



Article Research on Fault Diagnosis of Steel Surface Based on Improved YOLOV5

Weimin Liu¹, Yao Xiao¹, Aiyun Zheng^{1,*}, Zhi Zheng¹, Xiaojie Liu², Zhen Zhang¹ and Chen Li¹

- ¹ College of Mechanical Engineering, North China University of Science and Technology, Tangshan 063210, China
- ² College of Metallurgy & Energy, North China University of Science and Technology, Tangshan 063210, China
- Correspondence: zay@ncst.edu.cn

Abstract: Steel is an important raw material of fluid components. The technological level limitation leads to the surface faults of the steel, thus the key to improving fluid components quality is to diagnose the faults in steel production. The complex shape and small size of steel surface faults result in the low accuracy of the diagnosis, and the large size of the network leads to poor real-time performance. Therefore, aiming at the problems, an improved YOLOV5 is proposed. Firstly, to reduce the feature information loss, coordinate attention is used to improve YOLOV5, thus the diagnosis ability can be improved. Secondly, to further reduce the loss, a new connection is constructed in YOLOV5, and the detection ability can also be further improved. Thirdly, to improve the real-time performance of the fault diagnosis, YOLOV5 is improve the accuracy, the cosine annealing with warm restarts algorithm is used to optimize YOLOV5. The dataset of NEU-DET is verified and testified. The results show that improved YOLOV5 can diagnose steel surface faults with high efficiency and accuracy.

Keywords: YOLOV5; fluid components; fault diagnosis; steel surface; deep learning



Citation: Liu, W.; Xiao, Y.; Zheng, A.; Zheng, Z.; Liu, X.; Zhang, Z.; Li, C. Research on Fault Diagnosis of Steel Surface Based on Improved YOLOV5. *Processes* **2022**, *10*, 2274. https:// doi.org/10.3390/pr10112274

Academic Editor: Jie Zhang

Received: 1 October 2022 Accepted: 31 October 2022 Published: 3 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

Steel is an important raw material of fluid components, and it is widely used in hydraulic piston pumps [1–3], ball valves [4], Cylinders [5], Steam turbines [6], and other equipment. Due to the limitation of the equipment and technological level, it is inevitable that there are different types of surface faults in the processing. Therefore, in order to improve the steel quality, it is necessary to strengthen the supervision of surface faults in production. Traditional steel fault diagnosis is done by manual visual inspection, but the diagnosis accuracy and efficiency are low. With the rapid development of deep learning, it is introduced into the fault diagnosis of the steel surface. In the past few years, many scholars have studied various deep-learning methods to diagnose steel surface faults [7–10].

The large size of the deep learning networks leads to poor real-time performance. Therefore, lightweight networks such ShuffleNet [11,12], MobileNet [13,14], SqueezeNet [15], and EfficientNet [16] are proposed and widely studied. The ShuffleNet, MobileNet, SqueezeNet, and EfficientNet are also widely used in fault diagnosis. Nagy et al. proposed a new EfficientNet-B7 unit based on EfficientNet, and it is designed for small networks [17]. Liu et al. proposed a new binary network called BiShuffleNet by taking ShuffleNetV2 as the binary backbone network [18]. Liu et al. incorporated multi-scale design into MobileNetV2, and the computation cost is reduced [19]. Fu et al. used a global average pooling to replace the fully-connected layer in SqueegeNet to reduce the parameter amount [20]. Li et al. proposed a network called MobileNet-SSD by introducing a single-shot multibox detector, thus the parameter amount can be reduced [21].

Small fault diagnosis is a difficult and hotspot research in computer vision. Driven by deep learning, small target detection has made significant breakthroughs and has been successfully applied in defense security [22], intelligent transportation [23], and industrial automation [24–26]. Many scholars at home and abroad have conducted in-depth research on small target detection by using YOLOV5 networks. Qiu et al. proposed an image analysis model for small-target fault diagnosis of wind turbine blades based on YOLOV5. In the study, the multi-scale feature pyramid is used to achieve multi-layer feature fusion, which is much conducive to the detection of small targets [27]. Wu et al. used the multiscale anchor mechanism to improve the small target detection ability of YOLOV5 [28]. Zhang et al. proposed a model SOD-YOLO for wind turbine blade surface fault diagnosisbased YOLOV5, and a much small feature fusion layer is added to the YOLOV5, and a convolutional block attention module is added to the YOLOV5, thus the feature loss of small target defects can be reduced [29]. Gou et al. by using the encoding and decoding modules from the transformer added to the YOLOV5m and the detector was enhanced to dynamically adjust to objects of different scales by a multi-scale feature fusion structure for defects of different sizes [30]. Kim et al. proposed a method to improve the detection performance of small targets in aerial images by modifying YOLOV5 in order to obtain robust detection ability [31].

