

Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion

Yalin Wang¹ · Haibing Xia¹ · Xiaofeng Yuan¹ ·
Ling Li¹ · Bei Sun¹

Received: 10 November 2016 / Revised: 7 August 2017 / Accepted: 19 September 2017 /
Published online: 14 October 2017
© Springer Science+Business Media, LLC 2017

Abstract Inspecting steel surfaces is important to ensure steel quality. Numerous defect-detection methods have been developed for steel surfaces. However, they are primarily used for local defects, and their accuracy in detecting distributed defects is unsatisfactory because such defects are difficult to locate and have complex texture characteristics. To solve these issues, an improved random forest algorithm with optimal multi-feature-set fusion (OMFF-RF algorithm) is proposed for distributed defect recognition in this paper. The OMFF-RF algorithm includes the following three aspects. First, a histogram of oriented gradient (HOG) feature-set and a gray-level co-occurrence matrix (GLCM) feature-set are extracted and fused to describe local and global texture characteristics, respectively. Second, given the small number of samples of distributed defect images and the high dimensionality of the extracted feature-sets, a random forest algorithm is introduced to perform defect classification. Third, the feature-sets vary greatly in performance and dimensionality. To improve the fusion efficiency, OMFF-RF merges the HOG feature-set and the GLCM feature-set through a multi-feature-set fusion factor, which changes the number of decision trees that correspond to each feature-set in the RF algorithm. The OMFF factor is found by optimizing the fitting curve of the

✉ Xiaofeng Yuan
yuanxf@csu.edu.cn

Yalin Wang
ylwang@csu.edu.cn

Haibing Xia
haibingxia@csu.edu.cn

Ling Li
lilings@csu.edu.cn

Bei Sun
sunbei@csu.edu.cn

¹ School of Information Science and Engineering, Central South University, Changsha, China

classification accuracy of the test set using a stepping multi-feature-set fusion factor. In experiments, the effectiveness of the proposed OMFF-RF was verified using 5 types of distributed defects collected from an actual steel production line. OMFF-RF achieved a recognition accuracy of 91%, a result superior to support vector machine (SVM) and conventional RF algorithms.

Keywords Steel surface · Distributed defect recognition · Histogram of oriented gradient (HOG) · Gray-level co-occurrence matrix (GLCM) · Random forest (RF) · Optimal multi-feature-set fusion (OMFF)

1 Introduction

Steel is one of the most important metals because of its high yield and extensive applications. According to the report provided by the World Steel Association, in 2015, worldwide crude steel production was approximately 130 million tons per month [4]. Steel is widely used in a variety of applications such as the automobile, aviation, shipbuilding, electronics, and machine tool industries. As the technology and economy develop, the demand level for steel quality continues to increase, particularly in terms of surface integrity. However, a lot of problems like poor raw material quality and systematic process problems often cause the surface defects during the steel production process. Therefore, early inspection of steel surfaces by employing intelligent visual recognition methods is essential to prevent inferior steel products from being manufactured in large quantities.

Steel surface defects can be divided into local defects and distributed defects based on their appearance [16]. Local defects are small surface defects with simple shapes and clear boundaries, and they are easy to segment. In contrast, distributed defects have complex textures and fuzzy boundaries, cover large surface areas, and are difficult to segment. In the literature, methods to detect local defects such as holes [11, 30, 37], cracks [37, 43, 44] and scratches [10, 21, 36, 38, 44, 45, 50] on steel surfaces are well established and are primarily based on machine vision techniques. In most cases, these methods adopt the following frame construction. First, interference is removed from surface defect images through image-preprocessing algorithms, a step that is crucial for subsequent feature extraction. Then, the pretreated surface defect images are analyzed for visual features such as geometric shapes [10, 11, 41, 49], gray levels [10, 11, 41, 49, 52], and statistical texture features [8, 10, 13, 27, 32, 41, 45, 52]. These features are typically concatenated into a fusion feature-set (also called a simple multi-feature-set fusion [SMFF] in this paper), that better represents the surface defects. Recently, deep learning models [42] have been used to extract deep features. However, they are very time consuming during the extraction of deep features. Also, the features have high dimensionality, which hinders the efficiency of classification models. Finally, a classification algorithm such as a support vector machine (SVM) [1, 10–12, 21, 22, 33, 36, 40, 44, 50–52], neural network (NN) [8, 22, 23, 28, 33, 37, 41, 48, 49], fuzzy inference system (FIS) [5, 6, 9, 48, 52], learning vector quantifier (LVQ) [43] or self-organizing map (SOM) [30, 32] is utilized to classify defects on steel surfaces based on the fusion feature-set. However, few research studies have focused on complex steel surface distributed defects such as scale red, fold, heavy scale, and salt and pepper. Although some distributed defects have been considered in previous studies, their performance need to be further improved. Therefore, investigating additional distributed defect-recognition methods is essential.

Tracing this problem to its cause, specialized studies that focus on recognizing distributed defects on steel surfaces are both limited and unsatisfactory because several underlying complications exist:

1. Distributed defects are difficult to locate on steel surfaces. In general, these types of defects occupy most of the steel surface but are not completely connected. Local discretization exists, and finding a region of interest (ROI) in the image is difficult; as a result, distributed defects cannot be detected using simple methods such as threshold methods, edge detection, or other image-segmentation approaches.
2. Distributed defects have complex (often irregular) texture characteristics on steel surfaces. Because of these special texture characteristics, the frequently used geometric features cannot be used to detect these types of defects. Moreover, using only a single texture feature is unsuitable for feature extraction.
3. The commonly used SMFF functions poorly for defect recognition on steel surfaces. The SMFF causes surface defects to be insufficiently well classified because of its forced concatenation; this is particularly evident when the dimensionality and performance of the concatenated feature-sets are different. Moreover, the samples of distributed defect images are very limited.

In this paper, to overcome the aforementioned problems, a novel modeling framework is proposed for distributed defect recognition on steel surfaces. First, to handle the localization problem and the complex texture characteristics, two types of feature-description operators—HOG and GLCM—were utilized to extract features from distributed defect images. To take full advantage of the HOG and GLCM feature-sets, they were fused to ensure that the final fusion feature-set could simultaneously describe local and holistic information. Second, to resolve the problems associated with small samples of distributed defect images and high dimensionality of extracted feature-sets, an RF algorithm was introduced. Then, the fusion feature-set via SMFF was presented to the RF algorithm to classify the distributed defects. However, because of the serious imbalance in the dimensionality and performance of the HOG and GLCM feature-sets and the randomness of feature subset selection in the RF algorithm, a direct integration of these two feature-sets does not achieve satisfactory classification and identification performance. Therefore, a multi-feature-set fusion factor was introduced to merge the HOG and GLCM feature-sets. Using the curve fit of the RF classification accuracy achieved for the test set with a stepping multi-feature-set fusion factor, an OMFF factor was obtained by optimizing the above curve-fitting function. By changing the number of decision trees corresponding to each feature-set in the RF algorithm based on the OMFF factor, an OMFF-RF algorithm was obtained to improve the classification accuracy of distributed defects on steel surfaces. The novelties and contributions of this paper are listed below.

1. Aiming at practical industrial application, i.e., distributed defect recognition, the HOG feature-set was introduced and fused with the GLCM feature-set to better represent complex texture characteristics. The HOG feature-set describes local texture information, while the GLCM feature-set, which is frequently used to extract features of local defects, can capture global texture information. Thus, HOG and GLCM feature-sets were fused to improve the recognition performance of distributed defects. The experiment results show that the HOG feature-set and fusion feature-set can achieve a better performance.

2. Focusing on small samples of distributed defect images and high dimensionality of extracted feature-sets, the RF algorithm was first introduced to perform defect recognition on steel surfaces. Although other classification algorithms such as SVM can also solve the problem of small samples and high dimensionality, the RF algorithm has better classification ability because of bootstrap resampling with replacement during the establishment of training sets and optimal division via feature subsets.
3. Considering the difference in the dimensionality and performance of feature-sets, a new RF algorithm—namely, OMFF-RF—is proposed that can merge multiple feature-sets and achieve a commendable level of performance for distributed defect classification. To our knowledge, this is the first RF algorithm using multiple features that preserves each individual feature-set and optimally fuses all the feature-sets. The results demonstrate that OMFF-RF have a better performance at recognizing distributed defects.

The remainder of this paper is organized as follows. In Section 2, related work concerning background, methods to detect defects on steel surfaces as well as issues and motivation is introduced. Section 3 presents the details of the proposed recognition algorithm, including RF with simple multi-feature-set fusion (SMFF-RF) and OMFF-RF. Section 4 describes the experimental procedures used in this work, including image processing, feature extraction and classification, and presents the results of experiments and a discussion. Finally, Section 5 presents conclusions and future work.

2 Related work

2.1 Background

Defect on steel surfaces is one of the main factors affecting steel quality. Accordingly, surface inspection is of great significance to improve steel quality. Over the past decades, a massive number of recognition methods that employ diverse features and classification algorithms have been developed for classifying steel surface defects. Based on the identified defect types, the existing studies on steel surface defect recognition methods can be split into two categories: local defect recognition and distributed defect recognition. In addition, based on the data information used, the local defect recognition methods can also be divided into two groups: supervised and unsupervised.

