



Novel intelligent defects detection of boiler water walls in thermal power plants based on OFM_SSD[☆]

Yongming Han^{a,b}, Lei Wang^{a,b}, Jintao Liu^c, Liang Yuan^d, Hongxu Liu^{a,b,*}, Bo Ma^{e,*}, Zhiqiang Geng^{a,*}

^a College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

^b State Key Laboratory of Mechanical Transmission for Advanced Equipment, Chongqing University, Chongqing 400044, China

^c College of Systems Engineering, National University of Defense Technology, Changsha 410073, China

^d Shanghai Jiao Tong University, 200240, China

^e College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029, China

ARTICLE INFO

Keywords:

Optical flow method
Single shot multibox detector
Defects detection
Boiler water wall
Thermal power plants

ABSTRACT

The boiler is a critical component of conventional thermal power systems, where surface flaws in boiler water walls can significantly compromise safety and availability, potentially leading to substantial loss of life and property. Traditional detection methods, whether manual or based on machine learning, often prove costly, inefficient and time-consuming, failing to meet the stringent requirements for water wall inspection. Therefore, a novel surface defect detection model integrating an improved single shot multibox detector (SSD) with the optical flow method (OFM) (OFM_SSD) is proposed. The OFM enhances data sample diversity by augmenting the dataset derived from thermal power plants, while the incorporation of deconvolution techniques improves the model receptive field, thereby enhancing its ability to detect and classify small defects. Comprehensive experiments demonstrate that the OFM_SSD outperforms several existing models including the SSD model based on traditional expanded datasets (T_SSD), you only look once (YOLO), ordinary SSD, Regions with the CNN(R_CNN), and Deconvolution-only SSD (DSSD) in terms of accuracy in defect localization and classification. This advancement of the OFM_SSD not only reduces operational costs but also enhances detection capabilities, ultimately contributing to safer and more efficient operations within thermal power plants.

1. Introduction

The development of the electricity industry is inextricably linked to the enhancement of people's quality of life. Globally, thermal power generation is still the primary source of electricity in the world [1]. Economic development promotes the rapid consumption of electricity. Countries around the globe are vigorously embracing clean energy, such as wind and solar power, as the rapid progress of renewable energy continues. However, due to the inherent unpredictability of these clean energy sources and the limitations of energy storage technology, the traditional coal-fired power generation is indispensable to meet the demands of the electricity grid and ensure its reliable operation [2]. The boiler system is a significant part of the composition of the operation of the boiler system, which directly affects the production of the power industry [3]. Because large industrial boilers work in high temperature

and high-pressure environments [4], the power plant has a high probability of failure, which is extremely destructive after it occurs. In 2016, a power plant in Hubei, China, exploded due to a boiler problem in the power plant, causing serious economic losses and casualties [5]. A lot of power plant accidents are often caused by defects on the surface of the water wall [6]. Real-time monitoring of the water wall surface condition of industrial boilers, finding faults and solving these faults is particularly important for protecting people's lives and property. At present, there are few research contents on the surface defect detection of boiler water wall. The problems of high cost and low detection capability exist respectively based on hardware equipment detection and computer vision intelligent recognition [7]. The sample data of boiler water wall surface defects are small in number and difficult to collect, which further deepens the difficulty of detection. In order to solve the research problem in the actual industrial field, a novel intelligent defect detection

[☆] This paper was recommended for publication by Prof Guangtao Zhai.

* Corresponding authors at: College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China (H. Liu).

E-mail addresses: 2023700063@buct.edu.cn (H. Liu), mabo@mail.buct.edu.cn (B. Ma), gengzhiqiang@mail.buct.edu.cn (Z. Geng).

model based on an improved single shot multibox detector (SSD) integrating the optical flow method (OFM) (OFM SSD) is proposed.

The main contributions of this paper are summarized as follows:

- (1) The pixel correspondence between adjacent frames in video sequences is utilized to compute motion information, facilitating the generation of intermediate frames, which effectively augments the defect dataset.
- (2) The expanded dataset is employed to train and test an improved SSD method that incorporates deconvolution techniques.
- (3) The OFM SSD method is implemented for intelligent detection of boiler water wall surface defects in thermal power plants. Comparative analysis reveals that the accuracy of the OFM SSD exceeds that of traditional expanded dataset SSD (T.SSD), you only look once (YOLO), ordinary SSD, Regions with the CNN (R_CNN), and Deconvolution-only SSD(DSSD) by margins of 1.94 %, 4.56 %, 3.72 %, 3.38 %, and 2.93 %, respectively.

Furthermore, the proposed model demonstrates effective detection of water wall surface defects, thereby minimizing potential damages to facilities and reducing the costs associated with manual fault detection.

2. Related works

The boiler and steam turbine are essential for maintaining electricity in industrial production processes, as the failure of any individual component can halt the entire production line. Within the operational cost of a thermal power plant, boiler operation represents a significant expense. Meanwhile, the water wall plays a crucial role in the boiler system by absorbing radiant heat and protecting the furnace walls [8]. Most boiler shutdowns result from defects such as corrosion, cracks and crevices in the water wall tubes, leading to substantial losses for thermal power plants [9]. To mitigate these industrial costs, early warning systems for fault diagnosis of water wall surface defects can be categorized into two approaches: non-destructive testing methods utilizing hardware and intelligent defect identification based on computer vision [7].

In the first case, the main research methods for defecting the surface of the water wall include acoustic detection methods, magnetic flux leakage-based technology, and eddy current testing [10]. Lu et al. [11] proposed a pyramid auxiliary supervised U-Net model with Dual-Attention mechanism. Han et al. [12] proposed a novel multiscale variational autoencoder (MSVAE) based Regressor (REG) (MSVAE-REG) to predict production and save energy. Gao et al. [13] used the eddy current technology to detect the defect at the far end of the pipeline. Zhang et al. [14] proposed the DSSO-YOLO model, which introduces the coordinate attention mechanism into the C3 module of YOLOv5s to increase the accuracy of detection. However, these above methods have high cost and poor operability [15]. Meanwhile, the existing single method can detect the leakage after it occurs, which often brings serious damage to the production, and cannot give early warning in time along with certain operational limitations.