In image detection, the local important information is essential for detection accuracy. The self-attention was proposed in 2017 to get key feature information about an image [32]. Self-attention has been used widely all around the world [33–37]. Dong et al. proposed an improved YOLOV5 for vehicle detection. A convolutional block attention module is introduced to the YOLOV5, thus the detection accuracy is improved [33]. Zhang et al. proposed a squeeze-and-excitation attention mechanism module embedded into the YOLOV5, thus the detection accuracy is improved [34]. Zhang et al. introduced coordinate attention to YOLOV5, and the feature extraction ability can be strengthened, and the detection performance of leaf disease is improved [35]. Zhao et al. introduced the squeeze-and-excitation module of the attention mechanism to YOLOV5, and the depth of the feature map of the previous layer is a weighted average, and its accuracy is improved [36]. Ying et al. proposed an improved YOLOV5 based on an efficient channel attention mechanism, and the high-pressure steel wire braided hose is detected [37].

The small target detection accuracy and network lightweight are studied deeply in this study, and the improved YOLOV5 is proposed. Firstly, in order to reduce the loss of feature information, a CA-C3 structure and a new connection are proposed, and the feature information can be introduced from the backbone network to the neck one. Secondly, in order to reduce the network size, YOLOV5 is improved by the lightweight network ShuffleNetV2. Finally, to further improve the accuracy, the cosine annealing with warm restarts algorithm is used to optimize the above network. The dataset of NEU-DET is testified, and the results show that the improved YOLOV5 can detect steel surface faults with high efficiency and accuracy.

The rest of the study is organized as follows. In Section 2, the preliminaries of YOLOV5, coordinate attention, ShuffleNetV2, and the cosine annealing with warm restarts algorithm are presented. The architecture of improved YOLOV5 is constructed in Section 3. Section 4 displays the flowchart of improved YOLOV5. Section 5 shows the dataset of NEU-DET and the hardware and software of the computer. Section 6 details the experimental results of small target detection. In Section 7, the conclusions of this study are summarized.

2. Methodology

2.1. YOLOV5

Traditional image segmentation techniques have been relatively mature. However, these techniques require feature extraction for each fault. This results in high labor costs and low efficiency.

There are many target detection algorithms [38,39], and deep learning has been developed and studied by many researchers from 2012 to the present. Therefore, deep learning is a current research hotspot. YOLO is a kind of deep learning network [40], and it was proposed in 2016 [41], and it is highly efficient and has a good generalization ability for detecting small targets.

The YOLO series is a single-stage target detection method, which consists of the input, the backbone network, the neck, and the output. Mosaic data enhancement and adaptive anchor frame are approved at the input end in YOLOV5, thus it is good at highlighting small target features. In the backbone network, C3 is used to avoid gradient disappearance. A hybrid FPN + PAN is used to enhance the fusion of the shallow feature information. The binary cross-entropy loss is used for the classification loss in the training, and it is defined as follows:

$$Loss(obj) = GIOU_{loss} + \sum_{i=0}^{S \times S} \sum_{j=0}^{B} 1_{ij}^{obj} [C_i \log(C_i) + (1 - C_i) \log(1 - C_i)] \\ - \sum_{i=0}^{S \times S} \sum_{j=0}^{B} 1_{ij}^{noobj} [C_i \log(C_i) + (1 - C_i) \log(1 - C_i)] \\ + \sum_{i=0}^{S \times S} \sum_{i=0}^{B} 1_{ij}^{obj} \sum_{c \in classes} [p_i(c) \log(p_i(c)) + (1 - p_i(c)) \log(1 - p_i(c))]$$
(1)

where $S \times S$ represents the cell number, B denotes the number of predicted bounding boxes, 1_{ij}^{obj} and 1_{ij}^{noobj} represent the objects in *i* cell and *j* bounding box. C_i represents confidence values for the predicted *i* grid, and $p_i(c)$ represents probability values for predicted and actual targets in *i* grid. $GIOU_{loss}$ is the loss function of the bounding box. The architecture of YOLOV5 is shown in Figure 1.



Figure 1. The architecture of YOLOV5.