2.2 Methods to detect defects on steel surfaces

Substantial researches have focused on local defect-identification methods on steel surfaces. SVM is one of the most widely used supervised classification algorithms [1, 10–12, 21, 22, 33, 36, 40, 44, 50–52] for steel surface defect detection. In 2006, 46 geometric features and 8 gray-level features were extracted, and SVM was adopted as the defect classification method by Choi et al. [11]. This method can classify 5 defect types in rolling strip steel with accuracy values ranging from 87% to 94%. In 2010, Zhao et al. [52] introduced fuzzy functions into SVM (FSVM) and used it to perform surface defect detection on cold rolling strips. The extracted features included gray-level features, invariant moment features and texture features (based on GLCM). In 2011, 49 features, including 18 geometric features, 20 gray-level features, 7 texture features (based on GLCM) and 4 projection features (found using horizontal and

vertical edges), were selected by Chen et al. [10]. Spot planting and scratches were well classified (achieving an accuracy of 94%) using the above features and SVM. NN is another commonly used supervised classification approach [8, 22, 23, 28, 33, 37, 41, 48, 49] that also works well in classifying steel surface defects. In 2009, 6 gray-scale features, 4 GLCM features and 4 geometric features were used as inputs to a back propagation NN (BPNN) with 3 layers and 8 hidden layer nodes [41], achieving an accuracy of 97.19%. Yazdchi et al. [49] reported that a three-layer NN with 4 gray features and 6 geometric features could improve the classification accuracy of defect detection on steel surfaces to 97.9%. In 2010, 12 features were defined for each detected candidate region on steel bars by Liu et al. [28]. The overall classification rate of the proposed architecture using BPNN for 4 defects was 90.66%. A FIS [5, 6, 9, 48, 52], which is based on a supervised fuzzy logic-based classifier, has also played a significant role in steel surface defect classification. In [48], 9 appropriate features were chosen, and the FIS-fuzzy c-means (FCM) method was used. However, this strategy did not achieve good results, classifying only 82.46% of defects correctly. In 2010, Borselli et al. [6] used an FIS to analyze defects on flat steel surfaces using 4 features: num_regions, max_width, shape and brightness. This method worked well for 95% of the images. Many other supervised algorithms have also been applied to classify defects on steel surfaces. In 2006, decision tree-type discrimination logic was employed for 7 off-line sample defect types, resulting in a defect-detection rate of 95.5%, as reported by Sasaki et al. [38]. In 2008, a multivariate discriminant function model was established for defect inspection by Liu et al. [27]. In this method, surface images of a cold rolled steel strip were subdivided into blocks from which corresponding statistical features were extracted. The results were satisfactory, yielding a 91% detection rate. In the same year, linear discriminant analysis (LDA) [10, 13] was applied to classify defects based on a set of features. In 2011, extended Haar rectangle features (such as edge features, line features and center-surround features) were extracted by Yan et al. [45]. Four methods—k-nearest neighbor (KNN), BPNN, SVM and weak classifier adaptive enhancement—were used to classify steel surface defects, resulting in success rates of 80%, 82.86%, 88.57% and 94%, respectively.

All the above recognition algorithms are supervised. In addition, several unsupervised recognition algorithms exist that can classify local defects on steel surfaces. These unsupervised methods include LVQ [43] and SOMs [30, 32]. For example, in 2007, new frequency-domain features optimized by the genetic algorithm (GA) were proposed [43]. In this work, 54 frequency-domain features were used as the input vector of an LVQ neural network to recognize 11 surface defect types on hot-rolled strips with 84.56% accuracy. In 2006, a framework using a local binary pattern (LBP) and gray-level histograms with a SOM-based classifier was proposed by Maenpaa [32]. Overall, distinguishing different local defect types on steel surfaces is not very difficult. A review of existing steel surface-recognition methods is available in the literature [34].

However, there are few recognition methods for distributed defects. Although some of the steel surface-identification methods mentioned above have been utilized to detect various types of distributed defects, their efficacies are not high [1, 6, 8, 10, 16, 30, 38, 43, 44, 48, 49]. For example, in 2010, Martins et al. [30] developed 2 classification systems: One exploited the Hough transform as the detection method for 3 surface defect types on rolled steel with geometric shapes, resulting in an accuracy of approximately 98%. The other employed principal component analysis (PCA) to acquire features and SOMs to classify 3 surface defect types with complex shapes, achieving an overall classification rate of 77%. Apparently, using an SOM is relatively ineffective for classifying distributed defects. In the same year, SVM and

vector-valued regularized kernel function approximation (VVRKFA) were applied by Ghorai et al. [16] to classify 24 classes of flat steel surface defects—5 distributed defect types and 19 local defects. Based on their experimental results, we conclude that local defects can be readily identified, whereas the detection performance for heavy scale, salt and pepper and waviness, which are types of distributed defects, is relatively lower.

2.3 Issues and motivation

The poor performance of these recognition methods on distributed defects occurs principally because the selected features are not propitious for representing complex texture characteristics; an SMFF is less effective for steel surface distributed defect recognition because of its deficiencies in allowing each feature to project itself perfectly in classification algorithms. However, the machine learning domain includes many classification algorithms that have used multiple feature fusion techniques to address such issues. The crucial point in solving the problem is how to determine the weights for multiple features in classification algorithms and combine multiple features appropriately. For example, a hierarchical regression (HR) model was designed in reference [46] to utilize the preserved evidence derived from each individual feature, which was then cooperatively fused to gain a multimedia semantic concept classifier. Similarly, a multiple feature-hashing (MFH) technique was proposed for large-scale near-duplicate video retrieval by Song et al. [39]. The MFH preserved each type of feature and fused multiple features into a joint framework, i.e., a group of hash functions. In reference [15], an optimal graph-learning (OGL) technique that used multiple features was proposed to precisely encode the relationships among the data points. Then, this OGL was integrated with semi-supervised learning (SSL) to solve the multiple feature classification problem. The importance of each preserved feature was determined by the parameter α^t in the SSL algorithm. The above HR, MFH and OGL are classification algorithms that use late feature fusion; they first obtain the separate classification results from each feature and then combine these results for the final classification. Classification algorithms with late feature fusion preserve each type of feature, but require large amounts of computation during training. In contrast, classification algorithms that use early feature fusion fuse multiple features at the input stage of the classifier. For example, reference [17] first acquired the similarities of multiple features derived from random forest classifiers and then combined these similarities to an embedded version used as input to a random forest classifier. However, this early feature fusion approach did not preserve each individual feature in the classification algorithms and still required large amounts of calculation. Inspired by the success of classification algorithms with multiple feature fusion, an effective classification algorithm that can extract appropriate features for complex texture characteristics of distributed defects and consider multiple feature fusion is necessary to solve the aforementioned issues.

3 The proposed algorithm

To perform distributed defect recognition, in this paper, an OMFF-RF algorithm is proposed. An overview of the proposed algorithm is illustrated in Fig. 1. First, the HOG and GLCM feature-sets are extracted and fused to characterize the distributed defects. Then, an RF algorithm is introduced to perform distributed defect classification. Obviously, the SMFF-RF

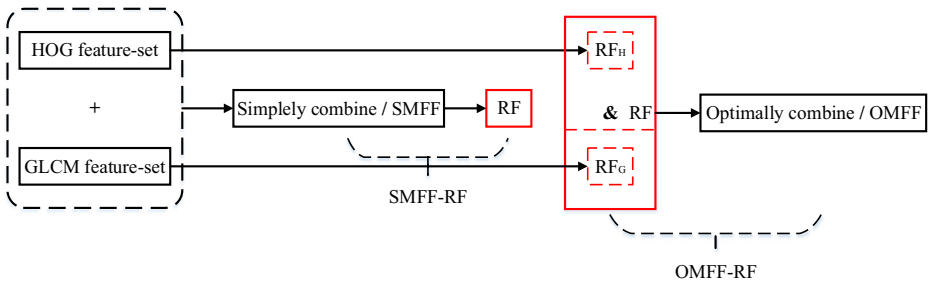


Fig. 1 An overview of the proposed algorithm. The HOG and GLCM feature-sets are extracted and an RF algorithm is introduced to perform distributed defect classification. The SMFF-RF is a classification algorithm with early feature fusion, while the proposed OMFF-RF belongs to classification algorithms with late feature fusion

fuses multiple feature-sets used for the input stage of the RF classifier, while the proposed OMFF-RF preserves each feature-set and optimally combines the classification results from separate sub-RFs (i.e., RF_H and RF_G) with each feature-set. The details of feature extraction and other recognition procedures are presented in the next section.

3.1 RF and SMFF-RF for distributed defect recognition

An RF algorithm was employed to establish an identification model for distributed defects on steel surfaces. The RF algorithm was first proposed by Breiman in 2001 [7] and is a classification algorithm that uses multiple decision trees to train and predict samples. The RF algorithm selects the category corresponding to the maximum vote based on the votes of the leaf nodes of multiple decision trees collected for each category. One advantage of the RF algorithm is that it inherits many of the decision tree properties such as proxy splitting of missing values, the ability to use non-normalized data and the ability to handle tag and numerical features. Another advantage of RF is that its prediction performance can be improved by applying multiple single-classification models (or single decision trees). The established method of training sets and the randomness of feature subset selection make each decision tree of the RF algorithm different. Thus, problems related to the small sample sizes and high dimensionality of extracted feature-sets are solved, and the phenomenon of over-fitting is eliminated. Other strengths, such as the ability to utilize out-of-the-bag (OBB) data to estimate the split performance and to exploit a similarity matrix to measure the proximity between two samples, are also evident. Furthermore, no redundant depiction occurs. In one word, the RF algorithm is one of the best classification algorithms, and it has been successfully applied in many fields. Therefore, this algorithm is selected to recognize distributed defects in this work.

