

Pipeline Magnetic Flux Leakage Image Detection Algorithm Based on Multiscale SSD Network

Lijian Yang, Zhujun Wang , and Songwei Gao

Abstract—In order to solve the problem of low detection accuracy of small targets in the SSD detection algorithm, a pipeline magnetic flux leakage image detection algorithm based on multiscale SSD network is proposed in this paper. The dilated convolution and attention residual module are introduced into the SSD algorithm to fuse the low-resolution high-semantic information feature map with the high-resolution low-semantic information feature map so as to improve the resolution of the low-resolution feature map and provide detailed features for small targets. Finally, the target location and category are obtained by regression algorithm. The experimental results show that the proposed algorithm can automatically identify the location of circumferential weld, spiral weld, and defect of magnetic flux leakage data. Compared with the original SSD algorithm framework, the improved algorithm has higher detection accuracy, 97.62%, 18.01% lower false detection rate, 18.36% lower false detection rate, better robustness, and obvious effect on small target detection.

Index Terms—Attention residual module, magnetic flux leakage in pipeline, residual network, SSD network, target detection.

I. INTRODUCTION

MAGNETIC flux leakage (MFL) detection technology plays an important role in long-distance oil and gas pipelines. Due to corrosion, abrasion, accidental damage, and other reasons, pipeline leakage accidents occur frequently, resulting in huge loss of life and property and environmental pollution. At present, the identification and analysis of MFL mostly adopts manual interpretation, but the amount of MFL data is huge and there are many types, so there are some problems such as time-consuming, leakage judgment, misjudgment, and so on. Therefore, how to detect the magnetic flux leakage signal efficiently and accurately is the focus of current paper.

Target detection is one of the important research directions in the field of computer vision. Traditional target detection meth-

Manuscript received March 9, 2019; revised May 22, 2019; accepted June 3, 2019. Date of publication July 2, 2019; date of current version January 4, 2020. Paper no. TII-19-0607. (Corresponding author: Zhujun Wang.)

The authors are with the School of Instrumentation Science and Technology, Shenyang University of Technology, Shenyang 110870, China (e-mail: yanglijian888@163.com; wangzhujun22@163.com; 568138924@qq.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TII.2019.2926283

ods use classifiers to classify features after feature descriptors are constructed to extract features. In the feature extraction stage, manual intervention is needed to obtain the target-related feature information in the original image input. Qiu *et al.* [1] applied genetic algorithm and back propagation (BP) neural network to quantitative identification of magnetic flux leakage of microcracks in pipelines; Tsukada *et al.* [2] proposed an unsaturated alternating current flux leakage test method based on magnetoresistance gradiometer to analyze internal cracks in steel; Liu *et al.* [3] proposed a window feature generated by MFL measurement to detect defects and confirm them. Kopp and Willems [4] used three-axis magnetic flux leakage method to detect and identify defects of different sizes; Cheng *et al.* [5] used pulse thermal induction to identify single defects near weld seam; Chen *et al.* [6] used radial basis function neural network as a defect prediction model for MFL and gradient descent method to update defect contour; Wang *et al.* [7] used the hidden Markov random field model and the Bayesian model; Liu *et al.* [8] proposed SSD model.

This algorithm relies heavily on the prior knowledge of feature designers has low efficiency and high error rate. Aiming at the shortcomings of traditional detection methods, target detection algorithm based on deep learning has become a new direction in the field of computer vision. Feng *et al.* [9] improved the recognition accuracy by adding local normalization to convolution neural network; Inkyu *et al.* [10] used encoder-decoder cascaded convolutional neural networks (CNN) to detect weed multispectral images; Arandjelovic *et al.* [11] proposed a new NetVLAD network to identify the location of a given inquiry photo; Yu *et al.* [12] introduced SSD model for multisaliency target detection; Tang *et al.* [13] fused single shot multibox detector (SSD) model with color histogram features and HOG features to achieve multiscale target detection; Tang *et al.* [14] proposed an end-to-end single convolution neural network to locate the target features in the image; Liu *et al.* [15] used gaussian mixture model (GMM) and CNN to improve the recognition accuracy; Gao *et al.* [16] proposed a target classification method based on LIDAR; Yang *et al.* [17] make use of context information to make the network have the ability to extract multiscale features; Wang *et al.* [18] proposed a color-guided target tracking algorithm based on particle filter to solve the problem of target loss in complex environment.

Liu *et al.* [19] proposed a convolutional neural network payload classification method and a recurrent neural network payload classification method for attack detection; Li *et al.* [20] proposed a deep convolutional computation model is proposed to learn hierarchical features of big data; Gao *et al.* [21] used the multibranch fusion of CNN to quickly extract the characteristics of high-speed information systems; Sakurai *et al.* [22] used CNN to estimation of image aspect ratio; Wang *et al.* [23] built a tree structure to learn fine-grained features; Wang *et al.* [24] fused convolution neural network and the hidden Markov model to detect multiple faults in mechanical systems; Xu *et al.* [25] proposed a new semisupervised multiconcept semantic image based on deep learning retrieval method; Liu *et al.* [26] combined “deep neural network” with data enhancement based on generated confrontation network to screen out low-quality confrontation samples; Wang *et al.* [27] used SSD network to detect SAR images; Akcay *et al.* [28] used CNN detection the background of X-ray baggage security images; Zheng *et al.* [29] proposed a novel electricity-theft detection method based on wide and CNN model; Cui *et al.* [30] proposed recognition and classification of malware images using CNN.

SSD network model can accurately and quickly detect multiple target objects of different scales, belongs to a single-stage detector. This paper uses this model to detect MFL image. However, the SSD model has the problem of poor detection ability for small targets. Therefore, this paper integrates the dilated convolution and residual module into SSD model, and proposes a multiscale SSD network model. The dilated convolution is used to extract more abundant low-level and high-level semantic information features of small targets, and then the image details and context semantic information can be effectively learned by fusing the features. Improve the detection accuracy of small target defects.