As hardware equipment continues to evolve and technology progresses, an increasing number of researchers are employing deep learning methods for intelligent defect detection in industrial equipment. In contrast to traditional detection methods that often suffer from limitations, the defect detection approach based on deep learning transforms the process of feature extraction for defects, resulting in a significant enhancement in defect detection performance [16]. Compared with traditional defect detection methods, the DL has achieved a better result [17], which can automatically extract features of different levels in the image with good robustness. Yan et al. [18] used variational autoencoder-based conditional Wasserstein GAN with gradient penalty (CWGAN-GP-VAE) method to diagnose various faults for chillers, which provided an effective means for the defect detection and achieved good results. Tan et al. [19] used a fast R_CNN method in the core components of the high-speed railway to detect high-speed

railway defects. However, the R_CNN requires a lot of repeated calculations to dispose of the image, so the speed of the R_CNN is slow, which cannot meet the requirement for the detection of boiler surface faults with high real-time and fast response speed. As a target detection algorithm, Han et al. [20] applied the YOLO to realize the surface defect of concrete building. Han et al. [21] used a novel pipeline attention integrating the long short-term memory (LSTM) to detect the leakage. But the accuracy and recall of the YOLO were not high [22]. The SSD algorithm based on the one-stage improved the speed of detecting faults on the boiler surface by giving up the full connection layer [23]. Sun et al. [24] applied the SSD network structure to motor image classification in EEG. Gong et al. [25] applied the SSD to track and detect small targets. However, the feature map representation ability of the shallow layer of the SSD algorithm was not strong for obtaining good semantic information [26]. Compared with general target detection, because of the small size of small targets, the shallow characteristic map has not undergone a lot of convolution pooling which has less semantic information [27]. Deconvolution is introduced to enrich the semantic information of features which used to enhance the detection accuracy and the ability of small object detection. In addition, there are fewer defect data sets collected in industrial production processes, so further augmentation data can be obtained to achieve good experimental results. The OFM first proposed by Arnheim [28] was a classic video frame interpolation algorithm, which converted the real-world motion field into the optical flow field in the two-dimensional image and filled in the missing motion frame to effectively expand the data set. In order to achieve effective defect detection on the small sample data set collected on the surface of the water wall in real, a novel OFM_SSD based intelligent detection method is proposed.

3. The proposed methodology

The intelligent fault diagnosis method utilizing computer vision facilitates automated detection, enhances the efficiency, and minimizes human error, offering significant advantages over traditional approaches [29]. However, current defect detection results based on limited sample data remain suboptimal. Therefore, the OFM_SSD method incorporates deconvolution operations into the SSD framework to improve the detection of small target defects.

3.1. The OFM

The optical flow comes from the relative motion of two objects. It is used to describe the instantaneous velocity of the moving object in space on the observation imaging plane. The core of the OFM is to calculate the information between two frames and use the information to synthesize the intermediate data frame according to the assumption of linear movement of pixels within the frame. The OFM needs to meet two basic assumptions. One is that the state of moving objects cannot change dramatically, that is, the displacement between frames cannot be large, which is the premise of the OFM [30]. The motion information of the object is calculated by using the corresponding relationship of pixels between adjacent frames in the video sequence. Described in I_t to I_{t+1} , where I_t represents the frame at time t . The basic assumption is that the brightness of the object surface is basically unchanged and has continuity in space, so the basic calculation method of the optical flow is determined. If the pixel at a certain point in the video sequence is as $E(x, y, t)$. According to the motion state of the object, the basic calculation method of the optical flow is determined [31] as shown below.

$$E_x u + E_y v + E_t = 0 \quad (1)$$

where $E_x = \frac{\partial E}{\partial x}$ is the partial derivative of image brightness to x . Similarly, the same can be obtained $E_y = \frac{\partial E}{\partial y}$, $E_t = \frac{\partial E}{\partial t}$. u and v are two rates in the x and y , which is the optical flow value of the pixel. They can be expressed by Eq. (2).

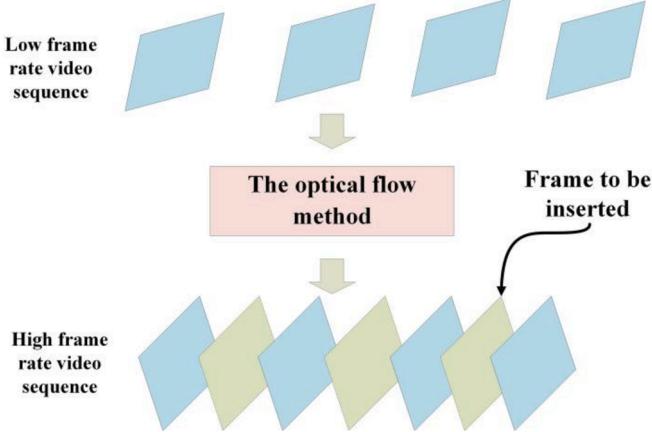


Fig. 1. The frame of insertion.

$$\begin{cases} u = \frac{dx}{dt} \\ v = \frac{dy}{dt} \end{cases} \quad (2)$$

Define two adjacent frames as I_0 and I_1 , in order to get the frame of intermediate time t , it is necessary to map the optical flow forward to time t , which is showed as Eq. (3).

$$u_t(\text{round}(x + t \times u_0(x))) = u_0(x) \quad (3)$$

where u_t represents the optical flow value at time t . The frame pixels generating the intermediate t time is expressed as s Eq. (4).

$$I_t(x) = \begin{cases} (1-t) \times I_0(x_0) + t \times I_1(x_1) & \text{if } Q_0(x_0) = 0, Q_1(x_1) = 0 \\ I_0(x_0) & \text{if } Q_1(x_1) = 1 \\ I_1(x_1) & \text{if } Q_0(x_0) = 1 \end{cases} \quad (4)$$

where x_0 denotes pixel coordinates of I_0 frames, and x_1 denoted as pixel coordinates of I_1 frame, which is obtained by Eq. (5).

$$\begin{cases} x_0 = x - t \times u_t(x) \\ x_1 = x + (1-t) \times u_t(x) \end{cases} \quad (5)$$

where $Q_1(x_1) = 1$ indicates that the pixels of a certain point are occluded and there is no corresponding optical flow. The OFM can obtain motion information between two frames, so as to insert the frame more accurately. The video interpolation is made to extend the data set in Fig. 1.

3.2. The improved SSD

As a classic algorithm framework, the SSD combining the regression idea can realize the effective detection of objects. In addition, the Anchor boxes mechanism can be used to perform multi-scale prediction. The SSD algorithm densely samples the image feature to be detected, uses the VGG-16 in the convolutional neural network as the network structure to extract features, and performs the regression and classification of the boxes. With the single point and multi-frame detection method, even the low-resolution image of the SSD algorithm can still achieve high-precision detection. The VGG-16 network model are used as convolution modules to extract features and generate feature maps. Then the multilayer convolution layer is connected to the pooled layer to provide additional feature extraction. Each candidate box has a corresponding category score, and the extracted default boxes are deeply learned through multi-task learning and non-maximum suppression (NMS) to quickly obtain the target object. The model composition of the SSD is shown in Fig. 2 [23].

However, the SSD model has poor ability and robustness to detect small targets [32]. In the actual production environment, defects have the characteristics of small targets and are difficult to identify. The traditional SSD uses layer by layer convolution to propagate the feature map after the feature layer. In the detection process, the receptive field of the feature map with large receptive field is often small, so to enlarge the size of the feature map with large receptive field, the deconvolution operation is introduced.