In SMFF-RF, the extracted multi-feature-sets are concatenated and considered as a whole (i.e., a fusion feature-set). Therefore, only the feature-set changes, not the essence of the RF algorithm itself. The operational principle of the SMFF-RF algorithm is quite similar to that of RF algorithm and is described below (see Fig. 2 for reference).

Step 1. Assuming that k decision trees are to be established in the RF algorithm, for every decision tree, adopt bootstrap resampling with replacement to randomly draw training sets $T_i (i = 1, \dots, k)$ with the same sample size from the total training set

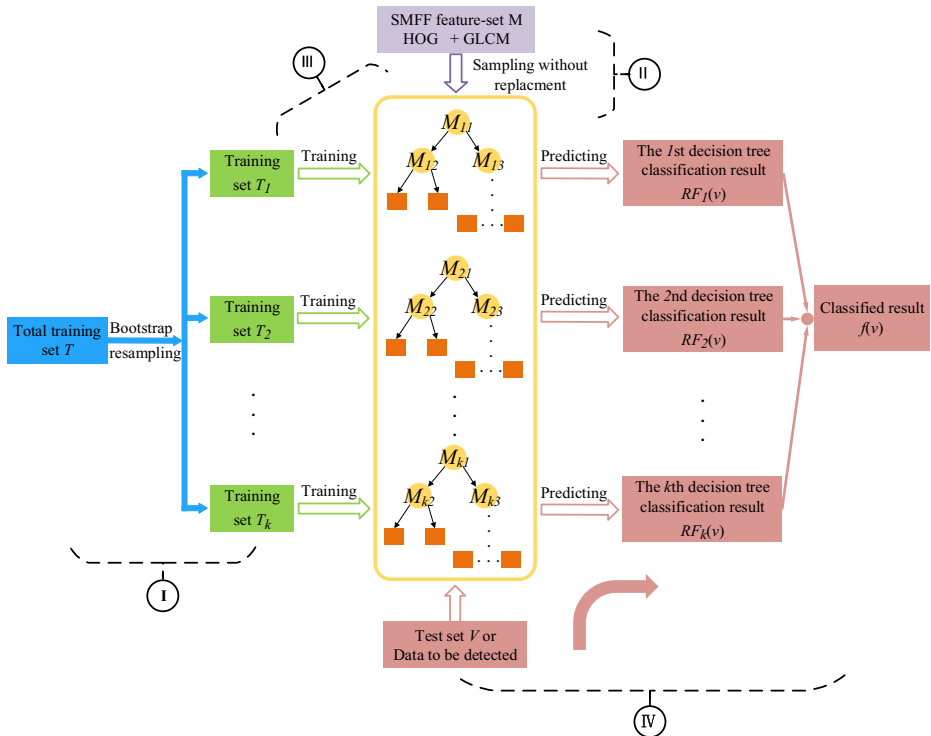


Fig. 2 A schematic diagram of the operational principle underlying the SMFF-RF algorithm. The different colors and numbers denote different modules and steps, respectively: blue and I indicate the process for producing training sets (step 1); light purple and II represent the sampling procedure for feature subsets (step 2); green and III indicate the training of the classifier (step 3); burgundy and IV depict the prediction of the test set (step 4); and gold or orange indicate the classifier established via the steps described above

T . Then, these training sets (T_i) are regarded as the training sets for the k decision trees (see I in Fig. 2).

Step 2. For each branch node of the i th decision tree, employ sampling without replacement to randomly select q feature variables from the total feature-set M (i.e., the SMFF feature-set [HOG + GLCM] consisting of HOG feature-set and GLCM feature-set) as a feature subset $M_{ij}(i = 1, \dots, k, j = 1, 2, \dots)$, where j expresses the number of feature subsets for each decision tree (i.e., the number of branch nodes). Typically, the number of variables in a feature subset M_{ij} is the square root of the dimensions of the total feature-set, M : $\sqrt{\dim(M)}$. Therefore, $j \leq \sqrt{\dim(M)}$ (see II in Fig. 2).

Step 3. Train each decision tree using T_i as a training set and M_i as a set of feature subsets. Then, choose the variable m_{ij} with the best classification ability from the feature subset M_{ij} to classify the samples by setting thresholds for each branch node $node_{ij}$. The representative choosing criterion is the Gini index:

$$Gini(node) = \sum_{c=1}^C p_c(1 - p_c) = 1 - \sum_{c=1}^C p_c^2 \tag{1}$$

where p_c is the proportion of samples belonging to class c in the node $node$. The best classification ability is guaranteed to minimize Gini split index:

$$Gini_{split}(node_{ij}) = p_l Gini(node_l) + p_r Gini(node_r) \quad (2)$$

where p_l and p_r denote the proportion of samples assigned to child nodes $node_l$ and $node_r$, of parent node $node_{ij}$. Thus, the variable m_{ij} with the best classification ability is chosen by

$$m_{ij} = \arg \min_{m \in M_{ij}} Gini_{split}(node_{ij}). \quad (3)$$

Repeat this process until classification is completed or all the feature subsets have been used. At the end, this process obtains an RF classifier $\{RF(T_i, M_i), i = 1, \dots, k\}$ where $RF(T_i, M_i)$ is the classification model of the i th decision tree (see III in Fig. 2).

Step 4. Input the test set V , collect the voting results of all decision trees, and search for the category that received the maximum number of votes (i.e., the classification result):

$$f(v) = \arg \max_c \sum_{i=1}^k I(RF_i(v) = c) \quad (4)$$

where v represents a sample from test set V , $f(v)$ is the final recognition result, $RF_i(v)$ expresses the classification result of i th decision tree, c is the class label, and I indicates the characteristic function (see IV in Fig. 2).

The introduction of the randomness of feature subset selection and the establishment of training set can equip the SMFF-RF algorithm with excellent anti-noise characteristics and avoids the over-fitting phenomenon. The SMFF-RF algorithm can handle data with high dimensionality; therefore, feature reduction is not required.

3.2 Improved OMFF-RF for distributed defect recognition

The SMFF-RF algorithm can fuse a variety of feature-sets, but it may cause a problem owing to the randomness of feature subset selection and the feature variables having independent and identical distributions. As the difference in the dimensionality of different types of feature-sets increases, the selection probability of feature variables in the higher-dimensionality feature-sets also increases, resulting in an extreme imbalance in feature-set fusion in which the low-dimensionality feature-sets will be meaningless. Another problem that must be considered, is that different types of feature-sets have different strengths and weaknesses. Therefore, the SMFF-RF algorithm must be improved to guarantee that different types of feature-sets can be fully fused and that the advantage of each feature-set can be adequately utilized.

In this paper, we primarily consider the fusion of two types of feature-sets. The HOG feature-set and the GLCM feature-set are used as an example. A multi-feature-set fusion factor, ε , is introduced to determine the proportion of decision trees that are separately allocated to the two feature-sets in the RF algorithm. This variable ε is defined as follows:

$$\varepsilon = \frac{n}{k} \quad (5)$$

where n denotes the number of decision trees occupied by HOG feature-set, and k is the number of decision trees to be established in the RF algorithm. Therefore, the proportion of decision trees dominated by the GLCM feature-set is $1 - \varepsilon$. The value of ε is between 0 and 1.

The difference between the improved RF algorithm (which involves a dynamic multi-feature-set fusion factor) and the SMFF-RF algorithm is that the branch nodes of n decision trees among all decision trees are fragmented with the feature subset $H_{ij}(i = 1, \dots, n, j = 1, 2, \dots)$ obtained from the HOG feature-set, whereas the branch nodes of the remaining $k - n$ decision trees are divided by feature subset $G_{(i-n)j}(i = n + 1, \dots, k, j = 1, 2, \dots)$ obtained from the GLCM feature-set (Fig. 3). Similar to the SMFF-RF algorithm, the variable numbers in the feature subsets H_{ij} and $G_{(i-n)j}$ are $\sqrt{\dim(H)}$ and $\sqrt{\dim(G)}$. This improvement is based on the fact that multi-feature-set fusion is accomplished by adjusting the proportions of the decision trees corresponding to the different feature-sets.

To make the improved RF algorithm effective for distributed defects, the OMFF factor ε_{max} must be determined. To do that, we selected different multi-feature-set fusion factor ε values via