II. PSEUDOCOLOR IMAGE OF PIPELINE MAGNETIC FLUX LEAKAGE

The magnetic flux leakage signal in the pipeline is acquired by magnetic flux leakage detector, and pseudocolor image is generated by pseudocolor processing technology, so that the complete magnetic flux leakage signal data can be displayed by color, which has better visual effect.

A. Principle of Internal Detection of Magnetic Flux Leakage in Pipeline

The magnetic flux leakage internal detection of pipeline is to use permanent magnet to magnetize the ferromagnetic pipeline sufficiently through a rigid brush to make it saturated or nearly saturated, and to judge whether the pipeline is defective by the change of magnetic induction line. Fig. 1 is the schematic diagram of pipeline magnetic flux leakage internal detection.

If there is no defect in the pipe wall, the magnetic induction line is parallel to the inside of the pipe; if there is defect in the pipe wall, the magnetic induction line will leak out of the pipe surface, resulting in magnetic flux leakage. The magnetic flux leakage signal collected by the magnetic sensor is imaged to determine the feature and location of the defect.

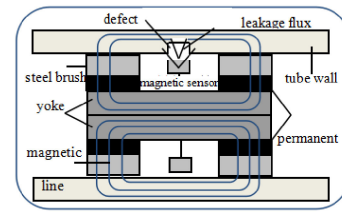


Fig. 1. Principle of pipeline magnetic flux leakage.

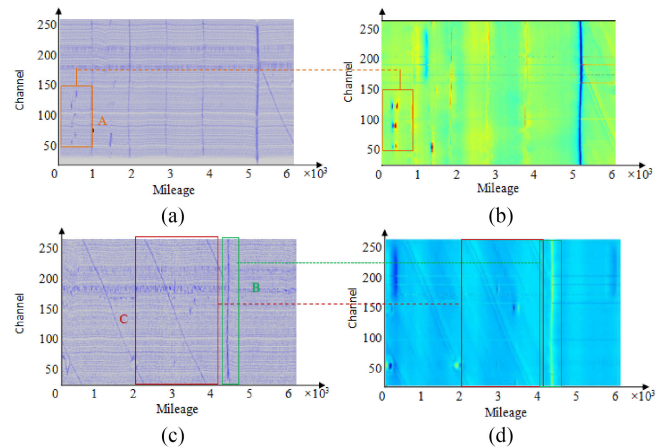


Fig. 2. Pseudocolor image of pipeline MFL internal detection. (a) MFL defects signal. (b) Pseudocolor image of MFL defects. (c) MFL weld signal. (d) Pseudocolor image of MFL weld.

B. Magnetic Leakage Data Imaging of Pipeline

There are three main visualization methods of magnetic flux leakage data: curve graph display, gray image display, and pseudocolor image display. However, the magnetic flux leakage signal is a three-dimensional (3-D) signal, so the curve and gray image cannot fully display all the data information. Pseudocolor processing technology is an effective image processing method for image enhancement. It can improve the resolution of image details and get clear and natural images. The color value can be used as the 3-D feature of MFL signal, so that the complete data of MFL signal can be displayed by color, as shown in Fig. 2.

Density segmentation is one of the simplest methods for pseudocolor enhancement. Its working principle is to divide the gray level of the image from 0 to M_0 into N intervals I_i ($i = 1, 2, \dots, N$), and assign a color C_i to each interval, so that a pseudocolor image can be generated. Assuming the gray level range of the original gray level image is $0 \leq f(x, y) \leq L$, the gray level range is divided into k segments, $[(I_0, I_1, \dots, I_k), I_0 = 0, I_k = L]$ to map each gray level into a color, and the mapping relationship is as follows:

$$\begin{cases} g(x, y) = C_i \\ I(i-1) \leq f(x, y) \leq I(i) = 1, 2, L, k \end{cases} \quad (1)$$

Among them, $g(x, y)$ is the output pseudocolor image; C_i is the mapping color.

Fig. 2 shows pseudocolor image of pipeline MFL signal data after image processing.

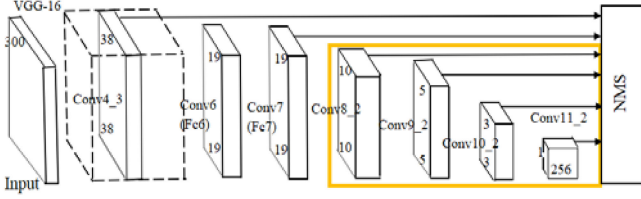


Fig. 3. Structural diagram of SSD model.

Fig. 2(a) and (b) shows magnetic flux leakage defect curves and pseudocolor graphs of pipelines, respectively. (a) represents pipeline defects. Pipeline defect refers to the damage caused by weather, misoperation, pipe material, external force, and other problems, which makes the pipeline defective with the growth of service.

Fig. 2(c) and (d) shows the curves and pseudocolor diagrams of pipeline MFL welds, in which B is circumferential weld and C is spiral weld. In long-distance oil and gas pipelines, welds are produced at the welds between the pipes. Welds is generally divided into spiral weld and girth weld.

III. IMPROVED IMAGE DETECTION ALGORITHM FOR MAGNETIC LEAKAGE IN SSD PIPELINE

SSD is scale sensitive and performs poorly in small-scale target detection. Dilated convolution and attention residual structure are introduced into the network to enhance the high-level semantic expression of underlying features, thereby improving the accuracy of small-scale target detection.

A. SSD Network Model

SSD network belongs to the single-step target detection algorithm and feedforward CNN. By using multiscale regional features of each position of the whole image, the confidence of the target category and the deviation between candidate frame and real frame are obtained to locate the target. The network structure is shown in Fig. 3.

SSD network consists of the first five layers of VGG-16 network and four additional convolution layers, there are 2, 2, 3, 3, 4, 2, 2, 2. Out of which 4, 6, 7, 8, and 9 convolution layers are output simultaneously for different prediction modules. The feature maps of different layers have the offset relative to the default box of the target and the confidence of different categories. The prediction module uses a set of 3×3 convolution kernels to predict the position of the target in the feature map. Finally, the location of the target in the image and the specific categories are obtained by nonmaximum suppression (NMS).