The deconvolution can make the semantic information on the shallow feature map more sufficient, so as to improve the ability of image representation and the detection capability. The deconvolution is the inverse operation of the convolution operation, which can increase the characteristics map resolution, and expand the receptive field. After the convolution operation with step size 2 is set, a 3×3 characteristic diagram is obtained.

$$M_{\text{out}} = d(M_{\text{in}} - 1) + k - 2p \quad (6)$$

In this equation, M_{out} is the size of the output, M_{in} is the characteristic input, d is the step size, k is the size of the convolution kernel, and p is the padding.

In this regard, a deconvolution layer is introduced to improve on the basis of the SSD model. The VGG network is a simple stack of traditional CNN structures, with a large number of parameters, but the ResNet has a stronger characterization capability with fewer parameters, which can accelerate network convergence, improve the accuracy, and avoid gradient disappearance and gradient explosion [33]. In this experiment, the ResNet is used instead of the VGG to obtain better results. Then, the residual network is added in a direct connection channel of the network, instead of learning a complete output, it learns the residual output of the previous network, without adding new parameters, and uses fewer

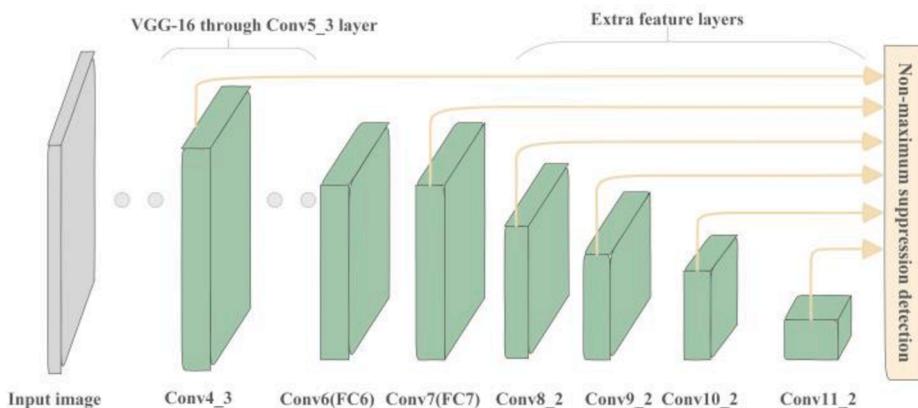


Fig. 2. The structure of the SSD network.

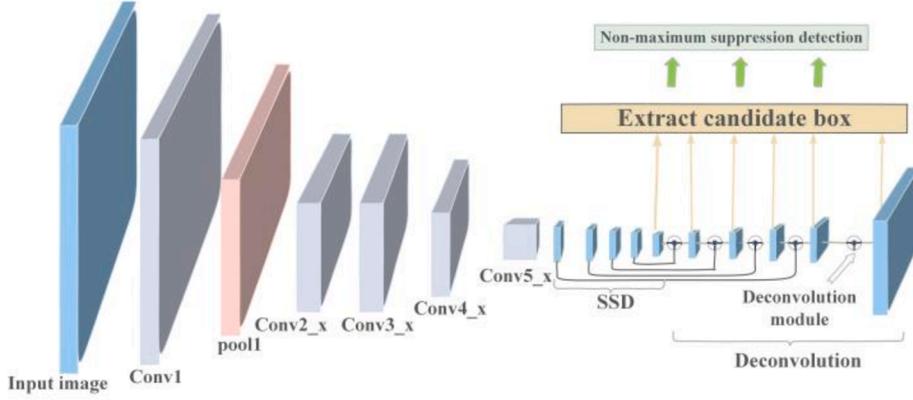


Fig. 3. The improved SSD model composition.

pooling layers, which greatly simplifies the goal and difficulty of learning.

Then the deconvolution module is established to realize the mapping and fusion of information from high-level features to low-level feature information. The small-size feature map is deconvolved, and then the feature maps are obtained from the deconvolution process and the convolution process are merged by the dot product. As shown in Fig. 3, the improved SSD composition is shown below.

The calculation method of the default box is shown in Eq. (7).

$$S_k = S_{\min} + \frac{S_{\max} - S_{\min}}{m-1}(k-1), k \in [1, m] \quad (7)$$

where m represents the prediction by m feature graphs, S is the size of the default box size, w_k^a represents the width and h_k^a represents the height of the default box, which can be expressed as follows.

$$\begin{cases} w_k^a = S_k \sqrt{a_r} \\ h_k^a = \frac{S_k}{\sqrt{a_r}} \end{cases} \quad k \in [1, m], a_r \in \left\{1, 2, 3, \frac{1}{2}, \frac{1}{3}\right\} \quad (8)$$

In this equation, a_r can be expressed as the aspect ratio of the default box.

The category of the default box is determined according to the confidence, and the detection result is determined by the NMS algorithm. The loss function of the convolution prediction can be calculated by the prediction of the confidence loss and the position loss. As shown in Eq. (9), the expression of the confidence loss is shown.

$$\begin{cases} L_{\text{conf}}(x, c) = -\sum_{i \in \text{Pos}}^N x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in \text{Neg}}^N \log(\hat{c}_i^z) \\ \hat{c}_i^p = \frac{\exp(\hat{c}_i^p)}{\sum_p \exp(\hat{c}_i^p)} \end{cases} \quad (9)$$

where x_{ij}^p is the indicator parameter, i represents the default box, and j represents the real box. \hat{c}_i^p and \hat{c}_i^z are the confidence of category p of the i -th candidate box and the confidence of the background class. The expression for the position loss prediction is obtained as follows.

$$\begin{cases} L_{\text{loc}}(x, l, g) = \sum_{i \in \text{Pos}} \sum_{m \in [\text{cx}, \text{cy}, \text{w}, \text{h}]} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m) \\ \text{smooth}_{L1}(x) = \begin{cases} 0.5x^2 \text{ if } |x| < 1 \\ |x| - 0.5 \text{ otherwise} \end{cases} \end{cases} \quad (10)$$

where cx and cy represent the coordinates of the position default box. \hat{g}_j^m is the deviation between the real box and the default box. w represents the width and h represents the height, respectively. l_i^m is the deviation. Therefore, the position prediction can be realized according to Eq. (10). The loss function can be designed based on the confidence prediction