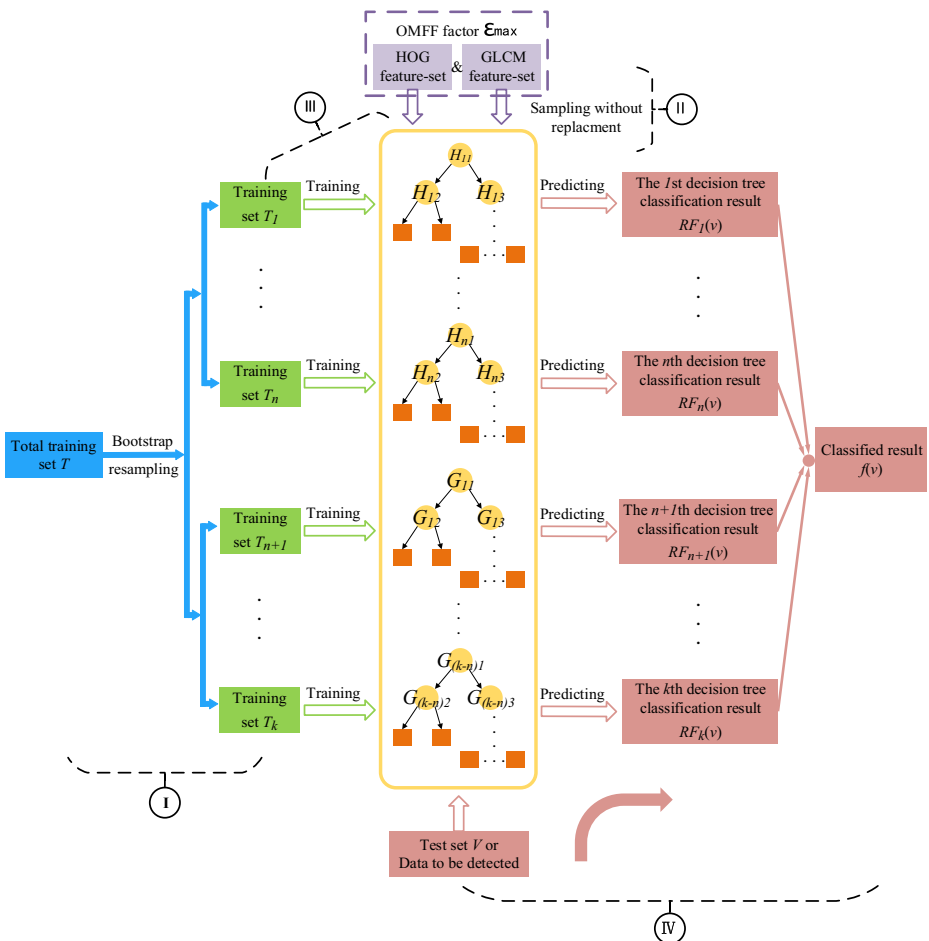


Fig. 3 A schematic diagram of the operational principle underlying the improved RF algorithm with the multi-feature-set fusion factor ε . All the representations are similar to those in Fig. 2, except that the HOG&GLCM feature-set is obtained when the OMFF factor ε_{max} is actualized

equal interval sampling to acquire the RF classification accuracies for the test set and then used a curve to fit the discrete points and obtain the RF recognition accuracy function rf of the test set. Finally, the OMFF factor ε_{max} was determined by optimizing the above function.

The detailed procedures are shown below.

Step 1. The discrete points with RF classification accuracy $Rf(\varepsilon_i)$ of the test set can be obtained using a stepping multi-feature-set fusion factor, ε_i , where the step length is m , as follows:

$$\left\{ (\varepsilon_i, Rf(\varepsilon_i)) \mid \varepsilon_i = i \times l, h = \frac{1}{l}, i = 0, 1, \dots, h \right\}. \quad (6)$$

Step 2. The RF recognition accuracy function rf of the test set can be obtained by curve fitting (CF) the above discrete points:

$$rf(\varepsilon) = g_{CF}(\varepsilon, \theta^*), \text{ where } \theta^* = \arg \min_{\theta} \sum_{i=1}^n [g_{CF}(\varepsilon_i, \theta) - Rf(\varepsilon_i)]^2. \quad (7)$$

Here, θ represents the parameters in the curve-fitting function.

Step 3. The optimization algorithm is used to find the ε at which the value of the RF recognition accuracy function of the test set is maximized, namely, the OMFF factor:

$$\varepsilon_{max} = \arg \max_{\varepsilon} rf(\varepsilon). \quad (8)$$

The method described above is the proposed OMFF-RF algorithm, which fuses the HOG and GLCM feature-sets with the OMFF factor ε_{max} (i.e., HOG&GLCM). This approach can be adapted for cases where three or more feature-sets must be fused using principles similar to the above. The advantage of the OMFF-RF algorithm is that it is suitable for fusing different feature-sets with larger or smaller differences in dimensionality. Moreover, the OMFF-RF algorithm inherits all the performance of the RF algorithm.

3.3 Implementation details

This subsection will discuss the implementation details of our proposed OMFF-RF algorithm. Figure 4 shows the flowchart of proposed algorithm. Table 1 provided the detailed implementation steps of the proposed modeling framework.

4 Experiments and results

In this paper, because of the limited numbers of some distributed defect types, only 5 types of representative distributed defects on steel surfaces were collected from an actual steel production line and analyzed: scale red, fold, heavy scale, rolled-in scale and salt and pepper (Fig. 5a-e). Images of these defects were used to build a model for distributed defect recognition on steel surfaces. We randomly selected 350 images (768×240 pixels) to construct the training set. Then, 132 images (768×240 pixels) that were completely different from those used in the training set containing the 5

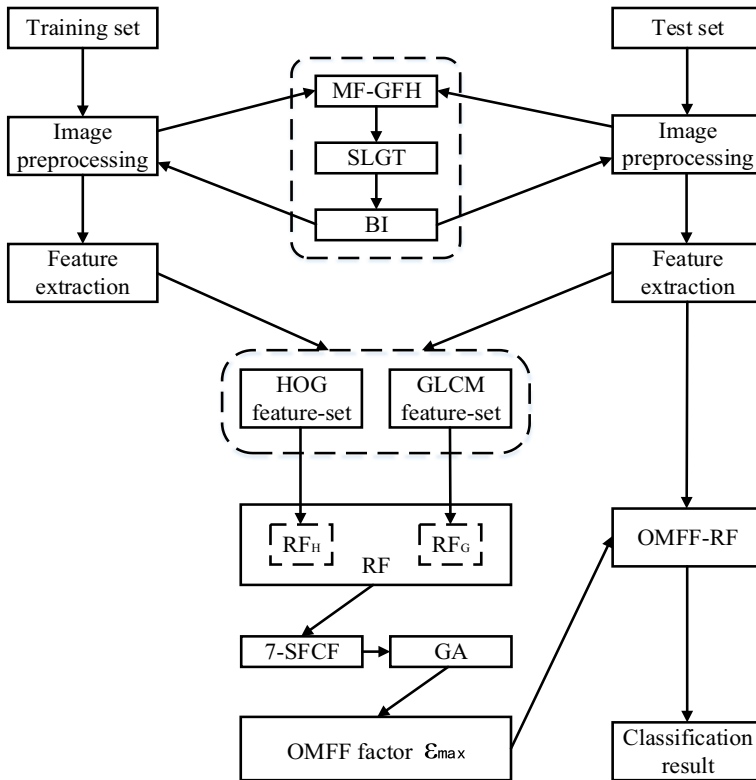


Fig. 4 The flowchart of the proposed OMFF-RF algorithm. The complete process consists of image preprocessing, feature extraction and classification. MF-GH, SLGT, BI, 7-SFCF and GA represent median filtering based on the gray frequency histogram, segmented linear gray transform, bilinear interpolation, 7th-order sine-function curve fit and genetic algorithm, respectively

types of distributed defects were used to create a test set. Table 2 lists information of the training and test sets for each type of distributed defect in the experiments. Moreover, the experimental platform was a single personal computer (PC) containing a Core i5–4460 central processing unit (CPU) running at 3.20 GHz and an 8G internal storage.

To verify the effectiveness of our proposed algorithm, after performing the same image preprocessing and feature extraction tasks, the proposed OMFF-RF was compared to SVM and traditional RF with HOG, GLCM, and SMFF (i.e., HOG + GLCM) feature-sets. Figure 4 depicts the experimental procedures to perform distributed defect recognition on steel surfaces using our proposed OMFF-RF. Full details of the experimental process, including image preprocessing, feature extraction and classification, are described below.

4.1 Image preprocessing

Many types of interference occur during the process of acquiring, transmitting, and converting steel surface images and they can be roughly divided into external and internal interference. External interference primarily involves lighting sources and electromagnetic waves. In contrast, internal interference comprises changes in the basic properties of photosensitive resistance and interior circuits that are generated during the photoelectric conversion process of the image-

Table 1 The implementation details of the proposed OMFF-RF algorithm

Algorithm OMFF-RF

Let $T = \{(h_1, g_1, y_1), \dots, (h_A, g_A, y_A)\}$ and $V = \{(h'_1, g'_1, y'_1), \dots, (h'_B, g'_B, y'_B)\}$ represent the training set and test set via image preprocessing and feature extraction respectively, where h, h' and g, g' denote the feature vectors extracted from HOG feature and GLCM feature respectively, and y is the corresponding image-class label.

For $\varepsilon = 0: l: 1$

 For $i = 1: n$

 1. Randomly draw training set T_i of size A from T via bootstrap resampling.

 2. Randomly select feature subset H_{ij} obtained from the HOG feature-set by sampling without replacement to form a set of feature subsets H_i .

 3. Use T_i and H_i to fit a tree:

 a. Find the variable h_{ij} with the best binary split among the feature subset H_{ij} for each branch node $node_{ij}$ according to formula (3).

 b. Repeat Step (a) until classification is completed or all the feature sets have been used.

 End

 For $i = (n + 1): k$

 1. Randomly draw training set T_i of size A from T via bootstrap resampling.

 2. Randomly select feature subset $G_{(i-n)j}$ obtained from the GLCM feature-set by sampling without replacement to form a set of feature subsets $G_{(i-n)}$.

 3. Use the T_i and $G_{(i-n)}$ to fit a tree:

 a. Find the variable $g_{(i-n)j}$ with the best binary split among the feature subset $G_{(i-n)j}$ for each branch node $node_{ij}$ according to formula (3).

 b. Repeat Step (a) until classification is completed or all the feature sets have been used.

 End

 Make a prediction at a test set V :

$$f(V) = \arg \max_c \left(\sum_{i=1}^n I(RF_i(v) = c) + \sum_{i=n+1}^k I(RF_i(v) = c) \right) \tag{9}$$

 where

$$RF_i = \begin{cases} RF(T_i, H_i) & i \leq n \\ RF(T_i, G_{(i-n)}) & i > n \end{cases} \tag{10}$$

 and get the RF classification accuracy $Rf(\varepsilon)$ of the test set.

