IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 2, 2019, accepted July 3, 2019, date of publication July 10, 2019, date of current version September 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927258

Automatic Welding Defect Detection of X-Ray Images by Using Cascade AdaBoost With Penalty Term

FENG DUAN[1], (Member, IEEE), SHIFAN YIN[1], PEIPEI SONG[2], WENKAI ZHANG[1], CHI ZHU[3], (Member, IEEE), AND HIROSHI YOKOI[1,4], (Member, IEEE)

[1]College of Artificial Intelligence, Nankai University, Tianjin 300350, China
[2]College of Engineering and Computer Science, The Australian National University, Canberra, ACT 2600, Australia
[3]Department of System Life Engineering, Maebashi Institute of Technology, Gunma 371-0816, Japan
[4]Department of Mechanics and Intelligence, The University of Electro-Communications, Tokyo 182-8585, Japan

Corresponding author: Feng Duan (duanf@nankai.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFE0129700, in part by the National Natural Science Foundation of China under Grant 61673224, and in part by the Tianjin Natural Science Foundation for Distinguished Young Scholars under Grant 18JCJQJC46100.

ABSTRACT Welding defect constitute a great danger to the safe usage of petroleum pipelines. In addition, manual welding defect detection has numerous deficiencies, such as subjectivity and inaccurate estimation of geometric parameters. Thus, we proposed an automatic welding defect detection system for X-ray images. In the system, five typical defects (cracks, lack of penetration, lack of fusion, round defects, and stripy defects) and non-defects were chosen for recognition. There are three stages in the system: defect extraction, defect detection, and defect recognition. In the first stage, background subtraction with an adaptive thresholding method was adopted to identify the potential defects. In the second stage, to extract real defects from the massive number of potential defects, the adaptive cascade boosting (AdaBoost) algorithm was employed for binary classification. Grayscale features and geometric properties of the defects were extracted for the classification. In the third stage, the AdaBoost algorithm was extended for multi-classification. In the process of distinguishing defects from non-defects, a high detection rate is necessary. To ensure the high true positive rate (TPR) and the low false positive rate (FPR), we proposed the cascade AdaBoost algorithm with penalty term. The accuracy of the defect detection was 85.5%, and the TPR was 91.66%. Moreover, three comparison tests of support vector machine (SVM), k-nearest neighbor (KNN), and random forest (RF) were employed to validate the superiority of AdaBoost. The experimental results indicate that the proposed detection system can be effective for defect detection.

INDEX TERMS AdaBoost, image processing, X-ray image, nondestructive testing, welding defects.

I. INTRODUCTION

Nondestructive testing is a widely used and important technique for defect detection in industry. According to statistics from the Central Intelligence Agency, there are a total of slightly less than 3.5 million km of pipelines in 120 countries globally in 2014 [1]. Currently, welding defect inspection is mainly performed through manual work. However, manual inspection can lack objectivity, consistency, accuracy, and efficiency. The evaluation results also tend to differ among operators of different levels. In addition, the evaluation work is typically subjective and time-consuming. Moreover, a novice operator usually lacks the experience necessary to make an accurate assessment of welding defects. The evaluation process is also mainly conducted by human eye observation. Thus, the operator cannot obtain precise geometric parameters, such as length, area, and density. These factors constitute major obstacles to making accurate assessments of welding quality. However, with the development of X-ray imaging techniques, inspection based on X-ray images is extensively adopted for the inspection of welding defects of petroleum pipelines. Automatic detection based on X-ray

The associate editor coordinating the review of this article and approving it for publication was Haiyong Zheng.

images offers major advantages over manual work in terms of labor consumption, time consumption, digitalized archiving, and objectivity. Therefore, this paper proposed an automatic welding defect detection system to replace manual inspection and improve detection performance.

In recent years, many researchers have focused on automatic inspection techniques based on X-ray images. It has been shown that X-ray images usually have high noise and low contrast, which introduces difficulty in image processing [2].

Regarding detecting welding defects, there are a variety of image processing methods. The background subtraction method [3]–[6], the wavelet technique [6], [7], and the texture feature based method [9]–[11] have been commonly used among a large number of detection methods. Background subtraction is used to dissociate defects and backgrounds in [6]. Li *et al.* [7] proposed the wavelet technique to detect three kinds of defects, including air-hole, foreign-object, and shrinkage-cavity defects. Li compared the three methods: second-order derivative and morphology operations, row-by-row adaptive thresholding, and wavelet transform technique in detecting defects and found that the third method can be effective for the three defects by testing 21 defects. Ngan *et al.* [9] presented a novel approach to detect welding defects based on texture features. They extracted the texture features from the welding defect regions. Since the extracted texture features have high dimensions and are partially redundant with each other, principal component analysis (PCA) was used to reduce the dimensions of these features.

Machine learning algorithms are widely utilized in distinguishing defects from non-defects. There are SVM [12], [13], gentle AdaBoost [15], Gaussian mixture model [16] neural network [17] and deep learning [18] for the defect detection. Shao used an SVM classifier to extract defects with three features. Haddad [12] firstly used the sparse coding for the defect detection.

As for multi-classification of defect types, multivariate statistical methods and machine learning algorithms are mainly employed. Jiang *et al.* [10] adopted three classifiers: SVM, neural network, and k-NN (k-nearest neighbor) to classify six defects (worm holes, porosity, linear slag inclusion, gas pores, lack of fusion, and crack) and non-defects. The neural network and SVM were proven to be better than k-NN. Hou *et al.* [19] proposed an automatic system to detect and recognize welding defects in radiographic images. An artificial neural network (ANN) and an adaptive-network-based fuzzy inference system (ANFIS) were used to classify four defects and non-defects. Zapata *et al.* [20] built a data-driven model that deals with four defects by applying multivariate statistics and machine learning methods. Moreover, they adopted wavelet packet decomposition principal component analysis (WPD-PCA) to extract features. Feedforward neural network (FNN) and SVM were also employed to estimate the welding status and classify defects. Zapata *et al.* [18] presented a model based on deep neural network used for the automatic detection of welding defects. Although the model can obtain a great classification accuracy rate of 91.84%, specific types of welding defects are not distinguished.