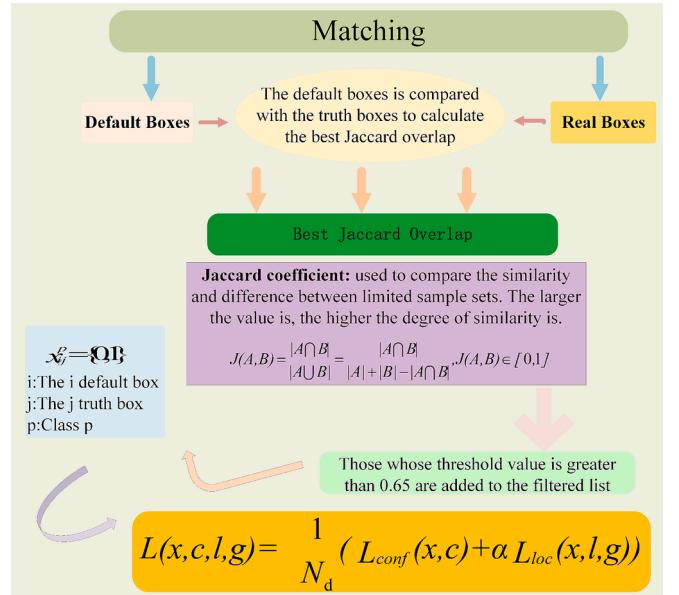


Fig. 4. The loss function graph.

and the location prediction as shown in Eq. (11).

$$L(x, c, l, g) = \frac{1}{N_d} (L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g)) \quad (11)$$

where $L_{\text{conf}}(x, c)$ is the confidence loss prediction, $L_{\text{loc}}(x, l, g)$ represents the location loss prediction, respectively. c represents the confidence of the default box, l represents the location information of the prediction frame, N_d represents the number of default boxes, g represents the location information of the real frame, and α is the weight parameter obtained using cross-entropy verification to describe weight relationship between the confidence loss and the location loss prediction. The value range is $(0, 1]$. The graphical representation can be seen as shown in Fig. 4.

3.3. The implementation process of the proposed method

Step 1: Collect the low frame rate video data set in the real.

Step 2: Apply the OFM to expand 25 frames of video to 60 frames for obtaining many extended data images of boiler water wall surface defects with the image size 300 * 300 dpi.

Step 3: Train and test on the expanded data set.

Step 4: The improved SSD model is introduced to identify the location of defects and output the recognition rate of surface defect.

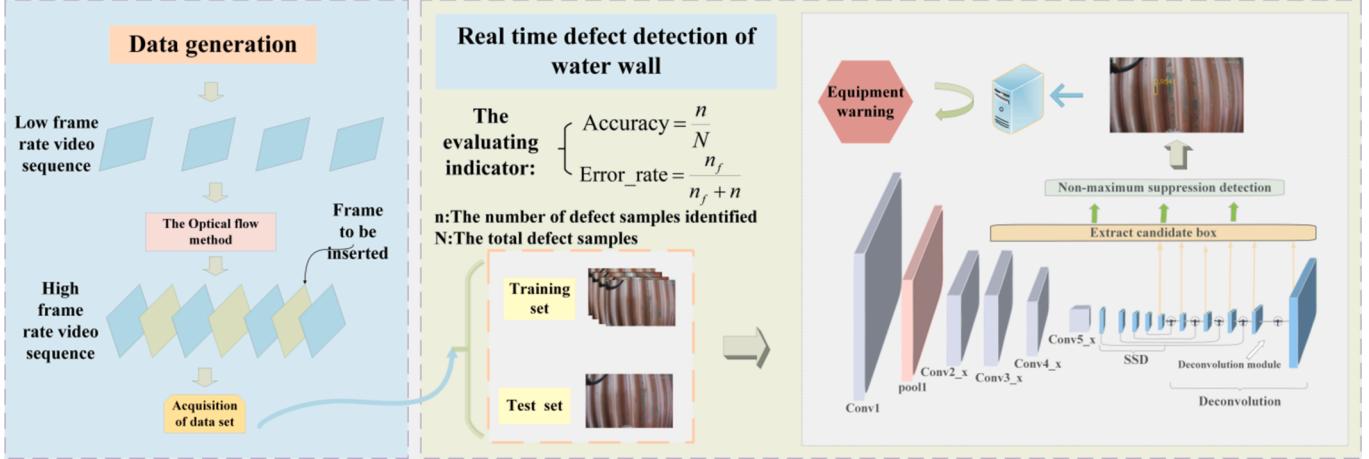


Fig. 5. The algorithm design graph.

The recognition rate of surface defects of the boiler water wall plays an important role in reducing accident losses. Therefore, the most important accuracy and the error detection rate are used as the criterion. The expression of the accuracy and the error detection rate is shown as in Eq. (12).

$$\left\{ \begin{array}{l} \text{Accuracy} = \frac{n}{N} \\ \text{Error_rate} = \frac{n_f}{n_f + n} \end{array} \right. \quad (12)$$

where n is the number of defect samples identified of boiler water wall that exceeds the threshold, N is the total defect samples of the boiler water wall, and n_f is the number that the algorithm judges to be a defect but the real situation is not a defect. Receiver operating characteristic curve (ROC) is a tool for performance measurement. The vertical axis is TPR and the horizontal axis is FPR. The calculation formulas of the two are as follows:

$$\left\{ \begin{array}{l} \text{TPR} = \frac{TP}{TP + FN} \\ \text{FPR} = \frac{FP}{TN + FP} \end{array} \right. \quad (13)$$

TP is the number of correctly classified and judged defects, FN is the number of defects but judged not to be defects, FP is the number of defects but recognized as defects, TN is the number of defects judged as defects, TPR and FPR are the horizontal and vertical axes of ROC curve, and the area under the curve (AUC) is used to evaluate the detection performance of the model. The AUC value is the sum of trapezoidal areas under the ROC curve as shown in Eq. (14).

$$\text{AUC} = \frac{1}{2} \sum_i^m (x_{i+1} - x_i)(y_i + y_{i+1}) \quad (14)$$

where, $x_{i+1} - x_i$ is the height of trapezoid, $y_i + y_{i+1}$ is the sum of the upper bottom length of the trapezoid plus the lower bottom length. The specific implementation of the proposed method is shown in Fig. 5.

After the data of the small sample is expanded, then the data is input into the SSD that introduces deconvolution. The deconvolution expands the receptive field so that the defects of small targets can be effectively