The target default box is defined as follows:

$$s_k = s_{\min} + \frac{s_{\max} - s_{\min}}{m - 1}(k - 1), k \in [1, m]. \quad (2)$$

Among them, s_{\min} and s_{\max} are set to 0.2 and 0.9, representing the lowest and highest scales, respectively. m is the number of convolution layers for prediction, with a value of 5. The default box can set different width ratios. The width and height $h_k^a = s_k / \sqrt{a_k}$ of each default box are sum, and the $w_k^a = s_k \sqrt{a_k}$

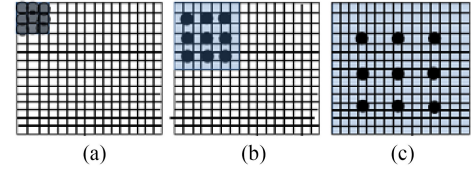


Fig. 4. Schematic diagram of dilated convolution with different. (a) Rate = 1. (b) Rate = 2. (c) Rate = 4.

center is $(\frac{i+0.5}{|f_k|}, \frac{j+0.5}{|f_k|})(i, j \in [0, |f_k|], |f_k|$ is the size of the k feature graph. Target default boxes can generate predictions covering various sizes and shapes of input.

B. Dilated Convolution and Attention Residual Module

1) *Dilated Convolution*: Dilated convolution, also known as inverse convolution, is the inverse process relative to convolution in CNNs. The feature graph is expanded into one-dimensional (1-D) vector, the convolution core is expanded into sparse matrix, and the dilated convolution operation is matrix multiplication. The calculation of dilated convolution is shown in

$$C^T \bullet \vec{Y} = \vec{X}. \quad (3)$$

Among them, X and Y represent 1-D vectors derived from the expansion of feature graphs, and C represent sparse matrices derived from convolution kernels.

Dilated convolution adds a parameter r (rate) to the original convolution. By controlling the size of rate, the convolution horizons of different sizes can be obtained. It expands the convolution core to the scale constrained by the expansion coefficient and fills the unoccupied area of the original convolution core with 0. Among them, R represents the expansion coefficient of convolution kernels. The effective convolution kernels obtained are $f_h + (f_h - 1)(r - 1)$ in height and $f_w + (f_w - 1)(r - 1)$, in which f_h represents the height of the original convolution and f_w is the width of the original convolution kernels. Dilated convolution can use different expansion coefficients in the same convolution core, as shown in Fig. 4.

Fig. 4(a) shows a 3×3 convolution nucleus with a receptive field size of 3×3 at the expansion coefficient rate = 1; Fig. 4(b) shows a 3×3 cavity convolution at the expansion coefficient rate = 2, and the receptive field size increases to 7×7 ; Fig. 4(c) shows a 3×3 convolution nucleus at the expansion coefficient rate = 4, and the receptive field size increases to 15×15 . Cavity convolution with expansion coefficients of 1 and 2 has the same effect as convolution cores of 7×7 and 15×15 . Therefore, the use of cavity convolution can reduce network parameters and improve the network through four different rates of convolution acting on the same input to feel the characteristics of different visual fields. The formula for calculating the size of the receptive field is as follows:

$$v = ((k_{\text{size}} + 1) \times (r_{\text{rate}} - 1) + k_{\text{size}})^2. \quad (4)$$

In order to avoid introducing too many transposed convolutions and increase the computational complexity, only low-level feature maps are improved. When the input image resolution is

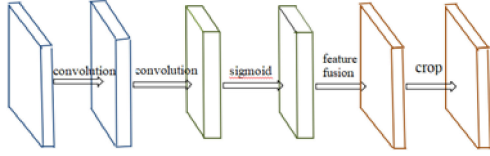


Fig. 5. Attention Residual Module.

300×300 , 75×75 high-resolution low-semantic information feature map is added to the detection layer, and the feature maps with resolution of 19×19 and 38×38 are transposed convolution, respectively. The feature extraction ability is improved to 16 and 17 convolution layers, respectively, compared with the original 7 layer convolution and 10 layer convolution layers.

After the low-resolution feature is transposed and convolved, the resolution of the feature map must be consistent with that of the high-resolution feature map in order to merge the feature map. Convolution kernel is 4×4 and step size is 2 for feature map of 19×19 resolution. Convolution kernel is 3×3 for feature map of 38×38 resolution, step size is 2 and expanded edge is 1. The resolution calculation formula of feature map after convolution is as follows:

$$O = S \times (L - 1) + H - 2 \times P. \quad (5)$$

Among them, O is the resolution of the transposed convolution output feature map; S is the step size; L is the resolution of the input feature map; H is the size of the convolution core; P is the supplementary size of the edge.

2) Attention Residual Module: Adding task-oriented subnetworks after detectors can effectively improve the detection accuracy. The saliency target detection method carries out saliency analysis of candidate targets through visual attention mechanism. Attention residual module is added to the original SSD network to enhance the selection of regions of interest.

Each attention module contains two branches: mask branch and trunk branch. The main branch can be any current convolution neural network model. The mask branch can process the feature map and output the attention feature map with the same dimension. Then, it combines the feature maps of the two branches by using the point multiplication operation to get the final output feature map.

In the shallow structure, the network's attention is focused on the background and other areas, while in the deep structure, the network's attention feature map focuses on the object to be detected, which makes the deep feature map have higher abstraction and semantic expression ability. The structure of residual attention module is shown in Fig. 5.

As shown in Fig. 5, the attention values between 0–1 are obtained by the convolution kernel sigmoid function with the convolution kernel size of 1×1 . The attention maps of each convolution layer feature map are learned and weighted with the original feature map.