2.2. Coordinate Attention Mechanism

Coordinate attention (CA) is a network architecture proposed by Hou et al. [42]. Because the location information is embedded into channels, the cross-channel feature information can be easily extracted, thus the target regions can be accurately located and identified. The architecture of the coordinate attention is shown in Figure 2.



Figure 2. The architecture of coordinate attention.

2.3. ShuffleNetV2 Architecture

The network ShuffleNetV2 is a kind of lightweight method, and it was proposed by Ma et al. [43] in 2018. ShuffleNetV2 is constructed based on basic and down-sampling modules, followed by a convolutional layer and an average pooling layer. When extracting the feature information, group convolution and channel shuffle are used to compress the model and greatly reduce the computation cost. Its architecture is shown in Figure 3.



Figure 3. The architecture of ShuffleNetV2: (a) Basic module; (b) Down-sampling module.

2.4. Cosine Annealing with Warm Restarts Algorithm

Deep learning models are often optimized by gradient descent algorithms. Setting too small a learning rate will lead to slow convergence of objective function, and setting one too large may lead to a local optimal solution. Cosine annealing with a warm restarts algorithm can realize periodic restarts in the decreasing process of the learning rate, so as to make the objective function jump out of the local optimal solution. The periodic restart method increases the learning rate suddenly and jumps out of the local optimal solution. Excellent, approaching the global optimal solution. The learning rate for each batch is written as follows:

$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})(1 + \cos\left(\frac{T_{cur}}{T_i}\pi\right)$$
(2)

where η_t denotes the current learning rate, η_{max} and η_{min} signify the maximum and minimum learning rates, respectively. T_{cur} is the epoch number executed since the last restart, and T_i signifies the epoch number that needs to be trained after *i*th restart. The change curve of the learning rate under the action of the algorithm is shown in Figure 4.



Figure 4. Learning rate curve of the cosine annealing with warm restarts.

3. Improved YOLOV5

3.1. CA-C3 Module

The shallow architecture of YOLOV5 leads to low accuracy and efficiency, thus the coordinate attention is introduced into the C3 module, and a new module CA-C3 can be obtained. The new module CA-C3 is shown in Figure 5. One part consists of a convolution layer and a number of Bottleneck modules, which reduces the parameter amount of Bottleneck modules. The other part consists of a convolutional layer and coordinates attention to enhance the feature extraction ability. Finally, concat operation is performed on the two parts to obtain the CA-C3 module. The main advantage of this module is to reduce the loss of feature information.



Figure 5. The architecture of CA-C3 module.

3.2. Improved YOLOV5

3.2.1. Improvement Based on the CA-C3 Module, New Connection and ShuffleNetV2

The size of a deep neural network has an impact on accuracy and efficiency, and its parameter amount and calculation cost of the network increase greatly with the size of the incensement. The recall rate may be low when YOLOV5 detects small targets, resulting in low accuracy. The rich feature information is contained in the backbone network, but the information is not fully used, which leads to low accuracy.

To address the problem of the small targets, YOLOV5 is improved by coordinate attention, and detection ability can be improved. To address the problem of low precision, a new connection is constructed in YOLOV5, and the detection ability can be improved. To address the problem of the lightweight, YOLOV5 is improved by the lightweight method ShuffleNetV2, and the network size is reduced.

Aiming at the above problems, YOLOV5 is improved by the CA-C3 module, new connection, and ShuffleNetV2. C3 module is replaced by the CA-C3 module to extract rich feature information, and Conv of the backbone module is replaced by ShuffleNetV2 to

construct a lightweight network, and Line I and line II are used to connect the backbone and neck networks to extract the rich feature information, thus improved YOLOV5 is proposed. Its architecture is displayed in Figure 6.



Figure 6. The architecture of improved YOLOV5.

The detail architecture information of improved YOLOV5 is listed in Table 1.