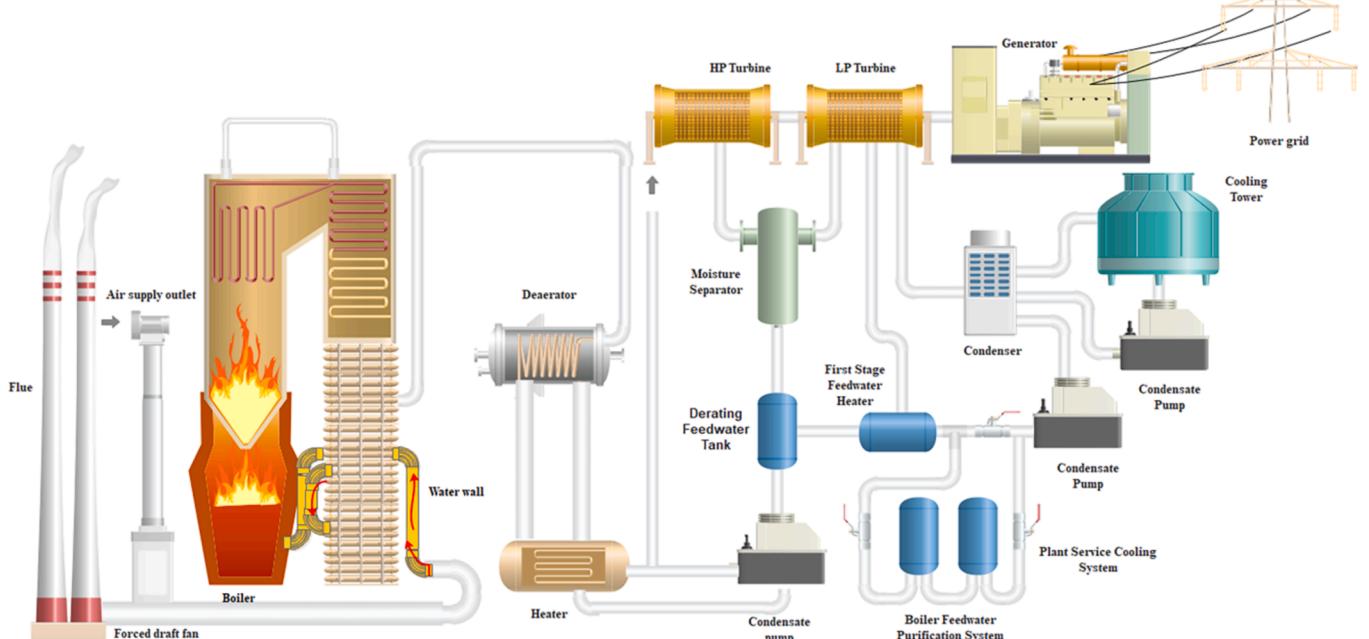


Fig. 6. The process flow of the power plant.

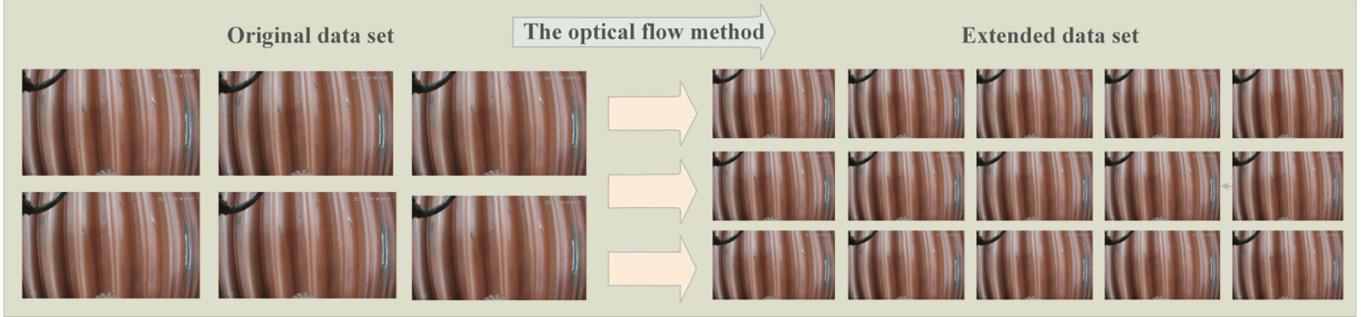


Fig. 7. The data augmentation graph.

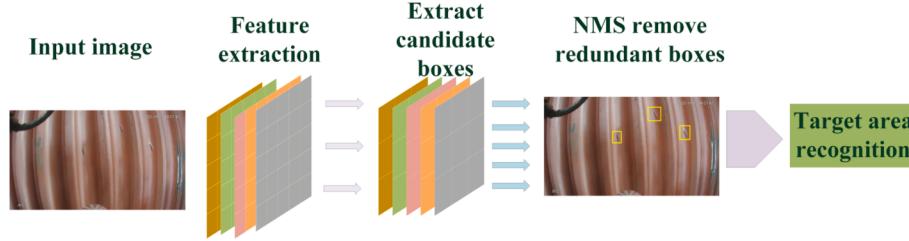


Fig. 8. The defect detection process.

detected, and then the defects are reported back to achieve the intelligent detection.

4. Experiments and results

4.1. Data collection and processing

The traditional thermal power plant is mainly composed of a boiler system, a turbine system and a power generation system [34]. The boiler system generates heat energy by burning coal and generates high-temperature steam by heating condensate, which is then transmitted to the steam turbine system to drive the rotor for generating induced electromotive force, and the power generation system transmits the electric energy to the power grid. As shown in Fig. 6, the main heated part of the boiler is composed of many steel pipes arranged around the furnace, which is the water wall.

Therefore, for the normal operation of the thermal power plant to reduce losses and ensure the safe production which is significant for the detection of water wall defects.

In this experiment, the data of different boiler surfaces are obtained from the intelligent robot on the boiler water wall. In order to solve the insufficient amount of data that may cause over-fitting of the surface defect model, the data set needs to be expanded and strengthened. The data is collected from different video data information collected by the wall-climbing robot from different boiler water wall surfaces. In this experiment, the OFM is used to expand the 25 frames of the video to 60 frames for obtaining 720 defect samples with the image size of 300 * 300 dpi, which greatly increases the number of initial samples. The OFM generates intermediate data frames between two pictures through the optical flow, and the two pictures are different in pixel points. Fig. 7 shows the situation after partial data augmentation.

The boxes of the proposed method are used to classify the default boxes as sliding windows to achieve detection and improve the detection speed. Then the extracted default boxes are obtained, in which the length and the width are different. Finally, the NMS is used to remove the redundant default boxes, thereby achieving the target detection.

Fig. 8 shows the specific detection process.

Table 1
The comparison of methods.

Methods	Accuracy	Error_rate	AUC	Time
OFM SSD	94.7 %	5.3 %	0.887	143.17 ms
T_SSD	92.9 %	7.1 %	0.865	143.29 ms
YOLO	89.9 %	10.5 %	0.859	145.31 ms
SSD	91.3 %	7.9 %	0.856	143.47 ms
R_CNN	91.6 %	8.8 %	0.839	146.35 ms
DSSD	92.0 %	7.1 %	0.870	143.21 ms

4.2. Experiments results

To assess the effectiveness and generalizability of the OFM_SSD, experiments focusing on surface defects of actual water walls are conducted. The experiments are conducted using an Intel Core i7-9700 CPU, an NVIDIA GeForce GTX 1600 GPU, and 16 GB of RAM, running Windows 10 64-bit. The software environment included Python 3.7, OpenCV 4.5.1, and PyTorch 1.7.1. The effectiveness of the OFM_SSD is proved by the accuracy of defect location detection, the classification effect of defects and the false detection rate of defects. The results of the accuracy, the error_rate, the AUC and the detection time per frame are shown in Table 1.