The output characteristics of the residual attention module are as follows:

$$H_{i,c}(x) = T_{i,c}(x) * M_{i,c}(x). \quad (6)$$

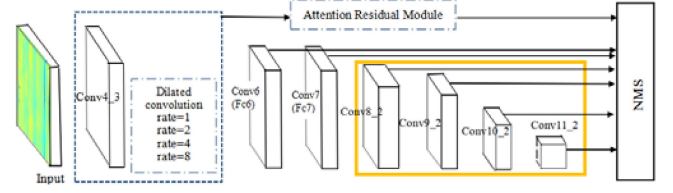


Fig. 6. SSD model structure diagram.

Among them, the output characteristic maps of the main branch are T_i , $c(x)$, the output characteristic maps of the mask branch M_i , and $c(x)$ are the values in the $[0, 1]$ interval.

In order to output the feature map with normalized weight, the mask branch needs to be followed by Sigmoid as the activation function. In mask branch, the processing of feature graph mainly includes forward downsampling and upsampling. The former is for fast encoding and acquisition of global features of feature graph, while the latter is mainly for combining the extracted global high-dimensional features after upsampling with the previously unsampled features in order to make context and high-low latitude features. They can be better grouped together.

C. Dilated attention residual (DAR)-SSD Network

The dilated convolution and attention residual modules are added to the SSD network, and a new SSD network model is obtained. The network structure is shown in Fig. 6.

From Fig. 6, we can see that the improved model adds dilated convolution to Conv4_3 layer in SSD network and changes it into multiscale dilated convolution layer. The image is convolved from dilation at 1, 2, 4 and 8, 4 scales. At the same time, the low-level features of the image are extracted using conv3_3 and conv4_3, and the conv3 layer stride is 4 and conv4_3 layer stride is 8. Then, the feature maps generated by the two convolution layers are noticed. The force residual module enhances the saliency of the target; the crop layer is added to adapt to any size of input, so that the size of the void convolved feature map is the same as that of the previous one. The two feature maps are multiplied pixel-by-pixel, and the fusion feature maps of the bottom and high-level semantic information are obtained.

For the training set m group labeled samples $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$, with $y(i) \in \{1, 2, \dots, k\}$, the weight vector $w \in R^n$ maps the eigenvector to the class label space by making the vector inner product with the input eigenvalue $x^{(i)} \in R^n$

$$f(x^{(i)}, W) = \sum_{j=1}^n w^{(j)} x_j^{(i)}. \quad (7)$$

Given the input image $I_{(i,j,k)}$, the training network f_1 parameter θ_1 , where W represents the network weight matrix, b represents the network bias vector, W_m and H_{m-1} are matrices, and the calculation formula is

$$f_1(I_{(i,j,k)}; \theta_1) = W_m H_{m-1}. \quad (8)$$

The output of the m hidden layer is calculated as

$$H_m = \max(0, (\text{pool}(W_m H_{m-1} + b_m))). \quad (9)$$

Among them, $m = 2$, $H_0 = I_{b(i,j,k)}$, b_m is the offset vector of layer m , W_m is the weight matrix. $\text{Pool}(\cdot)$ is the largest pooling operation and $\max(\cdot)$ is the activation function.

The features of network captured images are expressed as $f_2(I_{(i,j,k)}, \theta_2)$, which is combined with the output of low-level feature networks to express as follows:

$$F_c = [f_1(I_{(i,j,k)}, \theta_1), f_2(I_{(i,j,k)}, \theta_2)]. \quad (10)$$

Among them, the symbol F_c is used to represent the merged features, $I_{(i,j,k)}$ is input.

The feature F_c is transferred to the subsequent network, and then the visual feature information of multiscale feature map is effectively fused by using dilated convolution and modifying linear activation function. In multiscale networks, the output eigenvector is obtained, and the network training transforms the eigenvector F_c to the conditional probability. The conditional probability distribution of normalized prediction for each category C is calculated by using the soft max regression function

$$p(c | I_{(i,j,k)}, \theta) = \frac{e^{f_c(I_{(i,j,k)}, I_{L(i,j,k)}; \theta)}}{\sum_{b \in \{1 \dots n\}} e^{f_c(I_{(i,j,k)}, I_{L(i,j,k)}; \theta)}}. \quad (11)$$

Among them, p is conditional probability, c is category.

The training dataset is trained by the network dependency monitoring method using the minimum negative logarithmic likelihood function. The calculation formulas are as follows:

$$l_f(\theta) = - \sum \ln p(l_{(i,j,k)} | I_{(i,j,k)}, \theta). \quad (12)$$

Among them, $I_{(i,j,k)}$ is the correct class label for the position of (i, j) in image I_k . The stochastic gradient descent algorithm and inverse conduction iteration update error are used to learn the optimal parameters

$$\begin{cases} W^{(l)} = W^{(l)} - \alpha \Delta W^{(l)} \\ b^{(l)} = b^{(l)} - \alpha \Delta b^{(l)} \end{cases}. \quad (13)$$

Among them, $\Delta W^{(l)} = \frac{\partial l_f(\theta)}{\partial W}$, $\Delta b^{(l)} = \frac{\partial l_f(\theta)}{\partial b}$, and l is the number of network layers.

In SSD network, SmoothL1 loss function is used for position regression in training. The calculation formulas are as follows:

$$L_{\text{loc}}(x, l, g) = \sum_{i \in \text{Pos}} \sum_{m \in \{c, x, y, w, h\}} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m). \quad (14)$$

Among them, \hat{g}_j^m the relative offset between the real value and the regional candidate box in four dimensions (x, y, w, h) . l is the predicted offset, d is the regional candidate box. $x_{ij}^k = 1$, it indicated that the candidate box for the i region matched the j true value, otherwise is 0.

Confidence loss is a multiple category of *softmax*. The calculation formula is as follows:

$$\begin{cases} L_{\text{conf}}(x, c) = - \sum_{i \in \text{Pos}} X_{ij}^p \lg(\hat{c}_i^p) - \sum_{i \in \text{Neg}} \lg(\hat{c}_i^0) \\ \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \end{cases} \quad (15)$$

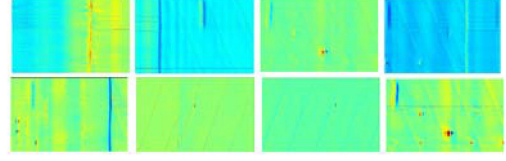


Fig. 7. Image of experimental sample.

where the default box i matches the true bounding box j of category p , it is x_{ij}^p otherwise is 0.