| | Layer | Size | Stride | Channel | Output Feature Size | Note |
|----|-------------------|--------------|--------|---------|----------------------------|-------------|
| | Input layer | \ | \ | \ | $224 \times 224 \times 3$ | |
| 0 | Convolution layer | 1×1 | 1 | 32 | 224 	imes 224 	imes 32 | |
| 1 | ShuffleNetV2 | \ | \ | 64 | 112 	imes 112 	imes 64 | |
| 2 | $CA-C3 \times 3$ | Ň | Ň | 64 | 112 	imes 112 	imes 64 | |
| 3 | ShuffleNetV2 | Ň | Ň | 128 | 56 	imes 56 	imes 128 | |
| 4 | $CA-C3 \times 6$ | Ň | Ň | 128 | 56 	imes 56 	imes 128 | |
| 5 | ShuffleNetV2 | Ň | Ň | 256 | 28	imes28	imes256 | |
| 6 | $CA-C3 \times 9$ | Ň | Ň | 256 | 28	imes28	imes256 | |
| 7 | ShuffleNetV2 | Ň | Ň | 512 | 14 	imes 14 	imes 512 | |
| 8 | $CA-C3 \times 3$ | Ň | Ň | 512 | 14 	imes 14 	imes 512 | |
| 9 | SPPF | Ň | Ň | \ | 14 	imes 14 	imes 512 | |
| 10 | Convolution layer | 1×1 | 1 | 256 | 14	imes14	imes256 | |
| 11 | Upsample | \ | \ | \ | 28	imes28	imes256 | |
| 12 | Concat | Ň | Ň | Ň | \ | Layer 6 |
| 13 | CA-C3 | Ň | Ň | 256 | 28 	imes 28 	imes 256 | - |
| 14 | Convolution layer | 1×1 | 1 | 128 | 28	imes28	imes128 | |
| 15 | Upsample | \ | \ | \ | 56 	imes 56 	imes 128 | |
| 16 | Concat | Ň | Ń | Ň | \ | Layer 4 |
| 17 | CA-C3 | Ň | Ň | 128 | $56 \times 56 \times 128$ | Output1 |
| 18 | Convolution layer | 3×3 | 2 | 128 | 28	imes28	imes128 | 1 |
| 19 | Concat | \ | \ | \ | \ | Layer 14, 6 |
| 20 | CA-C3 | Ň | Ń | 256 | $28 \times 28 \times 256$ | Output2 |
| 21 | Convolution layer | 3×3 | 2 | 256 | 14 	imes 14 	imes 256 | 1 |
| 22 | Concat | \ | \ | \ | \ | Layer10,8 |
| 23 | CA-C3 | Ň | Ň | 512 | $14 \times 14 \times 512$ | Output3 |

 Table 1. The detail architecture information of improved YOLOV5.

In this study, its initialization parameters are listed is Table 2.

| Table 2. 🕻 | The initialization | n parameters | of improv | ved YOLOV5. |
|------------|--------------------|--------------|-----------|-------------|
| | | | | |

| Parameters | Value | Note | | |
|-----------------------|------------------|---------|--|--|
| Size of input images | 224×224 | NEU-DET | | |
| Batch size | 8 | | | |
| Weight | yolov5s.pt | | | |
| Learning rate | 10 ⁻³ | | | |
| The iterations number | 300 | | | |

3.2.2. Optimization Based on Cosine Annealing with Warm Restarts Algorithm

To address the problem of low precision, YOLOV5 is improved by the cosine annealing with warm restarts algorithm for improving the precision.

In order to obtain good performance, the cosine annealing with warm restarts algorithm is applied to optimize the above-improved YOLOV5, thus YOLOV5 can be further improved. The optimization flowchart is shown in Figure 7.



Figure 7. The optimization flowchart of YOLOV5 based on the cosine annealing with warm restarts algorithm.

Firstly, the improved YOLOV5 is trained. Secondly, during the training process, the improved YOLOV5 is trained through cosine annealing with warm restarts algorithm. Thirdly, it determines whether the learning rate has met the restart condition. If the condition is not met, the learning rate is updated. If the restart condition of the learning rate is met, the learning rate is restarted. Lastly, it determines whether it has met the training cycle. If the condition is met, this training is terminated, and thus is an output of the improved YOLOV5 after optimization.

4. The Flowchart of Improved YOLOV5

A new network called improved YOLOV5 is proposed in this study. Firstly, in order to reduce the feature information loss of the network, a new CA-C3 structure is constructed. Secondly, to further reduce the feature information loss, a new connection is proposed. Thirdly, in order to reduce the network size and improve the accuracy, the network is

improved by the lightweight method ShuffleNetV2. Lastly, to further improve the accuracy of the network, the cosine annealing with warm restarts algorithm is used to optimize the above network. The flow chart is shown in Figure 8.



Figure 8. Flow chart of the proposed method.