As illustrated in Table 1, with regard to the principal metric of processing time, the proposed OFM_SSD method has a processing time of 143.17 ms, representing the shortest method among all models, which demonstrates that the OFM_SSD method exhibits excellent real-time performance while maintaining a high level of accuracy with 94.7 %.

Furthermore, the proposed OFM_SSD method demonstrates superior performance compared to the T_SSD, which employs rotation and shear Gaussian blur for dataset expansion. Specifically, the accuracy of the OFM_SSD is improved by 1.94 %, while the error rate is reduced by 25.4 %, and the AUC increases by 2.5 %. In comparisons with established defect detection methods such as the YOLO, the SSD, and the R_CNN, the accuracy rates of the OFM_SSD are enhanced by 4.56 %, 3.72 %, and 3.38 %, respectively. Concurrently, false detection rates decrease significantly by 49.5 %, 32.9 %, and 39.8 %, with corresponding AUC improvements of 3.3 %, 3.6 %, and 5.7 %. These results indicate that the OFM method effectively enhances defect recognition, particularly for

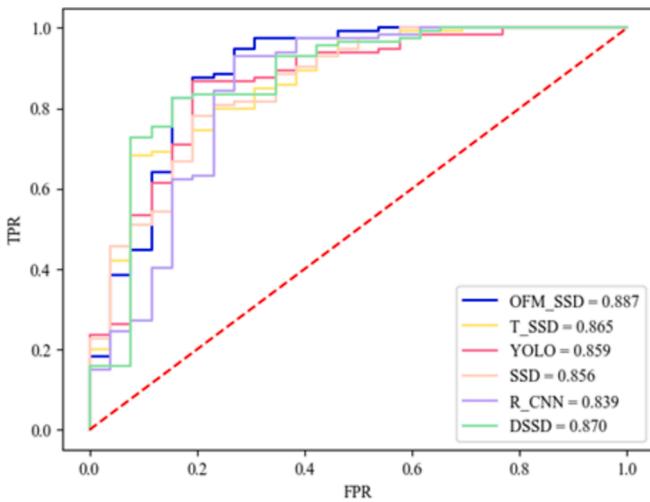


Fig. 9. The ROC curve comparison of models.

small sample datasets. Furthermore, when compared to the DSSD, the OFM_SSD shows an accuracy increase of 2.93 %, 25.4 % reduction in the error rate, and 2.0 % rise in AUC. This reinforces the notion that the incorporation of OFM bolsters detection capabilities by enhancing the receptive field of the feature area, thereby improving the overall detection performance. The ROC curves for each model are illustrated in Fig. 9, further validating the marked performance advancement achieved by the OFM_SSD through the integration of deconvolution techniques.

The detection results of the OFM_SSD in the actual water wall defect detection process are shown in Fig. 10.

Which in the number represents the accuracy of the defect recognition, and the yellow box is the defect position detected by the OFM_SSD, the T_SSD, the YOLO, the SSD and the R_CNN methods. From Fig. 10 that the recognition rate of defects based on the OFM_SSD is higher than

other algorithms. Moreover, the T_SSD, the YOLO, the SSD algorithms without extended data sets, the R_CNN and DSSD methods are easy to miss the identification and occur error detection of water wall defects.

It is evident from Fig. 11 that the dimensions of the defects in the inner wall of the boiler are insignificant, rendering them challenging to discern directly on the surface of the high-temperature and high-pressure boiler. Concurrently, the presence of corrosion may also exert a detrimental influence on the efficacy of the defect detection.

The OFM_SSD has the capacity to enhance the analysis of small sample and small target data, while also demonstrating a high level of detection capability for small targets. This results in the optimal performance in the surface defect detection. Therefore, SSDs with deconvolution operations contribute to an improvement in the capability for the detection of small targets. Furthermore, the OFM_SSD demonstrates superior resilience and reliability when compared to alternative models in the presence of identical corrosive environmental influences.

Boiler inner wall defect detection is a task targeted at fixed scenarios. Considering that the movement speed of the detection machine on the inner wall is not fast due to environmental factors. After deeply investigating the actual needs of enterprises in the industry, this application scenario does not require completely real-time detection speed to ensure the accuracy of the detection results and avoid resource waste caused by repeated detection of the same defect location. The light conditions remain essentially unaltered, as the detection machine is responsible for transporting the light when examining the interior of the boiler. Referring to the realities of the above application scenarios, the detection speed and efficiency of the proposed method fully satisfy the practical requirements while significantly outperforming other algorithms. Therefore, the experimental results prove that the OFM_SSD can better achieve the intelligent detection of real boiler water wall defects, reducing the loss to the factory after the defect occurs and the cost of manpower testing, so as to bring greater benefits to enterprises and factories.