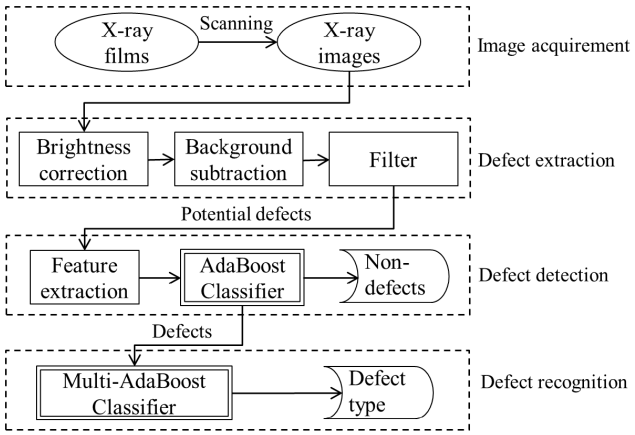


FIGURE 1. Architecture of the automatic welding defects detection system.

Taking advantage of the findings of previous studies, we proposed an automatic welding defect detection system. Defects are mixed with a large number of non-defects when using only image processing methods. This may result in a high rate of missing report, which is prohibited in industrial inspection. Therefore, it is critical to extract defects from non-defects by combining image processing methods with machine learning methods. During the welding defects detection process, detecting the welding defects is much more significant than the classification of welding defects. Therefore, we proposed the cascade AdaBoost [21] algorithm with penalty term to improve the detection true positive rate (TPR) of welding defects. The extended multi-class AdaBoost classifier was utilized to identify five kinds of welding defects.

The remainder of this paper is organized as follows. Section 2 presents an overview of the detection system. Section 3 briefly introduces the process of defect extraction. Section 4 explains the features selected in this paper. Section 5 introduces the proposed cascade AdaBoost algorithm and the extended multi-class AdaBoost classifier. Section 6 describes the experimental results and gives the analyses. Finally, conclusions are given in Section 7.

II. OVERVIEW OF THE PROPOSED SYSTEM

The automatic welding defect detection system deals with the detection and classification of five defects. The system architecture is shown in Fig. 1.

The system is divided into four stages, the image acquirement, defect extraction, defect detection, and defect recognition. In the defect extraction stage, we generated the background image. In addition, an adaptive thresholding method was applied to obtain the potential defects. Due to low contrast and high noise, many non-defects were detected. To reduce false alarms, defect detection was intended to remove non-defects. Some comparative features were computed by using the horizontal region comparison method.

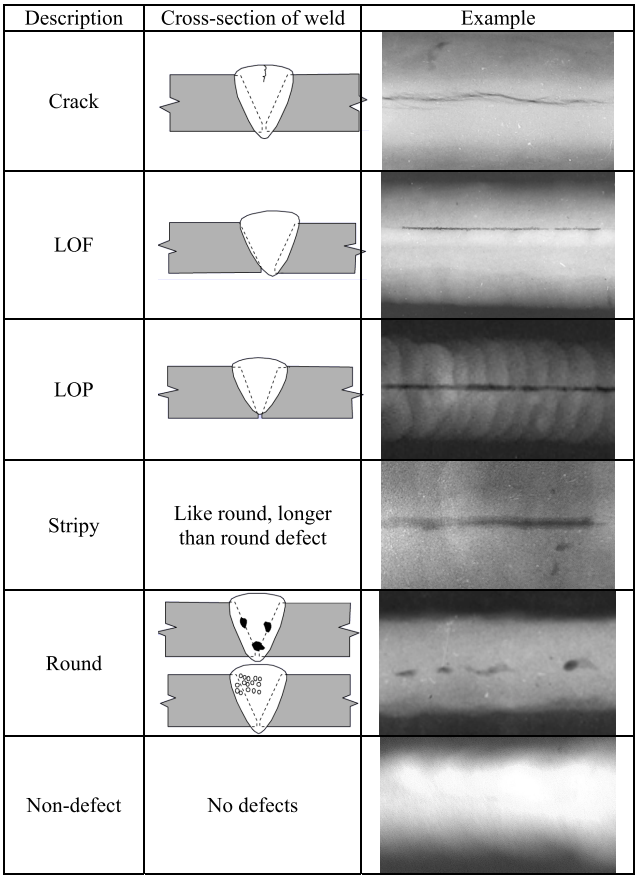


FIGURE 2. Examples of the five defects and non-defects.

Then, we proposed the cascade Adaboost classifier with penalty term distinguished defects from non-defects. Finally, the real defects were recognized by using the extended multi-class AdaBoost classifier.

There are a variety of types of welding defects. In this paper, we focus on five typical defects, which are cracks, lack of fusion (LOF), lack of penetration (LOP), stripy defects, and round defects. As shown in Fig. 2, cracks are usually low-contrast, and they have no definite shape. In addition, most of them have thin branches at the end. Generally, LOP is located in the center of the weld, and LOF is located in the edge of the weld. LOF and LOP are usually straight lines, and they differ in intensity distribution and positions in the weld. Round defects (including air holes and slag inclusions) are circles or ellipses. Stripy defects have an aspect ratio greater than 3. And stripy defects are uniform in intensity. LOF, LOP, and stripy defects possess numerous similarities in terms of grayscale and shape. Therefore, it is difficult to classify the three defects. Moreover, round defects are easier to detect because of their shape and strong contrast.

III. AUTOMATIC DEFECTS EXTRACTION

This section focuses on defects segmentation in X-ray images. This paper adopted the background subtraction method to extract potential defects. An important factor in defect segmentation is noise. Thus, this section includes three sub-sections: pre-processing, background subtraction method, and potential defect extraction.

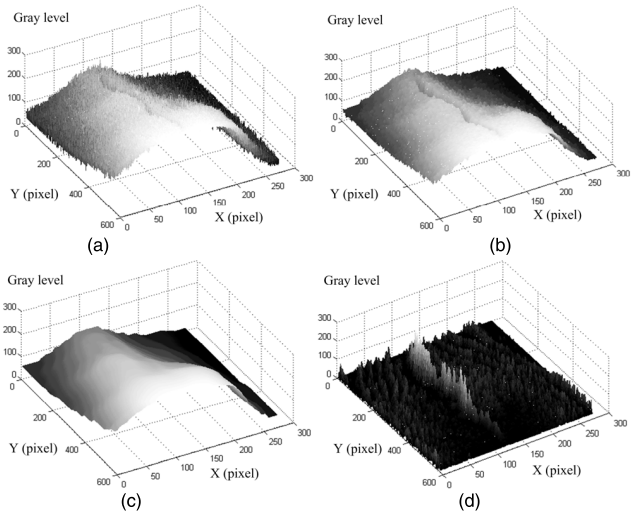


FIGURE 3. 3D graph of image processing. (a) Image after histogram equalization. (b) Smoothed image. (c) Modelling background. (d) Difference image.