5. The Dataset of NEU-DET

The dataset of NEU-DET [44] is used to verify the effectiveness and superiority of improved YOLOV5. In the dataset, each defect type has 300 images, a total of 1800 images. There are six defect types: crazing, inclusion, patches, pitted_surface, roll-in_scale, and scratches. The crazing defect is mainly featured by irregular lines. The inclusion defect is of different shapes. The patches defect is of irregular shapes. The pitted-surface defect is generally characterized by local pits. The rolled-in scale defect is manifested as convexes in many places. The scratches defect is mainly shown as long strips. These defects have the characteristics of complex shapes and small sizes. Therefore, their detection difficulty is increased. Figure 9 shows the images of defect samples.



Figure 9. The fault samples of 6 defect types.

This study is implemented on the Windows 10 operating system using the PyTorch framework in Pycharm. The device hardware is Intel(R) Core(TM) i5-8300H CPU @ 2.30 GHz and the GPU model is GTX 1050Ti. The software environments are CUDA11.4 and Python3.8.

6. Case Study

The ablation study is carried out to demonstrate the validity and superiority of improved YOLOV5, and the comparative results based on the three factors of precision, recall, and mAP are listed in Table 3.

| Methods | Cosine Annealing with Warm Restarts Algorithm | CA-C3 Module | Connections of Line I and Line II | ShuffleNetV2 | Precision | Recall | mAP |
|--------------------|--|-----------------|--------------------------------------|--------------|-----------|--------|-------|
| YOLOV5 | × | × | × | × | 0.626 | 0.818 | 0.758 |
| Case 1 | \checkmark | × | × | × | 0.723 | 0.796 | 0.791 |
| Case 2 | \checkmark | \checkmark | Х | × | 0.730 | 0.799 | 0.806 |
| Case 3 | \checkmark | \checkmark | \checkmark | × | 0.787 | 0.819 | 0.840 |
| Improved YOLOV5 | \checkmark | \checkmark | \checkmark | \checkmark | 0.822 | 0.817 | 0.859 |

Table 3. Comparative results of ablation study.

It can be known from Table 3 that Precision, Recall, and mAP show an overall upward trend with further improvements, and the analysis is detailed as follows:

Based on the comparison between Case 1 and YOLOV5, we can know that the three factors increase with the application of the cosine annealing with the warm restarts algorithm, thus the algorithm can help the objective function jump out of the local optimal solution.

With the construction of the new CA-C3 module in Case 2, the three factors show an increasing trend compared with case 1, thus the rich feature extraction can be extracted by the new CA-C3 module.

The new connections of Line I and Line II of Case 3 can show much more detailed feature information fused and preserved, thus the three factors are all bigger than those of Case 2.

Based on the comparison between improved YOLOV5 and Case 3, it shows that the parameter amount can be reduced and the accuracy can be improved with the application of the lightweight method ShuffleNetV2, thus the three factors become big.

In order to show effectiveness and superiority, they are compared in terms of the parameter amount and computation cost, and the results are listed in Figure 10.



Figure 10. The parameter amount and the model size based on the ablation study.

We can see from Figure 10 that the parameter amount and model size is not influenced by the cosine annealing with warm restarts algorithm by comparing YOLOV5 and Case 1. Based on Case 1 and Case 2, the parameter amount can be reduced by the CA-C3 module. In Case 3, since Line I and Line II are connected to other modules, the parameter amount and model size are all increased compared with Case 2. In order to eliminate the adverse effects of Case 3 on the parameter amount and model size, the introduction of the lightweight method ShuffleNetV2 has greatly reduced the parameter amount and model size, and the parameter amount and model size of improved YOLOV5 is the least, and they are even 23.48% and 22.3% lower than those of YOLOV5 respectively.

In order to prove the effectiveness of the improved YOLOV5, the PR curve is presented to verify. In the PR curve, the horizontal axis represents recall, and the vertical axis represents precision. The closer the curve gets to the top right, the better the fault results are.

As shown in Figure 11, the AP of crazing, inclusion, patches, pitted_surface, rolledin_scale, and scratches are improved by 95.3%, 7.9%, 0.2%, 8.4%, 2.6%, and 9.3% respectively. Among them, due to the crazing defect feature, there are a lot of faults in the small targets. Therefore, CA-C3 models are added in this paper, thus feature loss is reduced. In addition, the improved YOLOV5 is closer to the top right compared with YOLOV5, which indicates that the detection effects are better compared with YOLOV5.

To further verify the effectiveness and superiority of improved YOLOV5, it is compared with other models single shot multi-box detector (SSD), faster RCNN, YOLOV3 and the methods proposed in [45,46]. Their mAP are displayed in Figure 12, and their parameter amount and model size are shown in Figure 13.