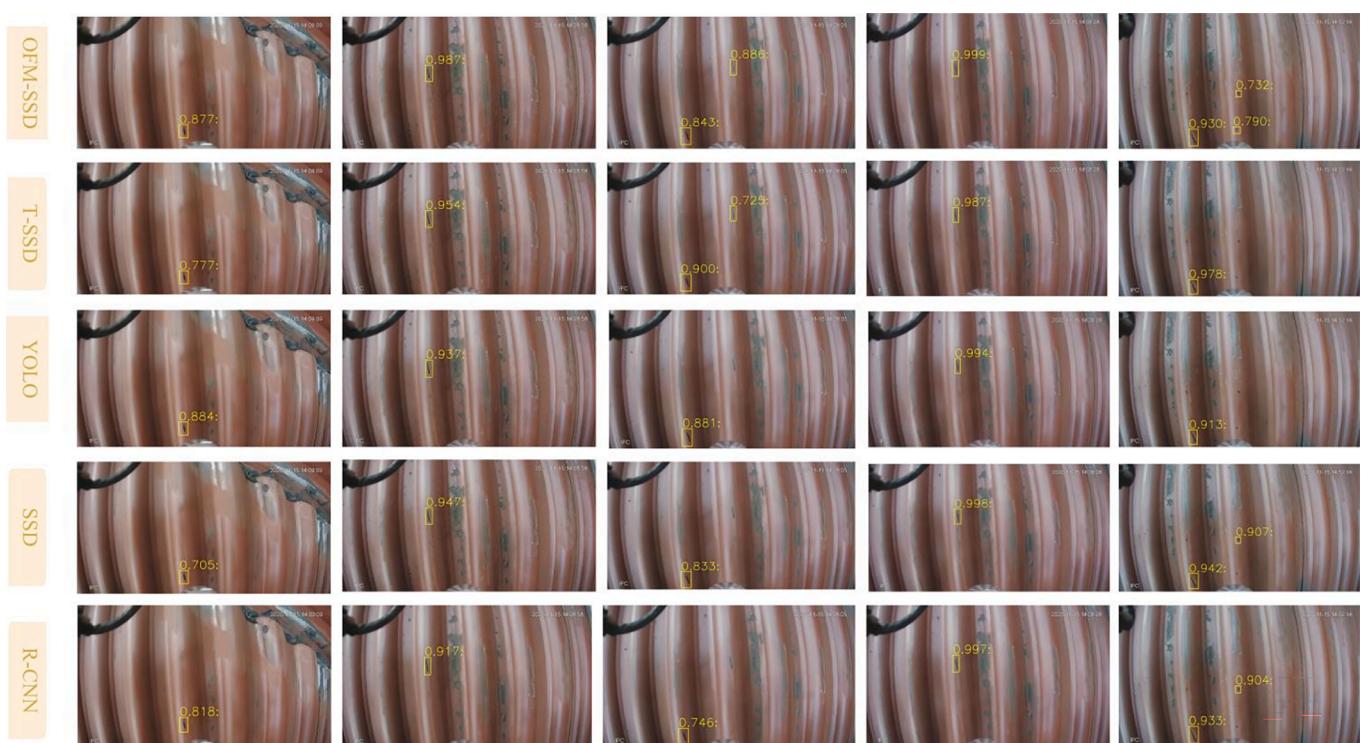


Fig. 10. The comparison graph of various algorithms.

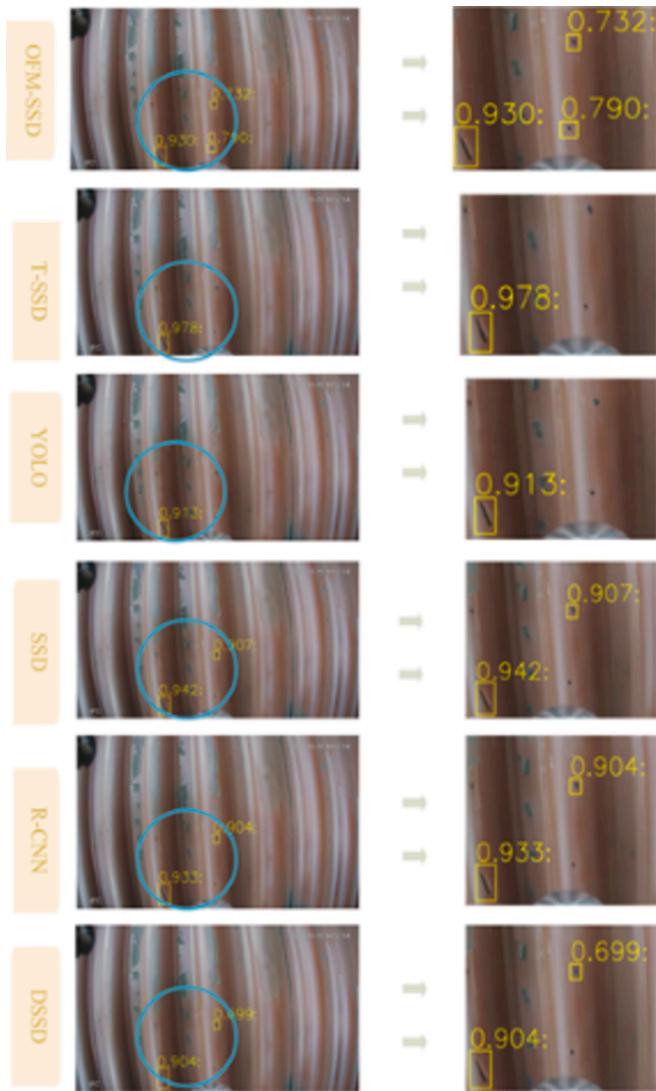


Fig. 11. The local detection of water wall defects.

5. Conclusion

This paper presents a novel intelligent detection model for the identification of surface defects on water walls, designated as the OFM_SSD. The OFM is employed to effectively expand the data set. Subsequently, the deconvolution operation based on the SSD facilitates the development of an intelligent defect detection system for the boiler water wall. The experimental results demonstrate that the OFM_SSD is capable of accurately detecting surface defects on the boiler water wall. Its accuracy and detection capability are superior to those of existing classical target detection algorithms, including the T_SSD, the YOLO, the SSD algorithms without extended data sets, the R_CNN and the DSSD, by 1.94 %, 4.56 %, 3.72 %, 3.38 % and 2.93 %, respectively. Furthermore, the OFM_SSD is capable of meeting the current requirements of smart factories. It enables industrial production departments to identify and address defects in a timely manner, thereby reducing the risk of accidents and associated losses. Concurrently, the OFM_SSD also effectively reduces the cost of defect detection in the boiler water wall, thereby facilitating greater profitability for the factory.

In the future word, to enhance the adaptability and portability of the proposed method, we plan to expand the dataset of defect types to encompass a broader range of potential surface defect scenarios, utilizing both simulations and actual data acquisition. Additionally, we aim to implement more complex structures in the deconvolution operation to

improve the model capacity for detecting defects on smaller targets.

CRediT authorship contribution statement

Yongming Han: Project administration, Methodology, Software, Writing – review & editing. **Lei Wang:** Data curation, Investigation, Methodology, Software, Writing – original draft. **Jintao Liu:** Visualization, Writing – review & editing. **Liang Yuan:** Conceptualization, Writing – review & editing. **Hongxu Liu:** Methodology, Resources, Writing – review & editing. **Bo Ma:** Project administration, Funding acquisition, Resources, Writing – review & editing. **Zhiqiang Geng:** Project administration, Funding acquisition, Resources, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work is partly funded by National Natural Science Foundation of China, China (62373035 and 62203038), and Open Fund of State Key Laboratory of Mechanical Transmission for Advanced Equipment, China (SKLMT-MSKFKT-202312) and Funds for First-class Discipline Construction of Beijing University of Chemical Technology, China (Grant No. XK2023-10).

