP@0.5:0.682 and F1 is 0.680. The results are shown in *figure 4*.



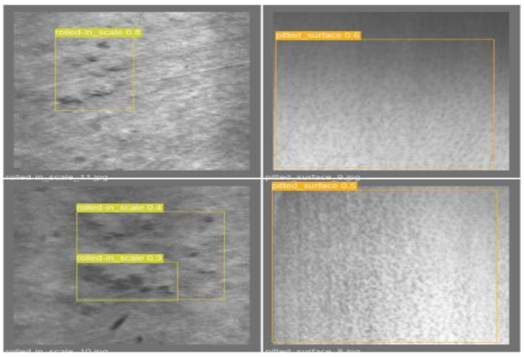
(a)



(b)



(c)



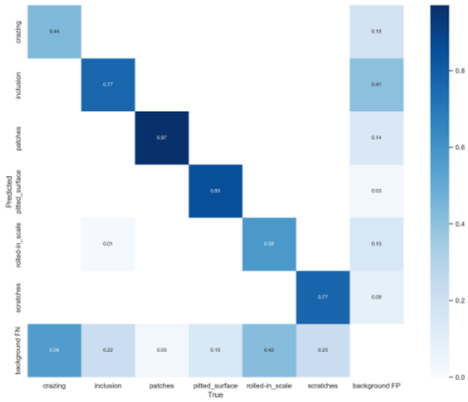
(d)

(a) Confusion matrix (b) PR curve diagram (c) F1 diagram (d) verification detection diagram

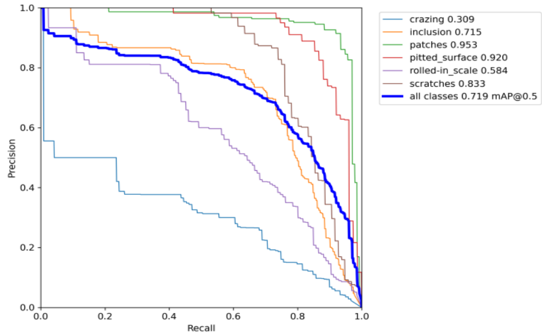
Figure 4: Yolov5 detection results

4.2.2 Improved Yolov5 Deep Learning Model

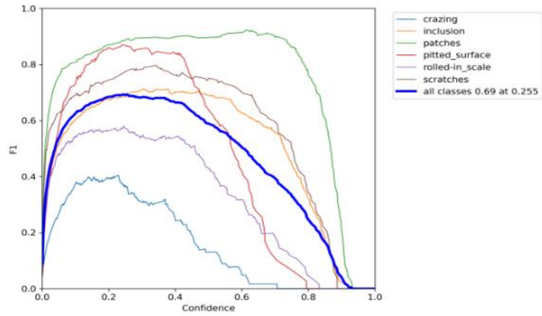
Improved C3CA-Yolov5: the best weight is obtained after 200 times of training. The results of the test set are as follows: mAP@0.50: 0.719 and F1 is 0.690. The results are shown in *figure 5*.



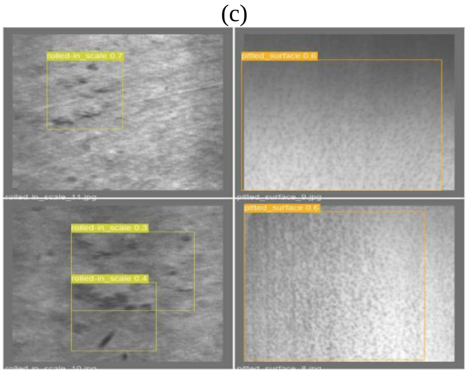
(a)



(b)



(c)



(d)

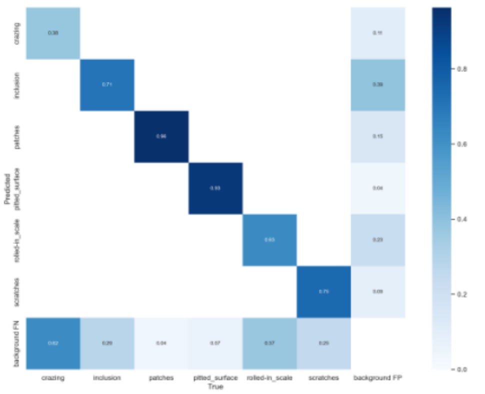
(a) Confusion matrix (b) PR curve diagram (c) F1 diagram (d) verification detection diagram