It can be known from Figures 12 and 13 that mAP, parameter amount, and model size are 85.9%, 20.51×10^6 , and 10.56×10^4 , respectively. They are 11.13%, 52.30%, 116.91%, 13.32%, 13.32%, and 3.87% higher than the SSD, Faster RCNN, YOLOV3, and YOLOV5, methods proposed by [45,46].

The conclusion is that the cosine annealing with warm restarts algorithm, attention mechanism, and connections can improve the accuracy. Furthermore, ShuffleNetV2 can reduce the size of the network and improve precision.



Figure 11. PR Curve: (a) YOLOV5; (b) Improved YOLOV5.



Figure 12. mAP of improved YOLOV5 and other methods [45,46].



Figure 13. The parameter amount and model size of improved YOLOV5 and other methods.

In order to further demonstrate the advantages of improved YOLOV5 in detecting small targets, the detection comparison results are shown in Figure 14.

Small targets detected by improved YOLOV5 in Figure 14b are marked with arrows, but the targets are not detected by YOLOV5 in Figure 14a. As can be seen from Figure 14a, although most of the defect targets can be detected by YOLOV5, there are still some small defect targets that are difficult to detect, because the feature loss of convolution operation results in low accuracy. However, all of the above small defect targets can be detected by the improved YOLOV5 with high accuracy.

improve 0.53
improve 0.57

Figure 14. Diagnosis results: (a) YOLOV5; (b) Improved YOLOV5.

Therefore, improved YOLOV5 performs better than YOLOV5 in detecting small targets with high accuracy and efficiency.

7. Conclusions

Steel is an important raw material of fluid components, and it is meaningful to diagnose steel faults. The complex shape and small size of steel surface faults lead to low accuracy, and the large size of the network leads to poor real-time performance. Aiming at the problems, an improved YOLOV5 is proposed. The dataset of NEU-DET is verified and testified, and the conclusions of this study are shown as follows.

(1) The new CA-C3 module is constructed based on the coordinate attention, and improved YOLOV5 can extract much feature information, thus the accuracy can be improved.

(2) A new connection is constructed, and the feature information can be directly input into the neck from the backbone, and much information can be preserved by improved YOLOV5, thus the accuracy can be improved.

(3) The lightweight method ShuffleNetV2 is used to improve YOLOV5, and the size of the improved YOLOV5 is reduced, thus the computation cost can be reduced.

(4) YOLOV5 is improved by the cosine annealing with the warm restarts algorithm, thus the accuracy of improved YOLOV5 can be improved.

Steel surface detection is mainly researched in this paper. However, it is required to construct lightweight architecture for rapid detection. In the future, we will continue to op-

timize YOLOV5, and adopt quantization recognition to optimize the method. Furthermore, many categories of steel surface fault will be added, thus the requirements of the industrial fields much met.