A. PRE-PROCESSING

Pre-processing of X-ray images intends to eliminate noise and enhance contrast. We used histogram equalization to improve intensity distribution, as shown in Fig. 4(b). Quantum noise and electronic noise exist in X-ray images. Hence, filtering is an important component of image de-noising. The median filter can preserve the edges and also reduce impulse noise. The average filter can remove the pulse and random disturbances. Since we need to remove the noise before extracting potential defects, we used the median filter and the average filter [6] for image smoothing.

In this paper, a 3×3 median filter template and a 3×3 average filter template were utilized for image filtering. It can be seen from Fig. 3(b) that the smoothing operation reduced the spot noise. In addition, Fig. 3(b) retains defect information for further processing.

B. BACKGROUND SUBTRACTION METHOD

Background subtraction [3], [6] is one of the commonly used methods for the automatic defect segmentation of X-ray images. The critical part of this method is background modelling. Compared with the welding image, the defect is much smaller. Therefore, when building the background image, the surrounding pixels are considered. In this paper, we utilized a large template median filter for background estimation. The median filter focuses on the median grayscale of the filter window. As shown in Fig. 4(a), the defect region is darker and smaller than its surrounding pixels. We used the median filter to build the background model. The background image is shown in Fig. 4(c).

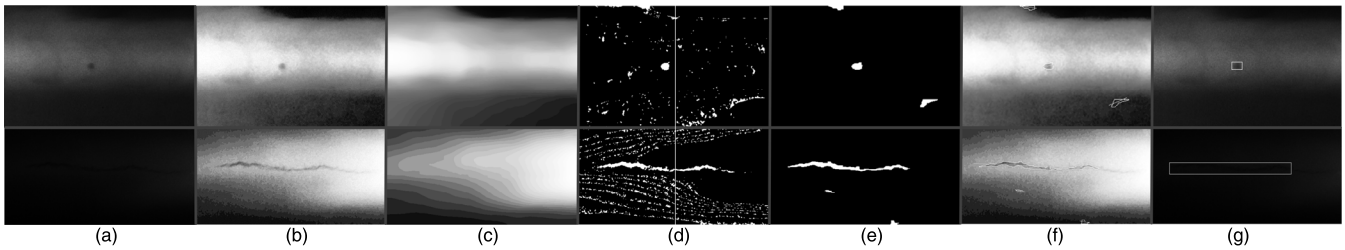


FIGURE 4. Illustrations of the defect detection process. Top row: an example of the round defect detection process. Bottom row: an example of the crack defect detection process. (a) Original image. (b) Histogram equalization. (c) Modelling background. (d) Binary image after the background subtraction method. (e) Potential defects obtained after area filter and morphological operation. (f) Edges of the potential defects and the major direction of each defect region. (g) The defects labeled by the white rectangle are extracted by using AdaBoost classifier. (Text information has been automatically removed).

In (1), the median filter template's size h_w was determined by the image resolution. $f(i, j)$ is the smoothed image, and $b(i, j)$ is the modeled background. In our experiment, h_w was set as 30 with an image resolution of 4000 × 1024. The difference image $f_s(i, j)$ can be computed in (2), which focuses on the darker part in the $f(i, j)$. The difference image is presented in Fig. 3(d).

$$b(i,j) = median(f(i+x, j+y), \quad x, y \in [-h_w, h_w]) \tag{1}$$

$$f_s(i,j) = \begin{cases} |f(i,j) - b(i,j)|, & if\ f(i,j) - b(i,j) < 0 \\ 0, & otherwise \end{cases} \tag{2}$$

To enhance the contrast of the image fs, we used linear gray stretch to extend the gray range. An adaptive threshold value was computed with the average grayscale g_{avg} and the standard deviation g_{std}. A binary image was obtained with the adaptive threshold T, as shown in Fig. 4(d). In addition, the adjusting parameter p can adjust the sensitivity of the defect's detection. p was set as three in the experiment.

$$T = g_{avg} + p \times g_{std} \tag{3}$$

C. POTENTIAL DEFECT EXTRACTION

The binary image contains some noise and edges that are not defects, which requires removal of false alarms. Noise is usually small in comparison to defects. Therefore, an area filter based on the connected region was used to remove the noise. The erosion operator was also applied to reduce the noise. Moreover, the dilate operator was utilized to enhance the defect's connectivity. Finally, we obtained the potential defect regions, as illustrated in Fig. 4(e). In the segmentation process, one defect may be divided into two or more defects. Therefore, if the distance in the weld direction between two defects (other than round defects) is shorter than a threshold, they will be merged into one.

IV. FEATURE EXTRACTION

The classifier is one of the most important components of the proposed automatic welding defect detection system. In this regard, feature selection is critical to classification performance. The features focus on two main aspects: the grayscale features and the geometric properties. By observing the X-ray images, one can see that the defect is relatively distinct from its neighborhood area. Because the weld seam is along the horizontal direction, the intensity is uniformly distributed in the horizontal direction. Thus, we adopted the horizontal region comparison method to calculate the gray difference.

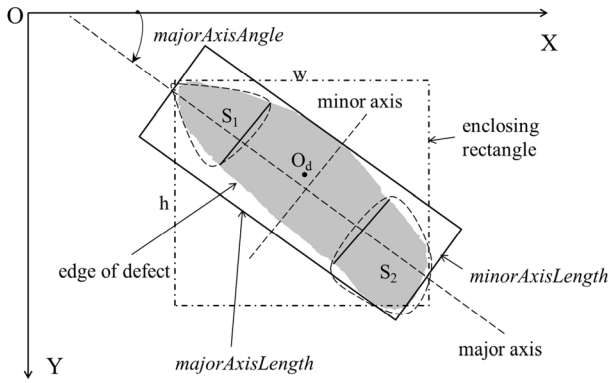


FIGURE 5. Illustration of the geometric parameters of the defect.

In X-ray images of pipelines, we can extract defects from potential defects by using some characteristics that describe the difference between the defect and its comparison region. In the experiment, we used two horizontal comparison regions: one is on its left and the other is on its right. The interval was set as 20 pixels. We put forward some features based on the gray difference between the potential defect region and the comparison region.

Taking advantage of the results of previous studies [22], [23], we extracted 61 features in the feature extraction step (refer to Fig. 8 for details). Fig. 5 depicts the definitions of some features. This section explains a part of the features' definitions.