The objective loss function is the weight of the loss of confidence and the loss of location, so the objective loss function is obtained

$$L(x, c, l, g) = \frac{1}{N} (\alpha L_{\text{loc}}(x, l, g) + L_{\text{conf}}(x, c)). \quad (16)$$

Among them, $L_{\text{conf}}(x, c)$ is a loss of confidence, $L_{\text{loc}}(x, l, g)$ is a loss of location, and N is the default number of boxes.

IV. EXPERIMENTS AND RESULTS ANALYSIS

A. Experiment

1) *Experimental Data*: The experimental data come from the real pipeline magnetic flux leakage detection data in the actual project. The data mileage is obtained by multiplying the sampling time by the number of data points collected per unit time of the probe. A total of 2000 magnetic flux leakage image samples including 1800 circumferential welds, 600 spiral welds, 700 defects, and 500 nontarget samples are included in the dataset for experiment and test. Fig. 7 shows an image of some experimental datasets.

In the training process, the algorithm needs to calibrate the positive and negative samples of classification. The positive and negative samples are determined by the target borders of the marked picture (ground truth) and the predicted target borders. If the intersection over union (IOU) closing value of the two is 0.6, the positive sample is set, otherwise the negative sample is set.

2) *Experiments and Parameter Settings*: Caffe framework was used to train the experiment. The training model is based on the improved SSD algorithm model. The small batch stochastic gradient descent method with momentum factor is used to train the network.

The size of the input image is 560×420 , the initial learning rate is 0.01, the number of iterations is set to 60 000, and the activation function is Relu. According to the performance of graphics card and mini-batch characteristics, batch size is set to 64, 128, and 256, respectively. The initial learning rate is 0.001, the momentum factor is 0.9, the weight attenuation is 0.005, the batch size is 32, the regularization coefficient is 0.0005, the confidence threshold is 0.6, and the NMS threshold is 0.6.

3) *Evaluation Index*: The classification ability of the network model is mainly evaluated by the accuracy, recall, and specificity. Among them, TP is the correct sample for forecasting, FP is the correct sample for forecasting.

1) Accuracy, the proportion of positive samples correctly predicted to positive classes of all predictions, the

TABLE I
MODEL PARAMETER SETTING AND TEST ACCURACY

Model	No.	Coefficient expansion	Image Size	Batch size	lr	Accuracy
SSD	a	1	300*300	64	0.001	83.67%
	b	1	500*500	128	0.0001	85.26%
	c	2	300*300	64	0.001	89.69%
	d	2	500*500	128	0.0001	90.23%
	e	4	300*300	128	0.0001	93.74%
DAR-SSD	f	4	500*500	256	0.0001	95.58%
	g	8	300*300	128	0.001	92.11%
	h	8	500*500	256	0.0001	95.58%

calculation formula is as follows:

$$\text{Accuracy} = \frac{TP}{TP + FP} \quad (17)$$

2) Recall rate and correct prediction of the proportion of weld image

$$\text{TPR} = \frac{TP}{TP + FN} \quad (18)$$

3) Specificity, incorrect prediction of weld image proportions

$$\text{FPR} = \frac{TP}{TP + FN} \quad (19)$$

B. Test Results and Analysis of Improved Model

1) *Effect of Expansion Coefficient on Accuracy:* Different expansion coefficients of convolution nuclei lead to different receptive fields, and the ability of extracting initial features is the same. Table I shows the setting of model parameters and the accuracy of measurement. Expansion coefficients of dilated convolution are set to 2, 4, and 8.

From Table I, the accuracy of model e with expansion coefficient 4 is 4.05% and 1.63%, which higher than that of model c with expansion coefficient 2 and model g with expansion coefficient 8. The accuracy of model f with expansion coefficient 4 is 5.35% and 1.26%, which higher than model d with expansion coefficient 2 and model h with expansion coefficient 8. It shows that the convolution kernel with expansion coefficient 4 has the best feature extraction ability, which can extract the information of interest from the original map and remove the redundant feature information.

According to Table I, when the expansion coefficient is 4, the image size is 500×500 , the learning rate is 0.0001 and batch size is 256, the network configuration parameters are optimal, and the model detection accuracy is 95.58%.

2) *Analysis of Detection Accuracy of Improved Model:* In order to test the convergence and average recognition accuracy of the improved model, the relationship between the accuracy and iteration times of the improved model on the dataset is compared, as shown in Fig. 8. The relationship between the number of girth welds, spiral welds, small defects, and the detection accuracy in a single image is compared and analyzed, as shown in Fig. 9.

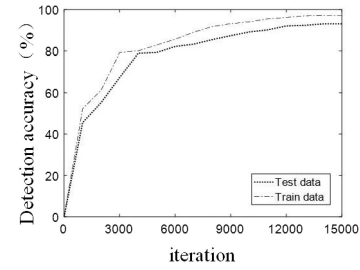


Fig. 8. Relation between detection accuracy and iteration.

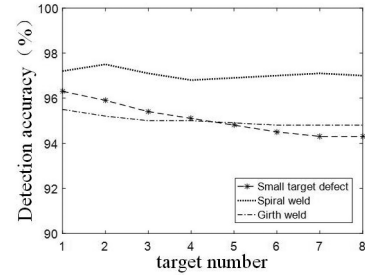


Fig. 9. Relationship between the number of targets and accuracy.

TABLE II
RESULT OF TRAINING SET LOSS VALUE AND ACCURACY COMPARISON

	SSD			DAR-SSD		
Iteration	5000	10000	15000	5000	10000	15000
Loss	1.622	1.341	1.292	0.672	0.405	0.318
Accuracy	0.521	0.693	0.727	0.750	0.893	0.953

As can be seen in Fig. 8, when the model is iterated to 15 000 times, the accuracy of datasets is higher than 90%, which shows that the improved model converges quickly and has high detection accuracy. As can be seen from Fig. 9, the less the number of target objects in a single image, the lower the probability of false detection and missed detection, the higher the detection accuracy. The number of small target defects has a great impact on the detection accuracy. when the small target defect of a single image is less than 4, the detection effect is better. So it is necessary to limit the number of sampling points in a single MFL image.