Author Contributions: Conceptualization, Y.X. and W.L.; methodology, Y.X. and A.Z.; software, A.Z. and Y.X.; formal analysis, W.L.; writing—original draft preparation, W.L., Y.X., C.L. and Z.Z. (Zhen Zhang); writing—review and editing, W.L., Z.Z. (Zhi Zheng) and A.Z., supervision, W.L. and A.Z.; project administration, W.L.; funding acquisition, W.L., Z.Z. (Zhi Zheng) and X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by S&T Program of Hebei (22282203Z), Hebei Natural Science Foundation (E2022209086, E2019209314), National Natural Science Foundation of China (52004096), Hebei Provincial Postdoctoral Science Foundation (B2020003033).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: https://github.com/magenasi/NEU-DET (accessed on 30 September 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Zhu, Y.; Li, G.; Tang, S.; Wang, R.; Su, H.; Wang, C. Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN. *Appl. Acoust.* **2022**, *192*, 108718. [CrossRef]
- Tang, S.; Zhu, Y.; Yuan, S. Intelligent Fault Identification of Hydraulic Pump Using Deep Adaptive Normalized CNN and Synchrosqueezed Wavelet Transform. *Reliab. Eng. Syst. Saf.* 2022, 224, 108560. [CrossRef]
- Tang, S.; Zhu, Y.; Yuan, S. A Novel Adaptive Convolutional Neural Network for Fault Diagnosis of Hydraulic Piston Pump with Acoustic Images. *Adv. Eng. Inform.* 2022, 52, 101554. [CrossRef]
- 4. Iravani, M.; Toghraie, D. Design a high-pressure test system to investigate the performance characteristics of ball valves in a compressible choked flow. *Measurement* **2020**, *151*, 107200. [CrossRef]
- 5. Darvishyadegari, M.; Hassanzadeh, R. Heat and fluid flow around two co-rotating cylinders in tandem arrangement. *Int. J. Therm. Sci.* **2019**, *135*, 206–220. [CrossRef]
- Cao, L.; Tu, C.; Hu, P.; Liu, S. Influence of solid particle erosion (SPE) on safety and economy of steam turbines. *Appl. Therm. Eng.* 2019, 150, 552–563. [CrossRef]
- Li, Z.; Tian, X.; Liu, X.; Liu, Y.; Shi, X. A two-stage industrial defect detection framework based on improved-YOLOV5 and optimized-inception-ResnetV2 models. *Appl. Sci.* 2022, 12, 834. [CrossRef]
- 8. Song, G.; Song, K.; Yan, Y. EDRNet: Encoder-decoder residual network for salient object detection of strip steel surface defects. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9709–9719. [CrossRef]
- 9. Tian, R.; Jia, M. DCC-CenterNet: A rapid detection method for steel surface defects. *Measurement* 2022, 187, 110211. [CrossRef]
- 10. Gao, L.; Zhang, J.; Yang, C.; Zhou, Y. Cas-VSwin transformer: A variant swin transformer for surface-defect detection. *Comput. Ind.* **2022**, *140*, 103689. [CrossRef]
- Yin, J.; Guo, L.; Jiang, W.; Hong, S.; Yang, J. ShuffleNet-inspired lightweight neural network design for automatic modulation classification methods in ubiquitous IoT cyber–physical systems. *Comput. Commun.* 2021, 176, 249–257. [CrossRef]
- 12. Ran, H.; Wen, S.; Wang, S.; Cao, Y.; Zhou, P.; Huang, T. Memristor-based edge computing of ShuffleNetV2 for image classification. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* 2021, 40, 1701–1710. [CrossRef]
- Kulkarni, U.; Meena, S.M.; Gurlahosur, S.V.; Bhogar, G. Quantization friendly MobileNet (Qf-MobileNet) architecture for vision based applications on embedded platforms. *Neural Netw.* 2021, *136*, 28–39. [CrossRef]
- 14. Zhao, L.; Wang, L. A new lightweight network based on MobileNetV3. KSII Trans. Internet Inf. Syst. (TIIS) 2022, 16, 1–15.
- 15. Panigrahi, S.; Raju, U.S.N. MS-ML-SNYOLOV3: A robust lightweight modification of SqueezeNet based YOLOV3 for pedestrian detection. *Optik* **2022**, *260*, 169061. [CrossRef]
- 16. Guan, S.; Chang, J.; Shi, H.; Xiao, X.; Li, Z.; Wang, X.; Wang, X. Strip steel defect classification using the improved GAN and EfficientNet. *Appl. Artif. Intell.* **2021**, *35*, 1887–1904. [CrossRef]
- 17. Nagy, A.M.; Czúni, L. Classification and fast few-shot learning of steel surface defects with randomized network. *Appl. Sci.* 2022, 12, 3967. [CrossRef]
- Liu, W.; Zhang, J.; Su, Z.; Zhou, Z.; Liu, L. Binary neural network for automated visual surface defect detection. *Sensors* 2021, 21, 6868. [CrossRef]
- 19. Lin, Z.; Ye, H.; Zhan, B.; Huang, X. An efficient network for surface defect detection. Appl. Sci. 2020, 10, 6085. [CrossRef]
- Fu, G.; Sun, P.; Zhu, W.; Yang, J.; Cao, Y.; Yang, M.Y.; Cao, Y. A deep-learning-based approach for fast and robust steel surface defects classification. *Opt. Laser Eng.* 2019, 121, 397–405. [CrossRef]