Feature 44: *reg*

The parameter *reg* indicates the relative grayscale of the defect region. *grayAvg* is defined as the average gray value of the defect's region. *grayAvgrec* is the average gray value of the rest of the extended rectangle. The extended rectangle is the enclosing rectangle extended outward five pixels.

$$reg = grayAvg/grayAvg_{rec} \tag{4}$$

Feature 46: *tightness*

Tightness presents the property of the shape. According to (5), the *tightness* of circles is the smallest among all 2D shapes. The *area* is the sum of pixels in the defect region. The *perimeter* is the number of pixels that belong to the boundary of the defect region.

$$tightness = perimeter^2/area \tag{5}$$

Feature 47 and feature 60: *vDisToCenter* and *hDisToCenter*

These two parameters indicate the relative position of the defect in the weld. As shown in Fig. 5, O_d is the gravity center of the defect region. Suppose that O_c is the center of the weld in the image. w_h is one half of the weld's length, and w_v is one half of the weld's width. d_h is the distance from O_d to O_c in the horizontal direction, and d_v is the distance from O_d to O_c in the vertical direction. *vDisToCenter* and *hDisToCenter* are computed as (6) and (7), respectively.

$$vDisToCenter = d_v/w_v \tag{6}$$

$$hDisToCenter = d_h/w_h \tag{7}$$

Feature 48: *majorAxisAngle*

The major axis angle is defined as the angle between the major axis and the x axis. The major axis indicates the main direction of the defect region in Fig. 5. The principle component analyzation method (PCA) was applied to compute the major direction of the defect. The minor direction is perpendicular to the major direction.

Feature 49 and Feature 50: *majorAxisLength* and *minorAxis Length*

The defect is projected to the major axis, and the projected length is the major axis length. The same is true for the minor axis length. The projection computation used radon transform.

Feature 51: *elongation*

elongation is the aspect ratio of the defect. It is an important variable to distinguish round defects between stripy defects.

$$elongation = majorAxisLength/\min \text{or } AxisLength \tag{8}$$

Feature 52: *degOfFill*

degOfFill refers to the degree of filling. It is also a characteristic of the shape. It is defined as (9).

$$\deg OfFill = area/(w \times h) \tag{9}$$

Feature 53: *flatness*

flatness is a value that indicates the smoothness of the defect's boundary and the defect's surface. $I(x, y)$ means the boundary of the defect region. (x_{avg}, y_{avg}) is the gravity center of the defect region.

$$\begin{cases} IP = \int\limits_{(x,y)\in Defect} I(x,y) \times \sqrt{(x-x_{avg})^2 + (y-y_{avg})^2}dxdy \\ flatness = \dfrac{IP^2}{area^2} \end{cases} \tag{10}$$

Feature 54: *sharpness*

One-fourth of the area is denoted as S1, and the other one-fourth of the area is denoted as S2 along the major axis in Fig. 5. The feature *sharpness* is defined in (11).

$$sharpness = (S_1 + S_2)/area \tag{11}$$

Feature 55: *sym*

sym describes the symmetry of the defect.

$$sym = |1 - S_1/S_2| \tag{12}$$

Feature 56: *bias*

bias is defined as the normalized distance from the gravity center (gc_x, gc_y) to the geometric center of the defect region (c_x, c_y).

$$bias = \frac{2 * \sqrt{(c_x - gc_x)^2 + (c_y - gc_y)^2}}{\sqrt{w^2 + h^2}} \tag{13}$$

Feature 57: *mFlatness*

Through radon transform, we obtain the projection at the major axis. *mFlatness* is defined as the standard deviation of the projection. It can indicate the smoothness in the vertical direction.

Feature 59: *vIntensityRatio*

It describes the gray distribution in the vertical direction. $g_{avg}(up)$ is the average gray value of the upper half. $g_{avg}(down)$ is the average gray value of the bottom half.

$$vIntensityRatio = \left|1 - \frac{g_{avg}(up)}{g_{avg}(down)}\right| \tag{14}$$

Feature 61: *moment*

The central *moment* is a common parameter to describe the contour's property.

$$moment = \sum_{i=1}^{n} (x_i - x_{avg})^2 (y_i - y_{avg})^2 \tag{15}$$

The feature vector is utilized to train classifiers for classification of defects, non-defects, and defect types.

V. CLASSIFICATION

According to the actual industrial application requirements, detecting the welding defects is much more significant than classifying the types of welding defects. Therefore, we proposed the cascade AdaBoost algorithm with penalty term to improve the detection TPR of welding defects. The extended multi-class AdaBoost classifier was utilized to identify five types of welding defects.

A. CASCADE ADABOOST ALGORITHM WITH A PENALTY TERM

In 1995, Freund and Schapire proposed the AdaBoost algorithm [21]. Freund and Schapire [22] explained that boosting is a general method for improving the accuracy of any given learning algorithm, and it does not suffer from overfitting.

The weak learner is the basic unit of the AdaBoost classifier. Kearns and Valiant [24] was the first to propose the

Given: $(x_1, y_1), \ldots, (x_m, y_m)$ where $x_i \in X, y_i \in Y = \{-1, +1\}$

Initialize $D_1(i) = 1/m_p, y_i = 1; D_1(i) = 1/m_n, y_i = -1$

For $t = 1, \ldots, T$:

- Train weak learner using distribution D_t.
- Get weak hypothesis $h_t: X \to \{-1, +1\}$ with error

$$\varepsilon_t = P_{ri \sim D_t}[h_t(x_i) \neq y_i].$$

- Choose:

$$\alpha_t = \frac{1}{2} \ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$$

- Update:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t}, & if\ h_t(x_i) = y_i \\ p \times e^{\alpha_t}, & if\ h_t(x_i) = -1, y_i = 1 \\ e^{\alpha_t}, & if\ h_t(x_i) = 1, y_i = -1 \end{cases}$$

where Z_t is a normalization factor (chosen so that D_{t+1} will be a distribution).

Output the final hypothesis:

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

FIGURE 6. The boosting algorithm AdaBoost with penalty term.

question of whether a "weak" learner can be "boosted" into an arbitrarily accurate "strong" learning algorithm. A "weak" learner is a classifier that performs just slightly better than random guessing. The AdaBoost algorithm provides the rules according to which weak learners are boosted into a strong classifier. In our experiment, we focus on two key points. One is sample distribution, and the other is each weak learner's weight. The original algorithm was introduced by You *et al.* [21]. The AdaBoost constructs a classifier by increasing the weight of the sample classified incorrectly. Mekhalfa and Nacereddine [15] used the Gentle Adaboost algorithm to classify four types of welding defects. In this paper, we have improved the Adaboost algorithm to improve the detection TPR of welding defects, as shown in Fig. 6. Moreover, five types of welding defects were recognized by the extended multi-class AdaBoost classifier.