Figure 5: C3CA-Yolov5m results

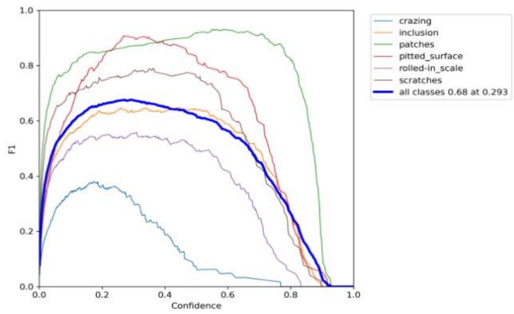
FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1140-1145 | e-ISSN: 2347-470X

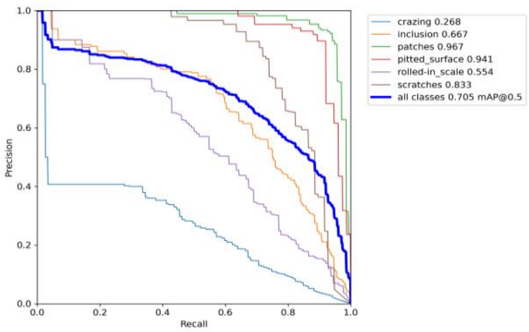
Improved C3ECA-Yolov5: the best weight is obtained after 200 times of training. The results of the test set are as follows: mAP@0.5: 0.680 and F1 is 0.705. The results are shown in *figure 6.*



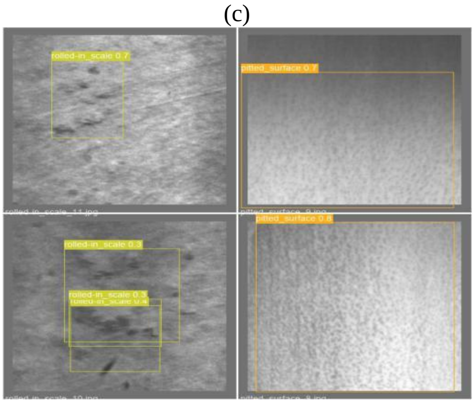
(a)



(b)



(c)



(d)

(a) Confusion matrix (b) PR curve diagram (c) F1 diagram (d) verification detection diagram

Figure 6: C3ECA-Yolov5m results

The comparison of detection results of three deep learning networks is shown in *table 1*. It can be seen from *table 1* that the improved Yolov5 performs well in the steel surface defect data set when the best model is obtained by training 200 rounds in this experiment. When the precision is the same, the Recall and mAP of improved Yolov5 perform better. Among them, compared with Yolov5, the number of C3CA-Yolov5 parameters decreased by 13.02%, the size of *pt* files decreased by 12.72%, and the reasoning time increased by 1.7ms; The number of C3ECA-Yolov5 parameters decreased by 13.36%, the size of *pt* files decreased by 13.22%, and the reasoning time increased by 0.3ms.

Table 1: Comparison of results of different network structures

Type	Yolov5	C3CA-Yolov5m	C3ECA-Yolov5m
Precision	0.689	0.689	0.680
Recall	0.671	0.702	0.680
mAP@.5	0.682	0.719	0.705
layers	290	416	332
parameters	20873139	18156075	18084141
inference	3.9ms	5.6ms	4.2ms
the size of *pt* files	40.1MB	35.0MB	34.8MB

5. CONCLUSION

In this paper, the steel surface defect images are classified by the deep learning classification model, and the target detection is carried out by using the training data of the improved Yolov5 deep learning neural network. The test on the steel surface defect data set shows that compared with the traditional Yolov5 deep learning neural network, the improved Yolov5m defect detection classification model has more advantages. The inference speed is not much different, the number of parameters and the size of *pt* files are reduced, and the Recall and mAP (Mean Average Precision) of target recognition are further improved. The improved deep learning neural network adopted in this paper provides a reference for the follow-up study of steel surface defects and promotes the development of intelligent industry. However, in actual production, the number of data sets is small, so how to further effectively enhance the data sets and realize deep learning defect detection in the case of small samples needs further research.

6. ACKNOWLEDGEMENTS

The authors declare no competing financial interest. The work was supported by grants from the industry university research project of Yulin science and Technology Bureau, China (cxy-2021-102-02). We also would like to express our sincere gratitude to the editor and anonymous reviewers for their valuable comments, which have greatly improved this paper.