C. Capability Comparison Between Improved Model and Original Model

1) *Loss Value and Accuracy Comparison:* In order to verify that the performance of the improved model is better than that of the original model, the loss value, accuracy, and ROC curve of the two models in the training process are compared and analyzed, as shown in Table II and Fig. 10.

It can be found from Fig. 10 and Table II, when the iterations is 10 000, improve model is close to convergence, and the loss and accuracy are also better. When the iterations reaches 15 000, the loss of improve model no longer increase, and the accuracy tends to be flat. The model has the best effect when the iterations is 15 000. The SSD model can not converge well when iterating to

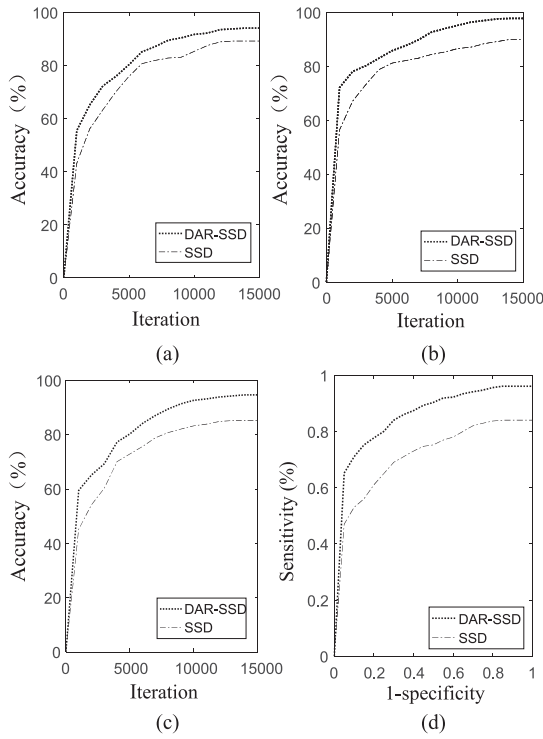


Fig. 10. Average accuracy comparison of test sets. (a) Spiral weld. (b) Girth weld. (c) Small target defect. (d) ROC.

15 000 times, which shows that the convergence of the improved model is better than that of SSD. Therefore, when the iterations is 15 000, the detection effect is the best. The detection accuracy of spiral weld, girth weld, and small defect is 93.2%, 97.6%, and 90.17%.

2) Comparison of Detection Accuracy: In order to verify the actual prediction effect of the optimal model in this paper, two single target and two multitarget images were randomly selected input the improved model and SSD model, the detection results in Figs. 11 and 12. Fig. 11 shows a single target detection result image with sampling points of 256×3200 , and Fig. 12 shows a multitarget detection result image.

In Fig. 11, it can be seen from the figure that the prediction accuracy of small defects in the improved model can reach more than 96%, and that in the original model, the prediction accuracy of small defects is 85%. The prediction of girth weld in the improved model can also reach 99%, but only 89% in the original model. By comparing the results, the accuracy of the improved model is higher than that of the original model.

In Fig. 12, two multiobjective image (a) and (b) are input into the improved model and SSD model, and the image (b)–(f) are detection results. It can be seen in Fig. 11(b) and (c), the predicted value of SSD model for small defect 1 is only 0.54. Small defect 2 and girth weld 1 are missed, and small defect 3 is misdiagnosed as girth weld. The predicted value of small defects in the improved model is higher than that in the SSD model. In Fig. 11(e) and (d), the SSD missed the detection of spiral weld 2, while the predicted accuracy of the improved model is higher

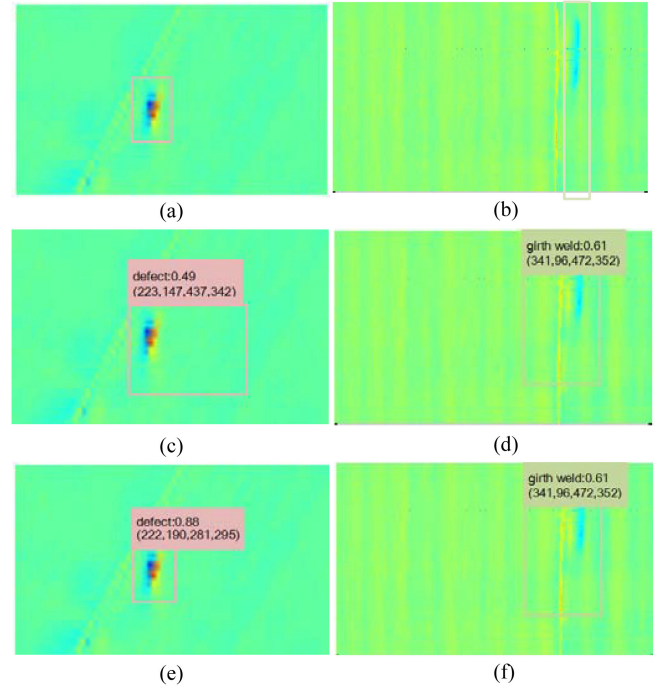


Fig. 11. Detection result of MFL image of single target. (a) Label image 1. (b) SSD detection result. (c) DAR-SSD detection result. (d) Label image 2. (e) SSD detection result. (f) DAR-SSD detection result.