We proposed the AdaBoost algorithm with penalty term to improve the TPR. The penalty term p is added to the classifier construction when a real defect is estimated as a non-defect. Consequently, the boosted classifier prefers the samples of the defect. This penalty term will improve the TPR by changing the sample distribution.

The weak learner that we chose is the binary tree, which deals with two-group classification problems. The weak leaner is presented in (16). x is the feature, $f(x)$ is the value of the feature, pr indicates the polar, and θ is the threshold. For each feature, the corresponding learner is trained to achieve the minimum classification error rate.

$$h(x, f, p, \theta) = \begin{cases} -1, & pr \times f(x) < p\theta \\ 1, & otherwise \end{cases} \quad (16)$$

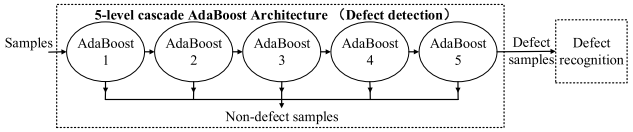


FIGURE 7. 5-level cascade AdaBoost architecture for defect detection.

The AdaBoost algorithm formed a powerful classifier by combining multiple weak learners. In addition, we can adjust the TPR by setting the penalty term p. Fig. 4(g) presents the defects estimated by the AdaBoost classifier.

The method of combining multiple strong classifiers is called a "cascade". The architecture of cascade Adaboost classifier is shown in Fig. 7. The structure of a cascaded classifier is essentially that of a degenerate decision tree. The setting of cascades can improve the performance of the classifier by increasing the TPR and decreasing the false positive rate (FPR). Suppose that more than half of the non-defects are rejected in each stage, and almost all of the real defects pass to the next stage. We can then obtain a good result after a number of stages. The number of stages in the experiment is set as five.

For each cascaded classifier, the increasing of TPR will always cause the increasing of FPR. Balancing these two parameters is important for the cascaded classifier. The tuning process of p is shown in (17). The parameter λ is set to adjust the TPR and the FPR, and the variable d is the step length.

$$p = p + d \times \lambda \times (1.0 - TP) + d \times (1.0 - \lambda) \times (0.5 - FP) \quad (17)$$

B. CLASSIFICATION FOR FIVE DEFECTS USING MULTI-CLASS ADABOOST

AdaBoost algorithm can effectively classify two classes, hence, the cascade AdaBoost algorithm with a penalty term deals with the detection of welding defects from potential defects.

However, after the detection process of welding defects or non-defects, we need to classify five types of welding defects. Therefore, we must generalize AdaBoost algorithm to adapt to multiple classes. In this paper, the nested dichotomy system is used to construct a multi-class AdaBoost classifier. The nested dichotomy system is a statistical model, which is used to decompose a multi-class problem into multiple two-class problems. The multi-class AdaBoost is constructed to classify five types of welding defects.

The structure of the multi-class AdaBoost classifier is illustrated in Fig. 9. There are four nodes that contain two 2-level cascade AdaBoost classifier, a 5-level cascade AdaBoost classifier and a 4-level cascade AdaBoost classifier, respectively. Node I is trained to classify the five types of welding defects into Crack type and the other four types of welding defects (LOF, LOP, Strip and Round). Node II is trained to classify the rest four types of welding defects into LOF type and the rest three types of welding defects (LOP, Strip and Round). Node III is trained to classify the rest three types

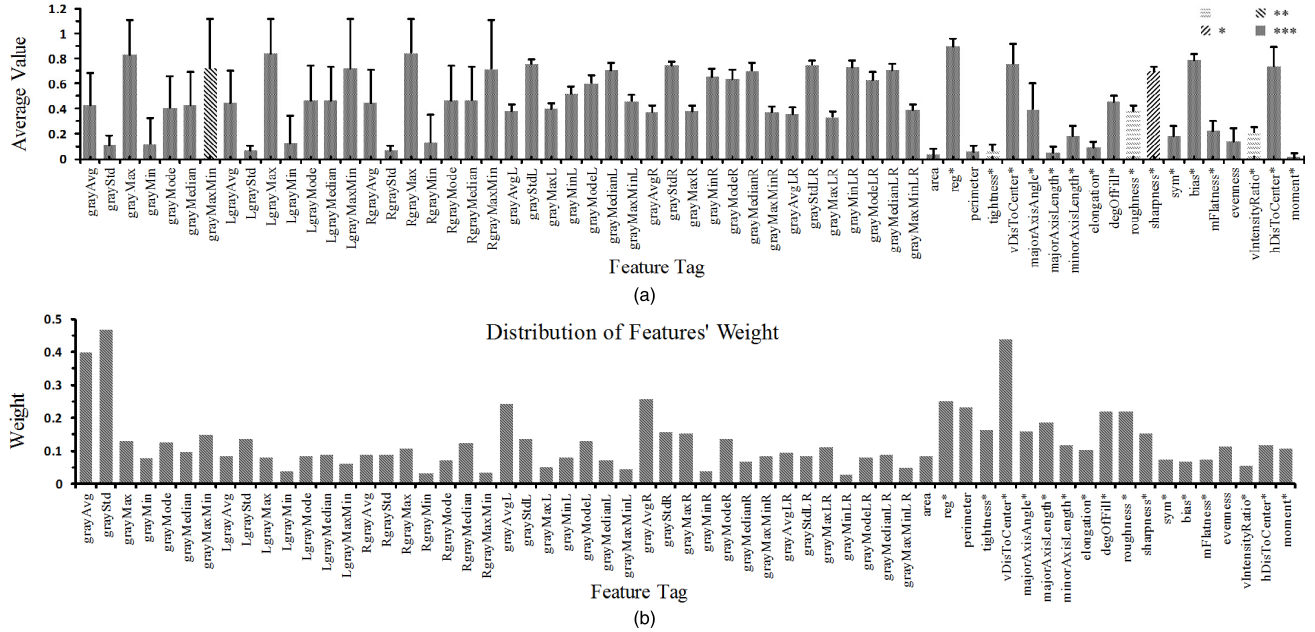


FIGURE 8. Analysis of features. (a) The data are normalized to 0-1. The vertical axis presents the average value for each feature. The horizontal axis shows features. The error bar is the standard deviation. The legend * presents the feature's significance level. One-way analysis of variance is used for the significance test. p is the significance value. *: 0.05 <p≤ 0.1, **: 0.01 <p ≤0.05, ***: p ≤ 0.01. (b) The vertical axis presents the weight of features. The weight of one feature is the weight of the corresponding weak learner in the AdaBoost classifier. The feature with * is described in detail in the feature section of IV.