References

- [1] J. Cui, T. Chai, X. Liu, Deep-neural-network-based economic model predictive control for ultrasupercritical power plant, *IEEE Trans. Ind. Informat.* 16 (2020) 5905–5913.
- [2] H. Ma, T. Peng, C. Zhang, C. Ji, Y. Li, M. Nazir, Developing an evolutionary deep learning framework with random forest feature selection and improved flow direction algorithm for NO_x concentration prediction, *Eng. Appl. Artif. Intell.* 123 (2023) 106367.
- [3] L. Wei, F. Fang, Hoo-LQR-based coordinated control for large coal-fired boiler-turbine generation units, *IEEE Trans. Ind. Electron.* 64 (2017) 5212–5221.
- [4] F. Mahmoudi, T. Najafabadi, A. Moarefianpour, Developing an operating instruction to attenuate pressure stress during fuel shortage events of a power boiler based on dynamic modelling, *Appl. Therm. Eng.* 200 (2022) 117565.
- [5] G. Fu, L. Zhou, J. Wang, M. Shi, Analysis of an explosion accident at Dangyang Power Plant in Hubei, China: causes and lessons learned, *Saf. Sci.* 102 (2018) 134–143.
- [6] A.H. Assefinejad, A. Kermanpur, A.M. Eslami, J. Kangazian, Failure investigation of water wall tubes in a drum boiler of a thermal power plant, *Eng. Fail. Anal.* 118 (2020) 104869.
- [7] J. Miao, J. Wang, Q. Miao, An enhanced multifeature fusion method for rotating component fault diagnosis in different working conditions, *IEEE Trans. Reliab.* 70 (2021) 1611–1620.
- [8] L. Wei, Q. Zhou, N. Li, Experimental study and simulation analysis of heat and deformation in the water walls of an opposed firing boiler under flexible operating conditions, *Appl. Therm. Eng.* 213 (2022) 118726.
- [9] Y. Zhang, Z.Y. Dong, W. Kong, K. Meng, A composite anomaly detection system for data-driven power plant condition monitoring, *IEEE Trans. Ind. Informat.* 16 (2020) 4390–4402.
- [10] A. Foudazix, A. Mirala, M.T. Ghasr, K.M. Donnell, Active microwave thermography for nondestructive evaluation of surface cracks in metal structures, *IEEE Trans. Instrum. Meas.* 68 (2019) 576–585.
- [11] Y. Lu, G. Zhang, S. Duan, F. Chen, A pyramid auxiliary supervised U-Net model for road crack detection with dual-attention mechanism, *Displays* 84 (2024) 102787.
- [12] Y. Han, Y. Wang, Z. Chen, Y. Lu, X. Hu, L. Chen, Z. Geng, Multiscale variational autoencoder regressor for production prediction and energy saving of industrial processes, *Chem. Eng. Sci.* 284 (2024) 119529.
- [13] W. Gao, D. Zhang, E. Zhang, Defect detection in the dead zone of magnetostrictive sensor for pipe monitoring, *IEEE Sens. J.* 21 (2021) 3420–3428.
- [14] Z. Zhang, L. Liu, X. Zhao, L. Zhang, J. Wu, Y. Zhang, Z. Li, DSSO-YOLO: A fast detection model for densely stacked small object, *Displays* 82 (2024) 102659.

- [15] Y. Tan, R. Cai, J. Li, P. Chen, M. Wang, Automatic detection of sewer defects based on improved you only look once algorithm, *Autom. Constr.* 131 (2021) 103912.
- [16] H. Yao, W. Yu, X. Wang, A feature memory rearrangement network for visual inspection of textured surface defects toward edge intelligent manufacturing, *IEEE Trans. Autom. Sci. Eng.* 20 (2023) 2616–2635.
- [17] G. Golcarenarenji, I. Martinez-Alpiste, Q. Wang, J.M. Alcaraz-Calero, Illumination-aware image fusion for around-the-clock human detection in adverse environments from unmanned aerial vehicle, *Expert Syst. Appl.* 204 (2022) 117413.
- [18] K. Yan, J. Su, J. Huang, Y. Mo, Chiller fault diagnosis based on VAE-enabled generative adversarial networks, *IEEE Trans. Autom. Sci. Eng.* 19 (2022) 387–395.
- [19] P. Tan, X. Li, Z. Wu, J. Ding, J. Ma, Y. Chen, Y. Fang, Y. Ning, Multialgorithm fusion image processing for high-speed railway dropper failure-defect detection, *IEEE Trans. Syst. Man Cybern. Syst.* 51 (2021) 4466–4478.
- [20] Y. Han, L. Wang, Y. Wang, Z. Geng, Intelligent small sample defect detection of concrete surface using novel deep learning integrating improved YOLOv5, *IEEE/CAA J. Autom. Sin.* 11 (2024) 545–547.
- [21] Y. Han, Z. Li, X. Hu, Y. Wang, Z. Geng, Novel long short-term memory model based on the attention mechanism for the leakage detection of water supply processes, *IEEE Trans. Syst. Man Cybern.: Syst.* 54 (2024) 2786–2796.
- [22] X. Zhang, T. Wan, Z. Wu, B. Du, Real-time detector design for small targets based on bi-channel feature fusion mechanism, *Appl. Intell.* 52 (2022) 2775–2784.
- [23] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, A. Berg, SSD: single shot MultiBox detector, *Eur. Conf. Comput. Vis.* 9905 (2016) 21–37.
- [24] B. Sun, X. Zhao, H. Zhang, R. Bai, T. Li, EEG motor imagery classification with sparse spatiotemporal decomposition and deep learning, *IEEE Trans. Autom. Sci. Eng.* 18 (2021) 541–551.
- [25] L. Gong, X. Huang, Y. Chao, J. Chen, B. Lei, An enhanced SSD with feature cross-reinforcement for small-object detection, *Appl. Intell.* 53 (2023) 19449–19465.
- [26] Z. Huang, Z. Yin, Y. Ma, C. Fan, A. Chai, Mobile phone component object detection algorithm based on improved SSD, *Proc. Comput. Sci.* 183 (2021) 107–114.
- [27] X. Ma, X. Li, X. Tang, B. Zhang, R. Yao, J. Lu, Deconvolution feature fusion for traffic signs detection in 5G driven unmanned vehicle, *Phys. Commun.* 47 (2021) 101375.
- [28] J.J. Gibson, The perception of the visual world, *Science* 113 (1951) 535.
- [29] Q. Zhang, J. Xu, M. Crane, C. Luo, See the wind: wind scale estimation with optical flow and VisualWind dataset, *Sci. Total Environ.* 846 (2022) 157204.
- [30] S.B. Block, R.D. da Silva, L.B. Dorini, R. Minetto, Inspection of imprint defects in stamped metal surfaces using deep learning and tracking, *IEEE Trans. Ind. Electron.* 68 (2021) 4498–4507.
- [31] B.K.P. Horn, B.G. Schunck, Determining optical flow, *Artif. Intell.* 17 (1981) 185–203.
- [32] Z. Chen, K. Wu, Y. Li, M. Wang, W. Li, SSD-MSN: an improved multi-scale object detection network based on SSD, *IEEE Access* 7 (2019) 80622–80632.
- [33] M.K. Panda, A. Sharma, V. Bajpai, B.N. Subudhi, V. Thangaraj, V. Jakhetiya, Encoder and decoder network with ResNet-50 and global average feature pooling for local change detection, *Comput. Vis. Image Understand.* 222 (2022) 103501.
- [34] Z. Su, G. Zhao, J. Yang, Y. Li, Disturbance rejection of nonlinear boiler-turbine unit using high-order sliding mode observer, *IEEE Trans. Syst. Man Cybern. Syst.* 50 (2020) 5432–5443.