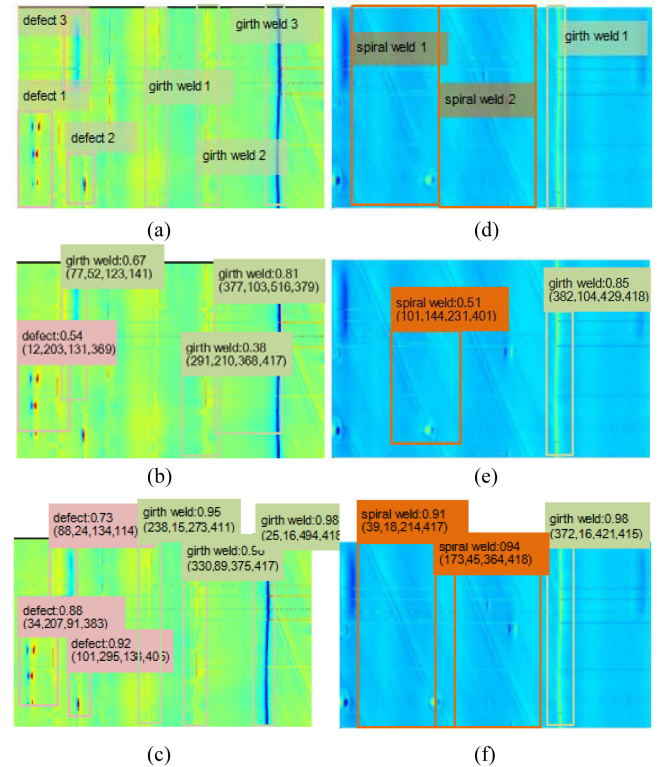


Fig. 12. Detection result of MFL image of multiple target. (a) Label image 3. (b) SSD detection result. (c) DAR-SSD detection result. (d) Label image 4. (e) SSD detection result. (f) DAR-SSD detection result.

TABLE III
MISTAKE DETECTION RATE AND MISSED DETECTION RATE

Model	Mistake detection rate (%)			Missed detection rate(%)		
	Spiral weld	Girth weld	Small defect	Spiral weld	Girth weld	Small defect
SSD	5.63	8.27	12.35	3.17	8.44	15.92
DAR-SSD	1.18	2.52	3.19	0.52	2.89	5.76

TABLE IV
IOU CONTRAST RESULTS

Model	IOU/Accuracy(%)					
	0.6	0.65	0.7	0.75	0.8	0.85
SSD	89.03	88.73	88.05	87.92	87.63	87.21
DAR-SSD	97.92	97.86	96.15	96.17	96.95	96.67

than 90%. By comparing the results of the two models, it can be found that the detection accuracy of multitarget images is lower than that of single-target images, and the detection accuracy of weld seam is higher than that of small defects. The detection ability of improved multitarget image model is higher than that of SSD model, especially for small target defect detection.

In order to improve the accuracy of the experimental results, 1000 MFL images were selected to input into the two models for verification, the detection results in **Table III**.

From **Table III**, it can be seen that the false detection rate and missed detection rate of the improved model are 1.18% and 0.52%, respectively, which are 4.45% and 2.65% less than those of the original model. The corresponding false detection rate and missed detection rate of spiral weld are 5.75% and 5.55% less than those of the original model, and the corresponding false detection rate and missed detection rate of minor defects are 9.16% and 10.16%. It can be seen that the improved model greatly reduces the false detection rate and missed detection rate, and greatly improves the detection ability of small target defects.

In the process of model evaluation, the average detection accuracy is calculated under different IOU conditions, and the results are shown in **Table IV**.

Table IV shows that under different IOU conditions, the accuracy of the improved model is always higher than that of the SSD. When IOU is raised to 0.85, the detection accuracy of the improved model is 9.46%, higher than that of the SSD. It shows that the improved model further learns the target's high semantic features, and thus, effectively generates high-quality boundary frames. Therefore, the improved model improves the low-detection accuracy of the single-stage detection model for small targets to a certain extent.

V. CONCLUSION

In this paper, MFL image detection algorithm based on deeplearning was proposed. Dilated convolution and attention residual module were added to the SSD model, which construct a multiscale feature extraction model with the multilevel semantic information, so as to improve the detection accuracy of small

targets. The improved model enlarged the receptive field of the convolution nucleus by hollow convolution, but did not increase the computational complexity. Through the attention residual module, more features of small target regions of interest were learned, thus improving the positioning confidence and accuracy.

The experimental results showed that the improved model can achieve more than 90% accuracy in detecting defects, better than the original model in detecting small target defects. The method based on deep learning proposed in this paper avoids the error caused by artificial feature selection, greatly reduces the false detection rate, and leakage detection rate.