 End

 Obtain the RF classification accuracy function r_f according to formula (7) using curve fitting.

 Find the OMFF factor ε_{max} via an optimization algorithm.

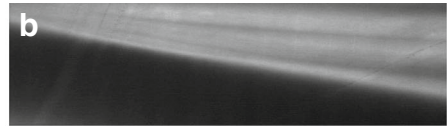
The dashed line is used for division

acquisition system. These interferences might result in the phenomena of polarized light, decreased definition and motion blur of steel surface images. In particular, because of the reflective characteristics of steel surfaces [29], the lighting power supply partially illuminates the steel surface images, which significantly affects subsequent operations. Thus, steel surface images require preprocessing before feature extraction. Image preprocessing consists of image restoration, image enhancement, image transformation and image segmentation. Image restoration and image enhancement are the primary considerations in this paper.

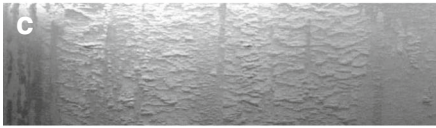
The main objective of image restoration is image de-noising to restore degraded images. The types of noise that arise during steel image acquisition include Gaussian noise, salt and pepper noise, and impulse noise. Particularly, Gaussian noise is a type of sensor noise or electronic circuit noise caused by low light or high temperatures, and it can negatively impact



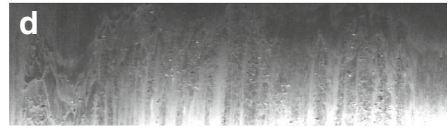
An example of a scale red defect



An example of a fold defect



An example of a heavy scale defect



An example of a rolled-in scale defect



An example of a salt and pepper defect

Fig. 5 Example images of the 5 distributed defect types on steel surfaces examined here. These images were collected from an actual steel production line. In this figure, panels (a–e) show scale red, fold, heavy scale, rolled-in scale, and salt and pepper defects, respectively

subsequent procedures. Thus, a median filtering [19, 24] process based on the gray frequency histogram (MF-GFH) was proposed to perform image de-noising. As shown in Fig. 6, the gray values of the steel surface image are concentrated in the range 40–160. These values primarily correspond to the gray values of backgrounds that have relatively few fluctuations associated with low-intensity noise. The application of traditional MF to steel surface images results in distortion. To overcome the limitations of the traditional MF, a filter center frequency f_{center} was introduced. By compiling statistics about the frequency of each gray level using a GFH, we acquired the maximum frequency f_{max} and, subsequently, adopted $f_{max}/2$ as the filter center frequency f_{center} :

$$f_{center} = \frac{f_{max}}{2}. \quad (11)$$

Then, using f_{center} as a threshold, MF is not performed when the frequency f of the gray level is not less than f_{center} (i.e., where the frequency primarily corresponds to the

Table 2 The number of each type of distributed defect images in the training and test sets

Distributed defect types	Number in training set	Number in test set
Scale red	80	30
Fold	56	21
Heavy scale	76	30
Rolled-in scale	64	24
Salt and pepper	74	27
Total	350	132

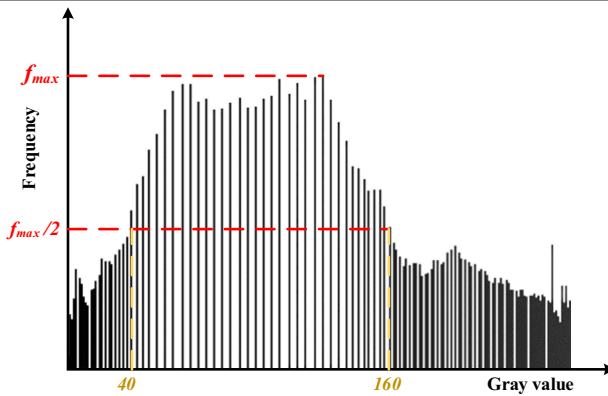


Fig. 6 The gray frequency histogram (GFH) of an example distributed defect image. The gray values of this example lie primarily in the range 40–160 and correspond mainly to the gray values of the background, which has a frequency exceeding $f_{max}/2$. Although the gray values of other distributed defect image backgrounds can fall outside this range (40–160), their frequencies are basically greater than $f_{max}/2$

image background). Otherwise, MF is required. The mathematical expression is as follow:

$$Y_{ij} = \begin{cases} \text{Med}_N \{X_{ij}\} & f_{X_{ij}} < f_{center} \\ X_{ij} & f_{X_{ij}} \geq f_{center} \end{cases} \quad (12)$$

where the median filter is represented by Med ; N is called the neighborhood window; X_{ij} and Y_{ij} are pixel values of image point (i, j) before and after filtering, respectively; and $f_{X_{ij}}$ is the frequency corresponding to the pixel value of the original image point (i, j) .

The purpose of image enhancement is to retain or highlight useful information; improve the image quality [14]; better display the ROI; and enhance the operational value of images by, for example, highlighting or enhancing edge information, contour information, or the contrast ratio. The non-uniformity of illumination [29, 31] and the reflective characteristics [29] of steel surfaces make steel surface images appear partially light or dark, which can interfere with the defect edge information of images. Thus, the segmented linear gray transform (SLGT) is applied to extend the background regions whose gray values are more concentrated and compress other regions to effectively eliminate the distortion of the edge information of defect images caused by the reflective characteristics and the non-uniform illumination. Compressing the gray-level range of the noise associated with steel surface images using the method described above generates processed images with less noise, greater uniformity and stronger contrast.

To facilitate feature extraction and reduce computational complexity, the original images were resized and normalized to 64×64 pixels by applying bilinear interpolation (BI).

4.2 Feature extraction

Feature extraction is a critical step before classification. Features for which high performance can be achieved and that can reflect the nature of images contribute significantly to image classification and recognition. As described in Section 1, the extracted features of steel surface images are commonly statistical features such as geometric shape features, gray features and statistical texture features. The use of other features such as morphological features, spatial

domain features, frequency domain features and fractal model features is comparatively rare. Because of the complex texture characteristics of distributed defects, geometric shape features cannot be used; instead, for this research, statistical texture features represent the best choice. Currently, the most commonly used methods for statistical texture feature extraction are GLCM [8, 10, 14, 20, 41, 52], LBP [20, 32, 47], HOG [2, 26, 35, 47], scale-invariant feature transform (SIFT) [3, 25, 40, 47], and Haar [26, 45, 47]; among these, HOG and SIFT are frequently used for object classification and facial recognition. To account for both the characteristics of distributed defects and the advantages and disadvantages of each texture feature, HOG and GLCM features, which describe the local and global texture information of images, respectively, are used to characterize distributed defects.

HOG, a feature description operator, is utilized for visual inspection by calculating the gradients of every pixel in a local image region and constructing a gradient direction histogram [47]. The information about the gray and texture changes of local image regions, especially edge feature information, can be included in the features extracted by this method without knowing the locations of the defects.

The basic steps of the algorithm are shown in the flow diagram in Fig. 7. Because of the image size (64×64 pixels) after image preprocessing, the algorithm works as follows (for reference, see Fig. 7): First, gamma correction is applied to the input images, and the gradient of each pixel is calculated. Next, 8×8 -pixel areas are identified as *cells*, for a total of 64 *cells* ($8 \times 8 = 64$). The gradient direction of every pixel in the *cell* is divided into 9 *bins* (e.g., $40^\circ - 60^\circ$ and $220^\circ - 240^\circ$ is a *bin*), and the direction gradient histogram of each *cell* is obtained by projecting the gradient direction onto each *bin*. Thus, the number of features in each *cell* is

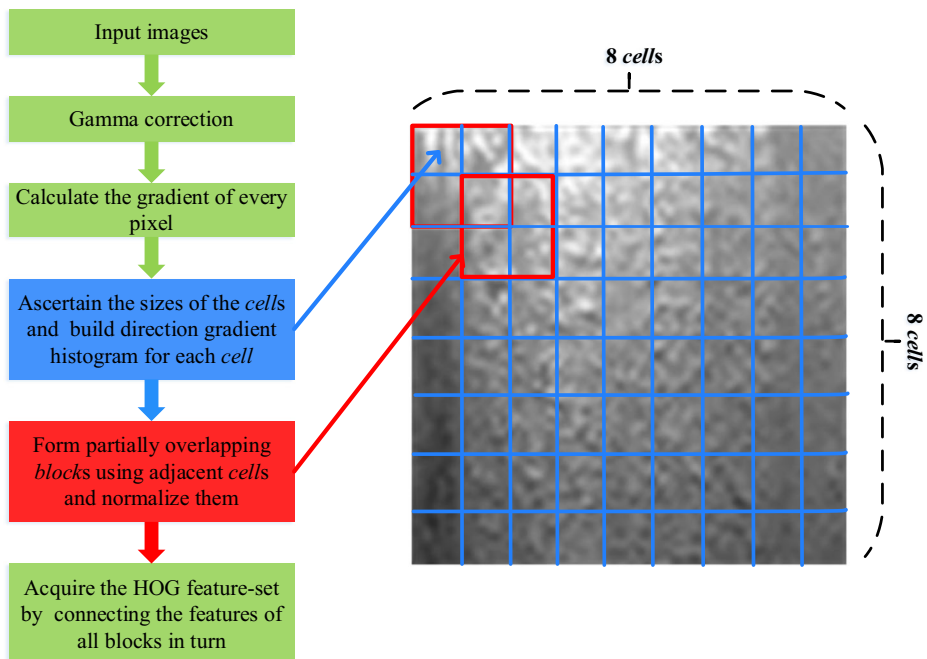


Fig. 7 A schematic diagram of the HOG feature-extraction process. The flow diagram on the left describes the basic steps to acquire the HOG feature-set, whereas the segmentation map on the right denotes the factual work. Blue indicates the treatment of *cells*, whereas red reflects the processing procedures for *blocks*