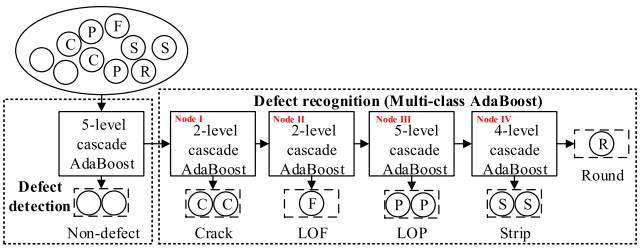


FIGURE 9. Structure of multi-class AdaBoost.

of welding defects into LOP type and the rest two types of welding defects (Strip and Round). Last Node IV is trained to classify the rest two types of welding defects into Strip type and Round type. Therefore, five types of welding defects can be classified by multi-class AdaBoost effectively.

VI. RESULTS AND ANALYSES

This section presents analyses of the classification results. The results of binary classification for defect detection and multi-class classification for the defect recognition are presented and analyzed.

As shown in Fig. 8(b), we obtained the features' weight. The weight of the weak learner is taken as the weight of the corresponding feature. The features with high weights can better describe the characteristics of defects, such as *vDisToCenter*, *grayAvg*, and *grayStd*. *vDisToCenter* is the relative position to the weld central line, *grayAvg* is the average gray value, and *grayStd* is the standard deviation. *grayAvgL*, *grayAvgR*, and *reg* are also good features. *grayAvgL* is the difference between the average gray value of the defect and the average gray value of the left comparison region. *grayAvgR* is the difference between the average gray value of the defect and the average gray value of the right comparison region. *reg* is the relative gray of the defect.

TABLE 1. Sample distribution of potential defects.

	Defects	Non-defects	Potential defects
Training	853	2307	3160
Testing	214	577	791
Sum	1067	2884	3951

For feature selection, a good result can be achieved by selecting the optimal feature number. Sort the features by weight and select the top n features for the classifier. We consider three evaluation parameters: classification accuracy, TPR, and FNR. The variable TPR is critical for industrial applications of automatic welding defect detection. Therefore, we emphasize the variable TPR when deciding the number of features. In the experiment, we extracted 3951 potential defects from 263 X-ray images. As shown in Table 1, there are 1067 defects and 2884 non-defects. Fig. 10 presents the classification results with different feature numbers. The optimal number of features is set as 55.

The parameters λ is adjusted to balance the TPR and FNR. The optimal λ can be decided according to the receiver operating characteristic curve (ROC curve). Fig. 11 illustrates the results of different λ. Five-fold cross validation is used to test the classifier. The average testing accuracy is shown in Fig. 11. It can be seen from Fig. 11 that the detection accuracy and the FNR decline when λ increases. Nevertheless, the TPR

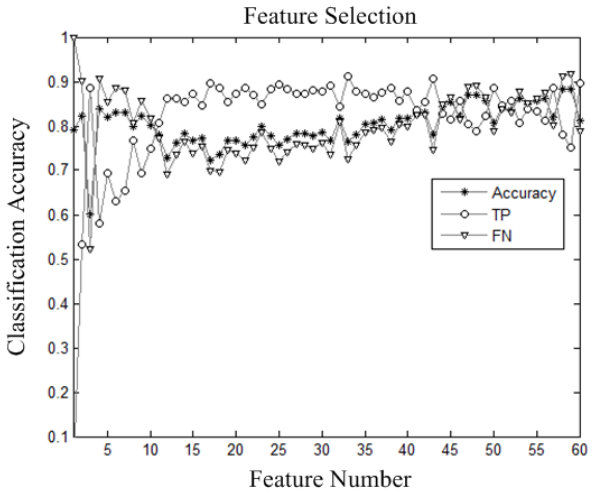


FIGURE 10. Illustration of feature selection. The figure shows the classification accuracy, the rate of TP, and the rate of FN with different numbers of features.

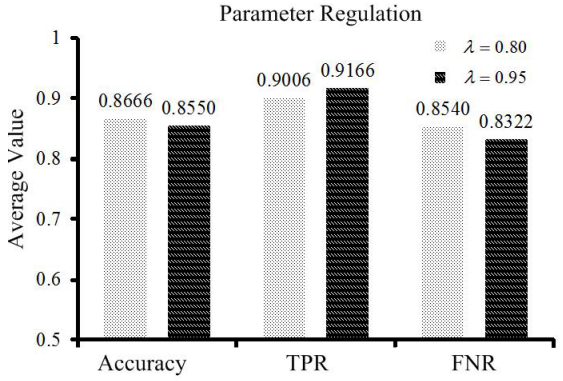


FIGURE 11. Classification performance of five-fold cross validation for tuning the penalty term. TPR, rate of true positive; FNR, rate of false negative.

increases at the same setting. In this paper, we improve the TPR by increasing λ.

For defect detection, the average testing accuracy of the five-fold cross validation is shown in Table 2. We implemented three comparison tests: SVM, KNN, and RF. The library OpenCV provides the realization of SVM and RF. The kernel of the SVM is a radial basis function. The parameters of the three classifiers have been optimized. The k of KNN is optimized by the training data. The number of trees in the forest is set to 20, and the prior probability is computed for classification. The average classification accuracy of the AdaBoost classifier is 85.5%. The average classification of RF is higher than that of the others. However, a large difference was found regarding TPR. The average TPR of the cascade AdaBoost classifier is 91.66%. The average TPR of the SVM is 71.04%, that of the KNN is 69.26%, and that of the RF is 81.53%. To obtain a high TPR, the cascade AdaBoost classifier was applied for defect detection.

For defect recognition, the extended multi-class AdaBoost was utilized for defect classification. The sample distribution

TABLE 2. Cross validation results of four classifiers.

	Classifier	Five Random Tests (%)					
		1	2	3	4	5	AVG
ACC	ADA	85.34	85.32	86.33	84.56	85.95	85.50
	SVM	89.00	86.33	88.35	87.97	88.23	87.98
	KNN	87.36	85.57	85.70	85.32	85.32	85.85
	RF	91.66	90.00	89.62	91.52	92.66	91.09
TPR	ADA	90.65	90.61	91.55	90.61	94.86	91.66
	SVM	76.17	64.79	73.24	69.95	71.03	71.04
	KNN	73.83	66.67	69.95	68.54	67.29	69.26
	RF	84.58	77.93	77.93	83.57	83.64	81.53
FNR	ADA	83.36	83.36	84.40	82.32	82.64	83.22
	SVM	93.76	94.28	93.93	94.63	94.62	94.24
	KNN	92.37	92.55	91.51	91.51	92.01	91.99
	RF	94.28	94.45	93.93	94.45	96.01	94.63

ACC, classification accuracy; TPR, rate of true positive; FNR, rate of false negative; AVG, average value of the five random tests; ADA, AdaBoost classifier; SVM, support vector machine classifier; KNN, k-nearest neighbor classifier; RF, random forest classifier.