- Li, Y.; Huang, H.; Xie, Q.; Yao, L.; Chen, Q. Research on a surface defect detection algorithm based on MobileNet-SSD. *Appl. Sci.* 2018, *8*, 1678. [CrossRef]
- Liu, H.B.; Zhong, H.; Karpowicz, N.; Chen, Y.; Zhang, X.C. Terahertz spectroscopy and imaging for defense and security applications. *Proc. IEEE* 2007, 95, 1514–1527. [CrossRef]
- Xu, X.; Zhao, J.; Li, Y.; Gao, H.; Wang, X. BANet: A balanced atrous net improved from SSD for autonomous driving in smart transportation. *IEEE Sens. J.* 2020, 21, 25018–25026. [CrossRef]
- Yan, J.; Wang, Z. YOLOV3 + VGG16-based automatic operations monitoring and analysis in a manufacturing workshop under Industry 4.0. J. Manuf. Syst. 2022, 63, 134–142. [CrossRef]
- Xing, J.; Jia, M. A convolutional neural network-based method for workpiece surface defect detection. *Measurement* 2021, 176, 109185. [CrossRef]
- Li, M.; Wang, H.; Wan, Z. Surface defect detection of steel strips based on improved YOLOv4. Comput. Electr. Eng. 2022, 102, 108208. [CrossRef]
- Qiu, Z.; Wang, S.; Zeng, Z.; Yu, D. Automatic visual defects inspection of wind turbine blades via YOLO-based small object detection approach. *J. Electron. Imaging* 2019, 28, 043023. [CrossRef]
- 28. Wu, W.; Liu, H.; Li, L.; Long, Y.; Wang, X.; Wang, Z.; Chang, Y. Application of local fully convolutional neural network combined with YOLOV5 algorithm in small target detection of remote sensing image. *PLoS ONE* **2021**, *16*, e259283. [CrossRef]
- Zhang, R.; Wen, C. SOD-YOLO: A small target defect detection algorithm for wind turbine blades based on improved YOLOV5. *Adv. Theory Simul.* 2022, *5*, 2100631. [CrossRef]
- Guo, Z.; Wang, C.; Yang, G.; Huang, Z.; Li, G. MSFT-YOLO: Improved YOLOV5 based on transformer for detecting defects of steel surface. Sensors 2022, 22, 3467. [CrossRef]
- Kim, M.; Jeong, J.; Kim, S. ECAP-YOLO: Efficient channel attention pyramid YOLO for small object detection in aerial image. *Remote Sens.* 2021, 13, 4851. [CrossRef]
- 32. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* (*NIPS*) **2017**, *30*, 5998–6008.
- Dong, X.; Yan, S.; Duan, C. A lightweight vehicles detection network model based on YOLOV5. *Eng. Appl. Artif. Intell.* 2022, 113, 104914. [CrossRef]
- 34. Zhang, Z.D.; Tan, M.L.; Lan, Z.C.; Liu, H.C.; Pei, L.; Yu, W.X. CDNet: A real-time and robust crosswalk detection network on Jetson nano based on YOLOv5. *Neural Comput. Appl.* **2022**, *34*, 10719–10730. [CrossRef]
- 35. Zhang, Z.; Qiao, Y.; Guo, Y.; He, D. Deep learning based automatic grape downy mildew detection. *Front. Plant Sci.* 2022, 13, 872107. [CrossRef]
- 36. Zhao, Z.; Yang, X.; Zhou, Y.; Sun, Q.; Ge, Z.; Liu, D. Real-time detection of particleboard surface defects based on improved YOLOV5 target detection. *Sci. Rep.* **2021**, *11*, 21777. [CrossRef]
- Ying, Z.; Lin, Z.; Wu, Z.; Liang, K.; Hu, X. A modified-YOLOV5s model for detection of wire braided hose defects. *Measurement* 2022, 190, 110683. [CrossRef]
- Zaidi, S.S.A.; Ansari, M.S.; Aslam, A.; Kanwal, N.; Asghar, M.; Lee, B. A survey of modern deep learning based object detection models. *Digit. Signal Process.* 2022, 126, 103514. [CrossRef]
- Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 1137–1149. [CrossRef]
- Diwan, T.; Anirudh, G.; Tembhurne, J.V. Object detection using YOLO: Challenges, architectural successors, datasets and applications. *Multimed. Tools Appl.* 2022, 1–33. [CrossRef]
- 41. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
- Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for Efficient mobile network design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.
- Ma, N.; Zhang, X.; Zheng, H.T.; Sun, J. ShuffleNetV2: Practical guidelines for Efficient CNN architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 116–131.
- 44. Song, K.; Yan, Y. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects. *Appl. Surf. Sci.* 2013, 285, 858–864. [CrossRef]
- Zhao, W.; Chen, F.; Huang, H.; Li, D.; Cheng, W. A new steel defect detection algorithm based on deep learning. *Comput. Intell. Neurosci.* 2021, 2021, 5592878. [CrossRef] [PubMed]
- Zhang, J.; Kang, X.; Ni, H.; Ren, F. Surface defect detection of steel strips based on classification priority YOLOV3-dense network. *Ironmak. Steelmak.* 2021, 48, 547–558. [CrossRef]