Table 3 Scalar GLCM features and their properties

Feature types	Properties
Energy	Reflects the uniformity of the gray-level distribution and the coarseness of the texture
Correlation	Reflects the main direction of the texture
Sum of Average	Reflects the luminance of the images
Inverse Different Moment	Reflects the regularity of the texture
Contrast Variance of Difference	Reflect the clarity of the images and the groove depth of the texture
Variance Sum of Variance	Reflect the cyclical nature of the texture
Entropy Sum of Entropy Difference of Entropy	Reflect the heterogeneity and complexity of the texture
Shadow of Clustering Prominence of Clustering	–
Maximal Probability	–

limited to 9. Then, adjacent $2 \times 2 = 4$ cells are selected to form a block, resulting in a total of 49 partly overlapping blocks $((8 - 1) \times (8 - 1) = 49)$. The features of each block normalized by the L2-norm method are derived from the serial features of all cells in a block; thus, $4 \times 9 = 36$ features exist in each block. Finally, the HOG feature-set of the distributed defect images is

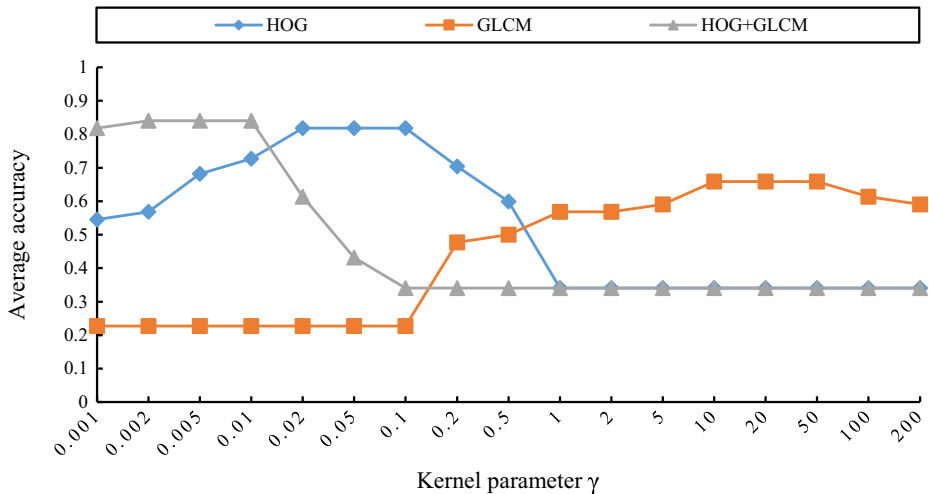


Fig. 8 The recognition accuracy of different RBF kernel parameter values of γ using SVM with the HOG, GLCM and SMFF feature-sets. The optimal RBF kernel parameter values of γ are within the ranges [0.02, 0.1], [10, 50] and [0.002, 0.01], respectively

Table 4 Recognition accuracy obtained with stepping multi-feature-set fusion factor ε

ε	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
Accuracy	0.7500	0.8636	0.8864	0.8636	0.8409	0.8636	0.8864	0.8636	0.8636	0.8409	0.8409
ε	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1	
Accuracy	0.8864	0.8409	0.8409	0.9091	0.8864	0.8409	0.8864	0.8864	0.8409	0.8182	

acquired by connecting the features of all *blocks* in turn, eventually resulting in a total of $49 \times 36 = 1764$ features.

The GLCM, which is a joint probability matrix, is utilized to describe the spatial correlations among pixels in texture images [14, 18]. The GLCM of an image can contain comprehensive gray information about the direction, adjacent spacing and magnitude of changes, and it is sensitive to changes in the texture of the entire image. Therefore, the GLCM is a feasible method for representing the overall features of distributed defect images on steel surfaces.

Assuming that the target image area is $f(x, y)$ and that S represents a set of ordered pixel pairs with a particular spatial relationship in the target image region, the value of each element in the joint probability matrix P is

$$P(f_1, f_2, d, \theta) = \frac{\#\{(x_1, y_1), (x_2, y_2) \in S | f(x_1, y_1) = f_1 \& f(x_2, y_2) = f_2\}}{\#S} \quad (13)$$

where the numerator on the right side of the equation indicates the number of pixel pairs whose distance is d , whose angle is θ and whose gray values are f_1 and f_2 . The denominator is the total number of pixel pairs ($\#$ is used to denote quantities), and P is a normalized GLCM. A different GLCM can be obtained by selecting a different distance d and angle θ .

The GLCMs of distributed defect images in 4 directions can be obtained by setting the distance d equal to 1 and the angle θ equal to $0^\circ, 45^\circ, 90^\circ$ or 135° . Indeed, the angle θ can be

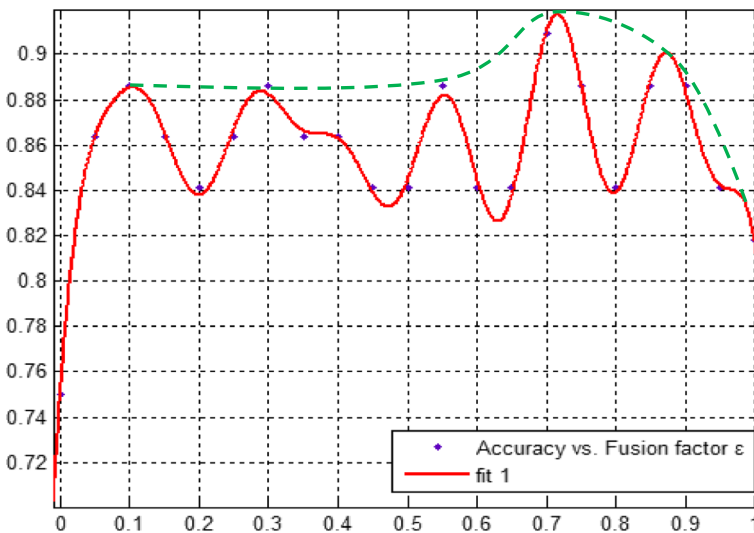


Fig. 9 The function curve of the discrete points of RF recognition accuracy with the stepping multi-feature-set fusion factor obtained using 7-SFCF. The green dashed line represents the envelope of the curve peaks

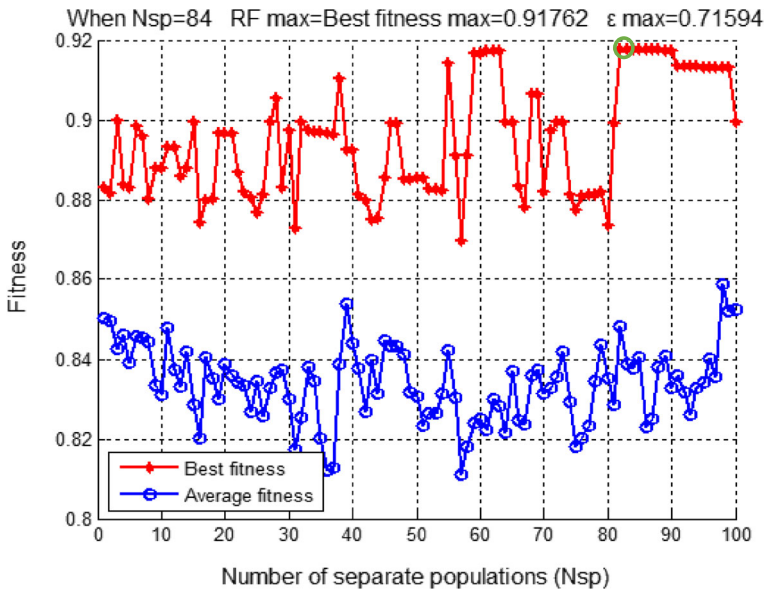


Fig. 10 The process of searching for the OMFF factor ε_{max} by adopting GA. The fitness reflects the recognition accuracy. To eliminate random search error, the recognition performances of different numbers of separate populations (or iterations) is displayed simultaneously. The point in the green circle is the rf_{max} (or best fitness)

expressed in terms of the distance d . That is, 0° , 45° , 90° and 135° can be represented as $[0 \ 1]$, $[-1 \ 1]$, $[-1 \ 0]$ and $[-1 \ -1]$, respectively. To more directly depict the texture properties with a GLCM, 14 scalar features [18] were applied to characterize the GLCM in every direction and, thus, better describe the distributed defects contained in images of steel surfaces (the concrete scalar features and their properties are listed in Table 3).

Accordingly, 56 texture features were abstracted by the GLCM.

In summary, 1764 HOG features and 56 GLCM features were extracted and normalized to values between 0 and 1. The subsequent multi-feature-set fusion is based on the HOG and GLCM feature-sets described above.