TABLE 3. Sample distribution of defects.

	Crack	LOF	LOP	Strip	Round	All
Training	22	19	134	102	576	853
Testing	6	5	33	26	144	214
Sum	28	24	167	128	720	1067

Crack, crack defect; LOF, lack of fusion; LOP, lack of penetration; Strip, stripy defect; Round, round defect.

TABLE 4. Classification accuracy of defect recognition.

Classifier	Average accuracy of five random tests (%)			
	ADA	SVM	KNN	RF
testing	74.98	75.63	76.19	81.54

ADA, AdaBoost classifier; SVM, support vector machine classifier; KNN, k-nearest neighbor classifier; RF, random forest classifier.

is shown in Table 3. The samples distribute unevenly among different types. The number of round defects is 30 times that of LOF defects.

Table 4 presents the results for the classification of defect types. The average testing classification accuracy of AdaBoost is 74.98%. The confusion matrix for defect type identification is illustrated in Table 5. For unevenly distributed samples, the classification performance of SVM and KNN are worse than that of AdaBoost and RF. One can see from Table 5 that SVM and KNN neglect the samples with small quantity, such as cracks and LOF defects. Moreover, the SVM cannot recognize LOF defects absolutely. In addition, AdaBoost performed better than RF in the identification of cracks, LOF defects, and LOP defects. AdaBoost can classify the unevenly distributed samples and put stress on particular types. The three defects with small quantities can cause more severe harm to the weld than stripy defects and round defects. Therefore, AdaBoost was selected for the recognition of defect type.

By comparing the AdaBoost with SVM, KNN and RF, the experimental results show that AdaBoost is the best classifier in defect detection and defect recognition. The proposed method can be utilized for feature selection. Specifically,

TABLE 5. Confusion matrix of five-classification for defects.

Classifier	Label	Estimated label				
		Crack	LOF	LOP	Strip	Round
ADA	Crack	2.2	0.4	1.4	0.6	1
	LOF	0.2	1.2	1.6	1.4	0.4
	LOP	0.6	2.4	21.2	4.2	5
	Strip	1	0.8	11.8	9.4	2.6
	Round	1	1	12.2	3.8	126
SVM	Crack	0.4	0	0	0.2	5
	LOF	0	0	0.6	0.8	3.4
	LOP	0	0	11.2	2	20.2
	Strip	0	0	2	7.6	16
	Round	0.2	0	0.6	1	142.2
KNN	Crack	1	0.2	0.6	1	2.8
	LOF	0	0.4	1	1	2.4
	LOP	0.2	0	18.2	3.8	11.2
	Strip	0.4	0.8	3.4	9	12
	Round	0.6	0.2	4.2	5	134
RF	Crack	1.2	0.2	0.6	1.2	2.4
	LOF	0	0.4	1.8	1.6	1
	LOP	0	0	18.8	7.2	7.4
	Strip	0	0	8.6	15.4	1.6
	Round	0.4	0.2	2.6	2.6	138.2

ADA, AdaBoost classifier; SVM, support vector machine classifier; KNN, k-nearest neighbor classifier; RF, random forest classifier; Crack, crack defect; LOF, lack of fusion; LOP, lack of penetration; Strip, stripy defect; Round, round defect.

it can improve the TPR and deal with the classification of unevenly distributed samples.

VII. CONCLUSION

The consequences of welding defects, such as oil leakage, constitute a major threat to the safe operation of pipelines. Hence, we proposed an automatic welding defect detection system. In X-ray images from industry, images with defects comprise a relatively small number. In addition, massive non-defects exist in the potential defects obtained by image processing. Moreover, the distribution of defect types is uneven. Defect inspection in industry requires high TPR. Therefore, the proposed system aimed at improving TPR and reducing FPR with uneven distributed samples. A total of 3951 potential defects were acquired by the background subtraction method. This paper made two main contributions:

1) Improving the TPR in defect detection. We proposed the AdaBoost method with penalty term. The average testing accuracy of defect detection was 85.5%, the TPR was 91.66%, and the FNR was 83.22%. Three contrast tests were implemented with the same data. Compared with SVM, KNN and RF, the TPR increased by more than 10%.

2) Classification for uneven distributed data. For defect recognition, the method placed different emphases on the defect according to the defect's severity. The classification performance of the proposed method was better than that of SVM, KNN, and RF in classifying the crack, lack of fusion, and lack of penetration.

In this paper, it is shown that the proposed method can be effectively utilized in defect detection, and it is practical in industrial applications.

REFERENCES

[1] *Statistics of Pipelines for Transporting All Over the World*. Accessed: Aug. 21, 2016. [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/fields/383.html

[2] H. Shi, J. Shao, D. Du, B. Chang, and H. Cao, "Noise reduction of the real-time X-ray image based on modified adaptive local noise reduction filter," in *Proc. 4th Int. Congr. Image Signal Process.*, vol. 4, Oct. 2011, pp. 1945–1949.

[3] J. Shao, D. Du, B. Chang, and H. Shi, "Automatic weld defect detection based on potential defect tracking in real-time radiographic image sequence," *NDT&E Int.*, vol. 46, pp. 14–21, Mar. 2012.

[4] J. Sun, C. Li, X. Wu, V. Palade, and W. Fang, "An effective method of weld defect detection and classification based on machine vision," *IEEE Trans. Ind. Informat.*, to be published. doi: 10.1109/TII.2019.2896357.

[5] B. Chen, Z. Fang, Y. Xia, L. Zhang, Y. Huang, and L. Wang, "Accurate defect detection via sparsity reconstruction for weld radiographs," *NDT&E Int.*, vol. 94, pp. 62–69, Mar. 2018.

[6] Z. Liao and J. Sun, "Image segmentation in weld defect detection based on modified background subtraction," in *Proc. 6th Int. Congr. Image Signal Process.*, vol. 2, Dec. 2013, pp. 610–615.

[7] X. Li, S. K. Tso, X.-P. Guan, and Q. Huang, "Improving automatic detection of defects in castings by applying wavelet technique," *IEEE Trans. Ind. Electron.*, vol. 53, no. 6, pp. 1927–1934, Dec. 2006.

[8] Y. Zhang, D. Lefebvre, and Q. Li, "Automatic detection of defects in tire radiographic images," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 3, pp. 1378–1386, Jul. 2017.