REFERENCES

- [1] Z. C. Qiu *et al.*, "Quantitative identification technology of micro-crack magnetic flux leakage based on GA-BP neural network," *J. Beijing Univ. Technol.*, vol. 36, no. 1, pp. 1101–1104, 2016.
- [2] K. Tsukada *et al.*, "Detection of inner cracks in thick steel plates using unsaturated AC magnetic flux leakage testing with a magnetic resistance gradiometer," *IEEE Trans. Magn.*, vol. 53, no. 11, Nov. 2017, Art. no. 2501305.
- [3] J. H. Liu, M. Fu, F. Liu, J. Feng, and K. Cui, "Window feature-based two-stage defect identification using magnetic flux leakage measurements," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 1, pp. 12–23, Jan. 2018.
- [4] G. Kopp and H. Willems, "Sizing limits of metal loss anomalies using tri-axial MFL measurements: A model study," *NDT E Int.*, vol. 55, pp. 75–81, 2013.
- [5] Y. Cheng, L. Bai, F. Yang, Y. Chen, S. Jiang, and C. Yin, "Stainless steel weld defect detection using pulsed inductive thermography," *IEEE Trans. Appl. Supercond.*, vol. 26, no. 7, Oct. 2016, Art. no. 0606504.
- [6] J. Chen, S. Huang, and W. Zhao, "Three dimensional defect inversion from magnetic flux Terhnnl leakage signals using iterative neural network," *IET Sri. Meas. Technol.*, vol. 9, no. 4, pp. 418–426, 2015.
- [7] H. Wang *et al.*, "A clustering approach for assessing external corrosion in a buried pipeline based on hidden Markov random field model," *Structural Safety*, vol. 56, pp. 18–29, 2015.
- [8] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Con. Comput. Vision*, 2016, pp. 21–37.
- [9] J. Feng, F. Li, S. Lu, J. Liu, and D. Ma, "Injurious or noninjurious defect identification from MFL images in pipeline inspect-ion using convolutional neural network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1883–1892, Jul. 2017.
- [10] S. Inkyu *et al.*, "WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming," *Remote Sens.*, vol. 10, no. 9, pp.1423–1448, 2018.
- [11] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "NetVLAD: CNN architecture for weakly supervised place recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 6, pp. 1437–1451, Jun. 2018.
- [12] C. Y. Yu, X. D. Xu, and S. J. Zhong, "An improved SSD model for saliency target detection," *J. Electron. Inf.*, vol. 11, no. 3, pp. 2554–2561, 2018.
- [13] C. Tang *et al.*, "Visual tracking method based on deep learning object detection," *Infrared Laser Eng.*, vol. 47, pp. 148–158, 2018.
- [14] T. Tang, S. L. Zhou, and Z. P. Deng, "Arbitrary-oriented vehicle detection in aerial imagery with single convolutional neural networks," *Remote Sens.*, vol. 9, pp. 1–17, 2017.
- [15] Z. Liu, Z. Wu, T. Li, J. Li, and C. Shen, "GMM and CNN hybrid method for short utterance speaker recognition," *IEEE Trans. Ind. Inform.*, vol. 14, no. 7, pp. 3244–3252, Jul. 2018.
- [16] H. B. Gao, B. Cheng, J. Wang, K. Li, J. Zhao, and D. Li, "Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment," *IEEE Trans. Ind. Inform.*, vol. 14, no. 9, pp. 4224–4231, Sep. 2018.
- [17] X. Yang, X. B. Gao, B. Song, and D. Yang, "Aurora image search with contextual CNN feature," *Neurocomputing*, vol. 281, pp. 67–77, 2018.
- [18] J. Wang, H. Zhu, S. Y. Yu, and C. X. Fan, "Object tracking using color-feature guided network generalization and tailored feature fusion," *Neurocomputing*, vol. 2017, pp. 387–398, 2017.

- [19] H. Liu, B. Lang, M. Liu, and H. B. Yan, "CNN and RNN based payload classification methods for attack detection," *Knowl. Based Syst.*, vol. 163, pp. 332–341, 2019.
- [20] P. Li, Z. Chen, L. T. Yang, Q. Zhang, and M. Jamal Deen, "Deep convolutional computation model for feature learning on big data in internet of things," *IEEE Trans. Ind. Inform.*, vol. 14, no. 2, pp. 790–798, Feb. 2018.
- [21] H. M. Gao *et al.*, "Multi-branch fusion network for hyperspectral image classification," *Knowl. Based Syst.*, vol. 167, pp. 11–25, 2019.
- [22] R. Sakurai, S. Yamane, and J.-H. Lee, "Restoring aspect ratio distortion of natural images with convolutional neural network," *IEEE Trans. Ind. Inform.*, vol. 15, no. 1, pp. 3244–3252, Jan. 2019.
- [23] Z. H. Wang, X. X. Wang, and G. Wang, "Learning fine-grained features via a CNN tree for large-scale classification," *Neurocomputing*, vol. 275, pp. 1231–1240, 2018.
- [24] S. H. Wang, J. W. Xiang, and Y. T. Zhong, "Convolutional neural network based hidden Markov models for rolling element bearing fault identification," *Knowl. Based Syst.*, vol. 144, pp. 65–76, 2018.
- [25] H. J. Xu, C. Q. Huang, and D. H. Wang, "Enhancing semantic image retrieval with limited labeled examples via deep learning," *Knowl. Based Syst.*, vol. 163, pp. 252–266, 2019.
- [26] W. Liu, Z. M. Luo, and S. Li, "Improving deep ensemble vehicle classification by using selected adversarial samples," *Knowl. Based Syst.*, vol. 160, pp. 167–175, 2018.
- [27] Z. C. Wang, L. Du, J. Mao, B. Liu, and D. Yang, "SAR target detection based on SSD with data augmentation and transfer learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 150–154, Jan. 2019.
- [28] S. Akcay, M. E. Kundegorski, C. G. Willcocks, and T. P. Breckon, "Using deep convolutional neural network architectures for object classification and detection within X-Ray baggage security imagery," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 9, pp. 2203–2215, Sep. 2018.
- [29] Z. B. Zheng, Y. Y. Yang, X. D. Niu, H.-N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Trans. Ind. Inform.*, vol. 14, no. 4, pp. 1606–1615, Apr. 2018.
- [30] Z. H. Cui, F. Xue, X. Cai, Y. Cao, G.-G. Wang, and J. Chen, "Chendetection of malicious code variants based on deep learning," *IEEE Trans. Ind. Inform.*, vol. 14, no. 7, pp. 3187–3196, Jul. 2018.



Lijian Yang received the B.S. degree from the Shenyang Institute of Mechanical and Electrical Technology, China, in 1981, and the M.S. degree from the Harbin Institute of Technology, China, in 1984.

He is currently a Professor with the School of Information Science and Engineering, Instrument Science and Technology and Doctoral Supervisor with the Shenyang University of Technology, Shenyang, China. His main research interests include nondestructive testing detection technology and theory of long oil and gas pipeline.



Zhujuan Wang received the B.S. and M.S. degrees in electronic and information engineering from Shenyang Aerospace University, Shenyang, China, in 2013 and 2016, respectively. She is currently working toward the Ph.D. degree in instrumentation science and technology with the Shenyang University of Technology, Shenyang.

Her research interests include nondestructive testing of pipelines.



Songwei Gao received the B.S. degree from the Shenyang Institute of Mechanical and Electrical Technology, China, in 1982, and the M.S. degree from the Shenyang University of Technology, China, in 1999.

She is a Professor with the School of Information Science and Engineering and Doctoral Supervisor with the Shenyang University of Technology, Shenyang, China. Her main research interests include nondestructive testing detection technology and theory of long oil and gas pipeline.