REFERENCES

[1] Shen Xizhong & Wu Di. (2022). Detection of small defects on the surface of aluminum alloy based on Yolo Journal of Zhejiang University of Technology (04), 372-380

[2] Lou Xudong, Cao songxiao, Jiang Qing, Xu Zhipeng & Tang Jianbin. (2022). Detection method of installation hole defect of expansion valve based on machine vision Machine tools and hydraulics (13), 90-96

[3] Li Ke, Wu Zhongqing, Ji Yong & Su Lei. (2022). Improving the detection method of weld bubble defect in u-net chip X-ray image Journal of Huazhong University of science and Technology (NATURAL SCIENCE EDITION) (06), 104-110. Doi: 10.13245/j.hust.220613

[4] Erozan Ahmet Turan and Bosse Simon and Tahoori Mehdi B. (2021). Defect Detection in Transparent Printed Electronics Using Learning-Based Optical Inspection. IEEE TRANSACTIONS ON VERY LARGE-SCALE INTEGRATION (VLSI) SYSTEMS, 29(8), pp. 1505-1517.

[5] Kento Nakashima et al. (2021). Defect detection in wrap film product using compact convolutional neural network. Artificial Life and Robotics, 26(3), pp. 1-7.

[6] Sun Yuhang et al. (2022). Vision-Based Detection of Bolt Loosening Using Yolov5. Sensors, 22(14), pp. 5184-5184.

[7] Chiriboga Matthew,Green Christopher M.,Hastman David A.,Mathur Divita,Wei Qi,Díaz Sebastían A& Veneziano Remi.(2022).Rapid DNA origami nanostructure detection and classification using the Yolov5 deep convolutional neural network. Scientific Reports (1).doi:10.1038/S41598-022-07759-3.

[8] Sozzi Marco et al. (2022). Automatic Bunch Detection in White Grape Varieties Using YOLOv3, YOLOv4, and Yolov5 Deep Learning Algorithms. Agronomy, 12(2), pp. 319-319.

[9] Nguyen HungCuong et al. (2022). Unified End-to-End Yolov5-HR-TCM Framework for Automatic 2D/3D Human Pose Estimation for Real-Time Applications. Sensors, 22(14), pp. 5419-5419.

[10] Q. Hou, D. Zhou and J. Feng, "Coordinate Attention for Efficient Mobile Network Design," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 13708-13717, doi: 10.1109/CVPR46437.2021.01350.

[11] Pikun Wang et al. (2022). Unmanned aerial vehicles object detection based on image haze removal under sea fog conditions. IET Image Processing, 16(10), pp. 2709-2721.

[12] Gu Ran,Wang Guotai,Song Tao,Huang Rui,Aertsen Michael,Deprest Jan... & Zhang Shaoting.(2020).CA-Net: Comprehensive Attention Convolutional Neural Networks for Explainable Medical Image Segmentation.. IEEE transactions on medical imaging. doi:10.1109/TMI.2020.3035253.

[13] Wang Q.,Wu B.,Zhu P.,Li P.,Zuo W. & Hu Q..(2020).ECA-Net: Efficient channel attention for deep convolutional neural networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. doi:10.1109/CVPR42600.2020.01155.

[14] Qu Zhong,Zhang Run & Bao Kanghua.(2021).A keypoint-based object detection method with wide dual-path backbone network and attention modules. IET Image Processing (8). doi:10.1049/IPR2.12152.

[15] oxu Qia, Qiao Xu, Liu Changchun, Wang Xinpei, Liu YuanYuan, Yao Lianke & Zhang Huan.(2021).Automated ECG classification using a non-local convolutional block attention module. Computer Methods and Programs in Biomedicine (prepublish). doi:10.1016/J.CMPB.2021.106006.

[16] Prof. Kalyani Kurundkar and Dr. G. A. Vaidya (2021), Application of HFPSO-TOPSIS approach for optimally locating and sizing of reactive power compensating devices for voltage control ancillary service. IJEER 9(3), 16-26. DOI: 10.37391/IJEER.090301.

[17] Swarnima Singh, Vikash Yadav (2021), An Improved Particle Swarm Optimization for Prediction of Accident Severity. IJEER 9(3), 42-47. DOI: 10.37391/IJEER.090304.

[18] Dr. J. S. Awati, Prof. S.S. Patil and Dr. M.S. Kumbhar (2021), Smart Heart Disease Detection using Particle Swarm Optimization and Support Vector Machine. IJEER 9(4), 120-124. DOI: 10.37391/IJEER.090405.

Author Details

Author 1: Fei Ren, male, doctoral candidate, mainly engaged in machine vision and artificial intelligence.
Author 2: GuangRong Wang, male, bachelor, senior engineer, and his main research field is intelligent factory and neural network.
Author 3: ZhiQi Hu, male, bachelor, senior engineer, mainly engaged in mechanical automation and neural network.

Author 4: MinMing Wu, female, doctoral candidate, associate professor, mainly engaged in machine learning and artificial intelligence.
Author 5: Madhavi Devaraj, corresponding author, female, doctor, professor, whose main research fields are artificial intelligence and modular computing theory.