[9] H. Y. T. Ngan, G. K. H. Pang, and N. H. C. Yung, "Performance evaluation for motif-based patterned texture defect detection," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 1, pp. 58–72, Jan. 2010.

[10] H. Jiang, Y. Zhao, J. Gao, and W. Zhao, "Weld defect classification based on texture features and principal component analysis," *Insight-Non-Destructive Test. Condition Monit.*, vol. 58, no. 4, pp. 194–200, 2016.

[11] I. Valavanis and D. Kosmopoulos, "Multiclass defect detection and classification in weld radiographic images using geometric and texture features," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 7606–7614, Dec. 2010.

[12] J. Shao, H. Shi, D. Du, L. Wang, and H. Cao, "Automatic weld defect detection in real-time X-ray images based on support vector machine," in *Proc. 4th Int. Congr. Image Signal Process.*, vol. 4, Oct. 2011, pp. 1842–1846.

[13] B. M. Haddad, S. Yang, L. J. Karam, J. Ye, N. S. Patel, and M. W. Braun, "Multifeature, sparse-based approach for defects detection and classification in semiconductor units," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 1, pp. 145–159, Jan. 2018.

[14] W. Yong and H. Guo, "Weld defect detection of X-ray images based on support vector machine," *IETE Tech. Rev.*, vol. 31, no. 2, pp. 137–142, 2014.

[15] F. Mekhalfa and N. Nacereddine, "Multiclass classification of weld defects in radiographic images based on support vector machines," in *Proc. 10th Int. Conf. Signal-Image Tech. Internet-Based Sys.*, Marrakech, Morocco, Nov. 2014, pp. 1–6.

[16] F. Mekhalfa and N. Nacereddine, "Gentle Adaboost algorithm for weld defect classification," in *Proc. Signal Process., Algorithms, Archit., Arrangements, Appl. (SPA)*, Poznan, Poland, Sep. 2017, pp. 301–306.

[17] N. Nacereddine, A. B. Goumeidane, and D. Ziou, "Unsupervised weld defect classification in radiographic images using multivariate generalized Gaussian mixture model with exact computation of mean and shape parameters," *Comput. Ind.*, vol. 108, pp. 132–149, Jun. 2019.

[18] J. Zapata, R. Vilar, and R. Ruiz, "Automatic inspection system of welding radiographic images based on ANN under a regularisation process," *J. Nondestruct. Eval.*, vol. 31, no. 1, pp. 34–45, Mar. 2012.

[19] W. Hou, Y. Wei, J. Guo, Y. Jin, and C. Zhu, "Automatic detection of welding defects using deep neural network," *J. Phys., Conf. Ser.*, vol. 933, Jan. 2018, Art. no. 012006. doi: 10.1088/1742-6596/933/1/012006.

[20] J. Zapata, R. Vilar, and R. Ruiz, "Performance evaluation of an automatic inspection system of weld defects in radiographic images based on neuro-classifiers," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8812–8824, 2011.

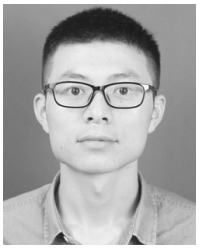
[21] D. You, X. Gao, and S. Katayama, "WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 628–636, Jan. 2015.

[22] Y. Freund, R. E. Schapire, and N. Abe, "A short introduction to boosting," *J. Jpn. Soc. Artif. Intell.*, vol. 14, no. 5, pp. 771–780, 1999.

[23] G. Wang and T. W. Liao, "Automatic identification of different types of welding defects in radiographic images," *NDT&E Int.*, vol. 35, no. 8, pp. 519–528, Dec. 2002.

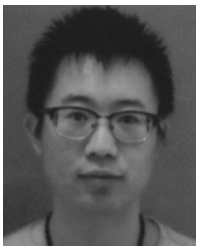
[24] N. Nacereddine and M. Tridi, "Computer-aided shape analysis and classification of weld defects in industrial radiography based invariant attributes and neural networks," in *Proc. 4th Int. Symp. Image Signal Process. Anal.*, Sep. 2005, pp. 88–93.
[25] M. Kearns and L. Valiant, "Cryptographic limitations on learning Boolean formulae and finite automata," *J. ACM*, vol. 41, no. 1, pp. 67–95, Jan. 1994.

FENG DUAN received the B.E. and M.E. degrees in mechanical engineering from Tianjin University, China, in 2002 and 2004, respectively, and the M.S. and Ph.D. degrees in precision engineering from The University of Tokyo, Japan, in 2006 and 2009, respectively. He is currently a Professor with Nankai University, China. His research interests include cellular manufacture systems, rehabilitation robots, and brain–machine interfaces.



SHIFAN YIN received the B.E. degree in electrical engineering and automation from Hangzhou Dianzi University, Hangzhou, China, in 2017. He is currently pursuing the master's degree with the College of Artificial Intelligence, Nankai University, Tianjin, China. His research interests include computer vision and pattern recognition.

PEIPEI SONG received the B.E. degree in intelligence science and technology and the M.S. degree in control science and engineering from Nankai University, Tianjin, China, in 2014 and 2017, respectively. She is currently pursuing the Ph.D. degree in engineering with The Australian National University, Canberra Australia. Her research interests include deep learning for computer vision and pattern recognition.



WENKAI ZHANG received the B.E. degree in automation from Nankai University, Tianjin, China, in 2016, where he is currently pursuing the master's degree with the College of Artificial Intelligence. His research interests include computer vision and deep learning.

CHI ZHU received the Ph.D. degree from The University of Tokyo, in 1999. He was a Researcher with the National Institute for Resource and Environment and Yokohama National University, in 1999 and 2000, respectively. He was a Postdoctoral Associate with the Massachusetts Institute of Technology, in 2001. He has been an Associate Professor with the Maebashi Institute of Technology, Japan, since 2007. His research interests include humanoid robots, walking-assisted robots, and power-assisted robots.

HIROSHI YOKOI received the Ph.D. degree in precision engineering from Hokkaido University, in 1993. He was an Engineer with Toyota Motor Corporation. He was with the Institute of Bioscience and Human Technology, AIST, as a Researcher, from 1993 to 1995. He was an Associate Professor with the Faculty of Engineering, Graduate School, Hokkaido University, from 1995 to 2004, and the Department of Precision Engineering, Faculty of Engineering, The University of Tokyo, from 2004 to 2009. He is currently a Professor with the Department of Mechanics and Intelligence, The University of Electro-Communications and a Visiting Professor with Nankai University. His current research interests include computational intelligence in robotics, artificial life, and medical engineering.