4.3 Classification

After image preprocessing and feature extraction, SVM, RF and the proposed OMFF-RF algorithm were employed to establish the classification model for distributed defects on steel surfaces. The HOG, GLCM and SMFF feature-sets (i.e., HOG + GLCM) were provided for

Table 5 The experimental recognition results of 5 types of distributed defects based on the SVM with the HOG feature-set, conducted using the optimal RBF kernel parameter $\gamma = 0.05$

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	15	0	0	9	6	50
Fold	0	21	0	0	0	100
Heavy scale	0	0	27	0	3	90
Rolled-in scale	0	0	0	24	0	100
Salt and pepper	0	0	6	0	21	77.78

Table 6 The experimental recognition results of 5 types of distributed defects based on the SVM with the GLCM feature-set, conducted using the optimal RBF kernel parameter $\gamma = 20$

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	30	0	0	0	0	100
Fold	0	9	0	3	9	42.86
Heavy scale	0	0	30	0	0	100
Rolled-in scale	6	0	0	12	6	50
Salt and pepper	6	9	0	6	6	22.22

SVM and RF. The fusion of the HOG and GLCM feature-sets (i.e., HOG&GLCM) with the OMFF factor ε_{max} , which determines the proportions of the two feature-sets in the decision trees, was utilized with the OMFF-RF algorithm. During the encoding process, the function *forest.train* (the parameters such as *_min_sample_count* (5), *_max_tree_count* (100), *_max_depth* (50), *_nactive_vars* and *_max_categories* were set in *CvRTPParams*) and the function *CvSVM.train* (the parameters including *svm_type* (C_SVC), *Cvalue* (10) and *kernel_type* (RBF) were also set in *CvSVMParams*) of the *letter_cecog.cpp* file in OpenCV were applied to train the classifiers. The classification results can be obtained by using the trained classifier and calling the functions *forest.predict* and *CvSVM.predict* from the *letter_cecog.cpp* file. The OMFF-RF training and prediction functions are available as an original program. It should be noted that most of the above parameter values were acquired via experience. However, the optimal RBF kernel parameter γ of SVM and the OMFF factor ε_{max} of OMFF-RF were obtained through parameter optimization experiments as detailed below.

- 1) *Optimization of the RBF kernel parameter, γ* : To find the optimal RBF kernel parameter γ , a non-uniform grid search in the form of (1, 2, 5) was used to test the recognition performance of SVM with HOG, GLCM and SMFF (i.e., HOG + GLCM) feature-sets (see Fig. 8). Figure 8 clearly shows that the optimal RBF kernel parameter values of γ for SVM with the HOG ($\dim(HOG) = 1764$), GLCM ($\dim(HOG) = 56$) and SMFF ($\dim(SMFF) = 1764 + 56$) feature-sets are within the ranges [0.02, 0.1], [10, 50] and [0.002, 0.01], respectively. Generally, for high-dimensional feature-sets, the settings of RBF kernel parameter γ are relatively small so that feature weights will decay faster. This reduces the high-dimensional subspace to a low-dimensional subspace. In contrast, for low-dimensional feature-sets, to map the data into a linear separable form, the settings of RBF kernel parameter γ are relatively large. This is the reason why the above optimal RBF kernel parameter γ of SVM increases as the dimensionality of the feature-sets decreases. In this paper, the optimal RBF kernel parameter values of γ for SVM with HOG, GLCM and SMFF (i.e., HOG + GLCM) feature-sets were set to 0.05, 20 and 0.005, respectively.

Table 7 The experimental recognition results of 5 types of distributed defects based on the SVM with SMFF (i.e., HOG + GLCM) feature-set, conducted using the optimal RBF kernel parameter $\gamma = 0.005$

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	21	0	0	6	3	70
Fold	0	21	0	0	0	100
Heavy scale	0	0	27	0	3	90
Rolled-in scale	3	0	0	21	0	87.5
Salt and pepper	0	0	6	0	21	77.78

Table 8 The experimental recognition results of 5 types of distributed defects based on the RF with the HOG feature-set

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	21	0	0	6	3	70
Fold	0	21	0	0	0	100
Heavy scale	0	0	30	0	0	100
Rolled-in scale	6	0	0	18	0	75
Salt and pepper	0	0	9	0	18	66.67

2) *Analysis of the OMFF factor, ε_{max}* First, the RF recognition accuracy with the stepping multi-feature-set fusion factor $\varepsilon = (0 : 0.05 : 1)$ was obtained as listed in Table 4. Second, by fitting the discrete points in Table 4 to a 7th-order sine-function curve fit (7-SFCF), the RF recognition accuracy function was acquired (Fig. 9). This curve has multiple peaks, or in other words, the synthetic recognition accuracy obtained with different values of the multi-feature-set fusion factor ε fluctuates, resulting from differences in the representation capabilities of diverse feature-sets for different distributed defects. Based on the envelope of the curve peaks, the overall trend of the accuracy function is described by the slope, and the OMFF factor value ε_{max} ranges from 0.65 to 0.75 when the function value is maximized (i.e., when the classification accuracy is maximized). This result is not unexpected because HOG feature-set are better than GLCM feature-set for recognizing distributed defects (also see Table 12). Finally, by employing the GA [2] to optimize the above function, the OMFF factor was solved; that is, $\varepsilon_{max} \approx 0.7159$ (Fig. 10). Given that 100 decision trees were built and that the number of decision trees corresponding to each feature-set should be an integer, we set $\varepsilon_{max} = 0.71$, where the value of the recognition accuracy function $r_{f_{max}}$ is equal to 0.917. Under these circumstances, the actual classification accuracy is 0.9091. This result is acceptable because the difference (approximately $0.86\% = (0.917 - 0.9091)/0.917$) between the actual and theoretical accuracies is negligible. In conclusion, the OMFF factor ε_{max} of the proposed OMFF-RF algorithm is 0.71, and the corresponding recognition accuracy is 0.9091.

4.4 Results and discussion

In this section, the experimental outcomes are presented, and the distributed defect recognition capacities of the SVM, RF and OMFF-RF algorithms with diverse feature-sets are compared.

Table 9 The experimental recognition results of 5 types of distributed defects based on the RF with the GLCM feature-set

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	27	0	0	0	3	90
Fold	0	12	0	9	0	57.14
Heavy scale	0	0	27	0	0	100
Rolled-in scale	0	0	0	21	3	87.5
Salt and pepper	12	0	0	6	9	33.33

Table 10 The experimental recognition results of 5 types of distributed defects based on the RF with SMFF (i.e., HOG + GLCM) feature-set

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	27	0	0	0	3	90
Fold	0	21	0	0	0	100
Heavy scale	0	0	27	0	3	90
Rolled-in scale	6	0	0	18	0	75
Salt and pepper	0	0	6	0	21	77.78

The experimental recognition results of the 5 types of distributed defects obtained via the SVM algorithm with the HOG, GLCM and SMFF (i.e., HOG + GLCM) feature-sets, conducted using the optimal RBF kernel parameter γ in Section 4.3, are shown in Tables 5, 6 and 7, respectively. The experimental recognition results of 5 types of distributed defects obtained via the RF algorithm with the HOG, GLCM and SMFF (i.e., HOG + GLCM) feature-sets are shown in Tables 8, 9 and 10, respectively. The experimental recognition results of the 5 types of distributed defects obtained via the OMFF-RF algorithm, conducted using the OMFF factor value ε_{max} in Section 4.3 to fuse the HOG and GLCM feature-sets (i.e., HOG&GLCM), are shown in Table 11. Finally, Table 12 presents the comparative experimental results of all the recognition methods regarding average accuracy and overall runtimes (in seconds). Here, the overall runtime reported in the last column of Table 12 reflects the entire runtime including image preprocessing, feature extraction, classifier training, and test set prediction in OpenCV.

The recognition capacities of the HOG and GLCM feature-sets can be derived by comparing Tables 5 and 6 and Tables 8 and 9. To present the differences more clearly, Fig. 11 shows the comparative recognition accuracy of the HOG and GLCM feature-sets using the two different classification algorithms on each type of distributed defect. It is clearly shown that the HOG feature-set works well for fold, heavy scale and rolled-in scale, while the GLCM feature-set performs extremely well for scale red and heavy scale. However, both feature-sets underperform for salt and pepper, although the HOG feature-set is superior to the GLCM feature-set for this distributed defect type. This phenomenon suggests that the inspection capabilities of diverse feature-sets for different distributed defects are different. Moreover, according to Table 12, when using the HOG feature set, both SVM and RF produced better performances than they did when using the GLCM feature-set. This is the reason that the HOG feature-set was introduced and fused with the GLCM feature-set to represent the complex texture characteristics of distributed defects.

In the SVM classifier, the comparative recognition capacities of SMFF (HOG + GLCM) feature-set and all the individual feature-sets are drawn from Tables 5, 6 and 7. For a visual

Table 11 The experimental recognition results of 5 types of distributed defects based on the OMFF-RF algorithms, conducted using the OMFF factor value $\varepsilon_{max}=0.71$ to fuse the HOG and GLCM feature-sets (i.e., HOG&GLCM)

Defect types	Scale red	Fold	Heavy scale	Rolled-in scale	Salt and pepper	Accuracy (%)
Scale red	27	0	0	0	3	90
Fold	0	21	0	0	0	100
Heavy scale	0	0	30	0	0	100
Rolled-in scale	6	0	0	18	0	75
Salt and pepper	0	0	3	0	24	88.89

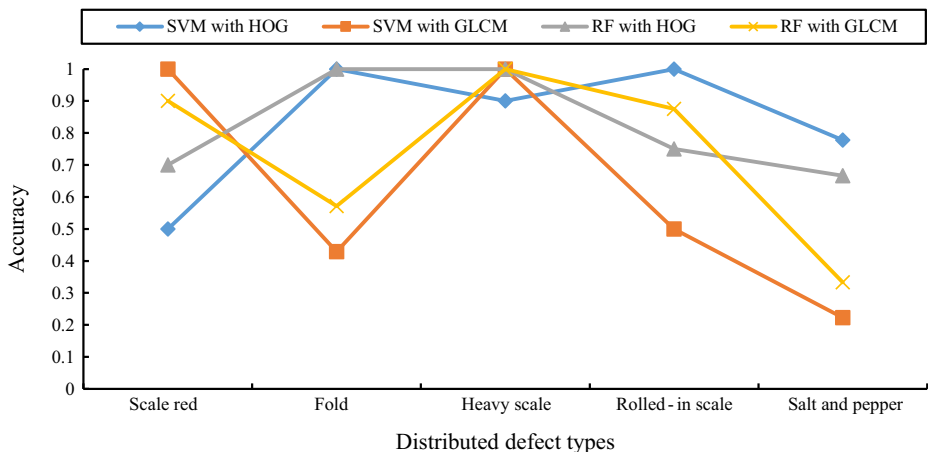
Table 12 The comparative experiment results of all recognition methods on average accuracy and overall runtimes in seconds

Classifier	Features	Average accuracy (%)	Overall runtime
SVM	HOG	81.82	3.681
	GLCM	65.91	41.091
	HOG + GLCM	84.09	56.303
RF	HOG	81.82	8.502
	GLCM	72.73	43.477
	HOG + GLCM	86.36	60.840
OMFF-RF	HOG&GLCM	90.91	9.537

The bold value denotes the best performance in the column

comparison, Fig. 12 shows the comparative recognition accuracy of the SMFF (HOG + GLCM) feature-set and all the individual feature-sets using SVM for each type of distributed defect. From Fig. 12, the SMFF feature-set relies upon the strengths of the better-performing feature-set to compensate for the deficiencies of the other feature-set, especially for scale red, fold, rolled-in scale and salt and pepper. This finding demonstrates that feature fusion can increase the effectiveness of classification algorithms. This improved effectiveness is also apparent from the results of the traditional RF classifier listed in Tables 8, 9 and 10 and in Fig. 13. In addition, from Table 12, regardless of what feature-set is used, the SVM algorithm's classification performance is worse than that of the RF algorithm introduced to perform defect recognition on steel surfaces. The possible cause for this phenomenon is that the RF algorithm establishes training sets via bootstrap resampling with replacement and optimally splits via the feature subsets to resolve the problems of small samples and high dimensionality.

By analyzing Tables 7, 10 and 11 and Fig. 14, we can draw several conclusions about our proposed OMFF-RF algorithm. First, the OMFF-RF algorithm retains the best recognition accuracy of the other two classification algorithms when using the SMFF feature-set for scale red and fold. Second, compared to the other two classification algorithms when using the SMFF feature-set, OMFF-RF improves the recognition accuracy for heavy scale and salt and pepper. These results benefit from the fact that the OMFF-RF algorithm preserves each individual feature-

**Fig. 11** The comparative recognition accuracy of the HOG and GLCM feature-sets using the two different classification algorithms for each type of distributed defect

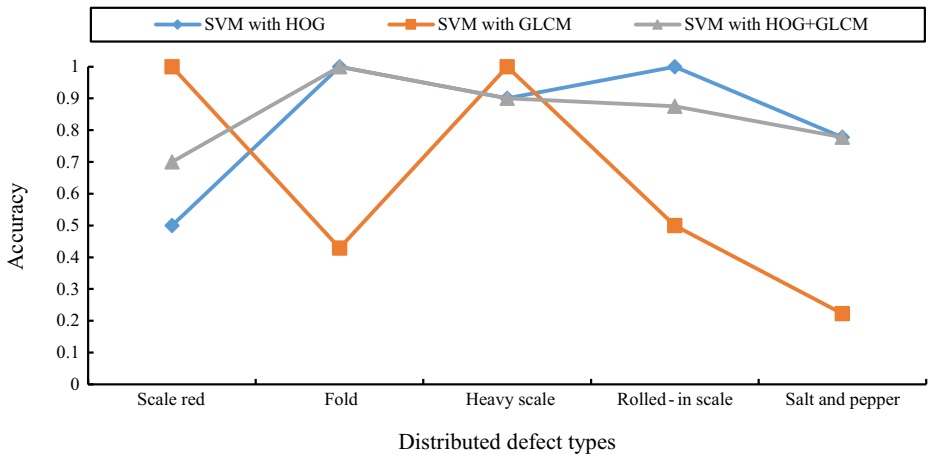


Fig. 12 The comparative recognition accuracy of the SMFF (HOG + GLCM) feature-set and all the individual feature-sets using SVM for each type of distributed defect

set and fuses all the feature-sets in an optimized manner particularly when the dimensionality and performance of the feature-sets are different. According to Table 12, our method achieves results 8.11% and 5.27% higher than SVM and RF with SMFF feature-set, respectively.

From Table 12, the runtime of the SMFF feature-set is visibly longer than the runtimes of all the single-feature-sets because of its higher dimensionality. The runtime of the OMFF feature-set is between those of the two feature-sets using the RF algorithm because of the invariant total number of decision trees and the variable feature fusion factor ϵ . The overall time (9.537 s) included the processing runtimes of both the training set (350 images) and the test set (132 images). Because the training runtime is longer, the prediction runtime for a single image is considerably less than $9.537 \text{ s}/482 = 19.79 \text{ ms}$; this value is very suitable for practical industrial applications.

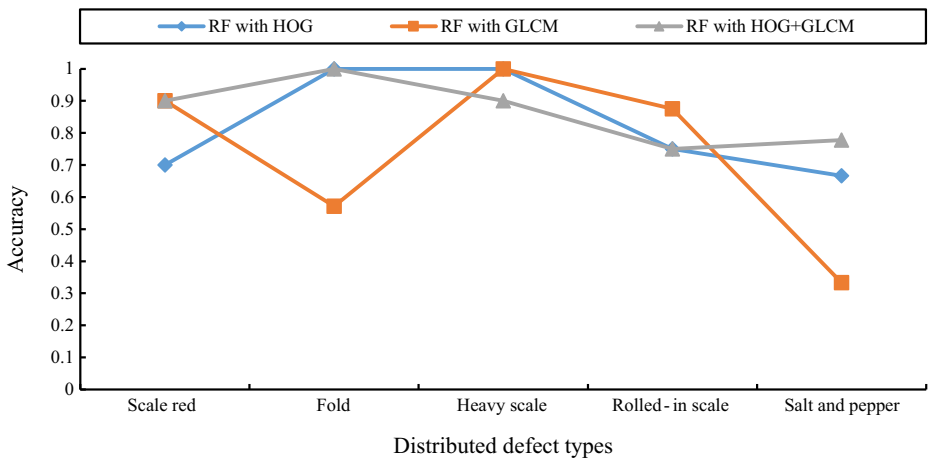


Fig. 13 The comparative recognition accuracy of the SMFF (HOG + GLCM) feature-set and all the individual feature-sets using RF for each type of distributed defect

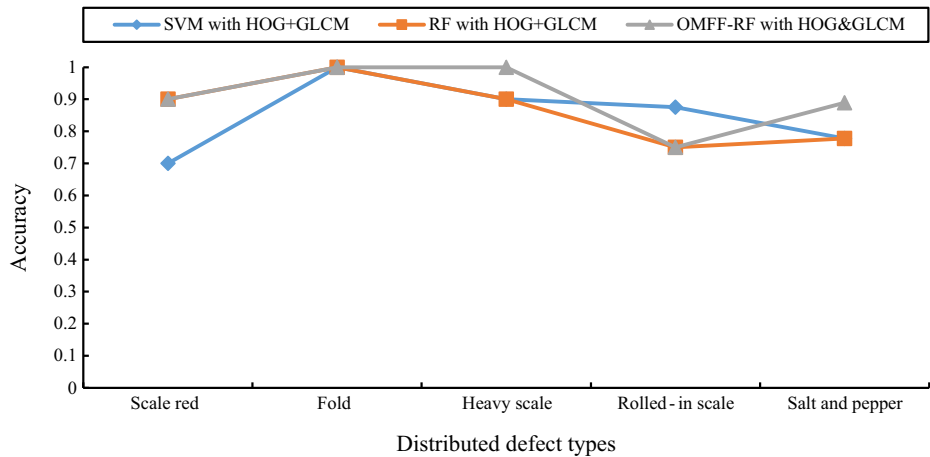


Fig. 14 The comparative recognition accuracy of the proposed OMFF-RF algorithm and other classification algorithms with the SMFF (HOG + GLCM) feature-set for each type of distributed defect

5 Conclusions and future work

In this paper an OMFF-RF algorithm that fuses the HOG and GLCM feature sets with the OMFF factor was presented to distinguish 5 types of distributed defects on steel surfaces acquired from an actual steel production line. The results of the described experiments showed that the OMFF-RF algorithm has great potential for distributed defect recognition and that its image-processing runtime is relatively short.

However, in this work, the types and sample quantities of distributed defects on steel surfaces were limited, and the recognition capabilities of the OMFF-RF algorithm for local defects and free defects were not considered. These issues will be considered in future work. For this algorithm to be utilized in practical industrial applications, the OMFF factor, which is determined based on test results, must be periodically reviewed in terms of historical data. In the future, we plan to investigate other texture features to characterize distributed defects on steel surfaces.

Acknowledgements This work is supported by the Major Program of the National Natural Science Foundation of China (Grant No. 61590921), the National Natural Science Foundation of China (Grant No. 61273187), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (Grant No. 61321003) and the Fundamental Research Funds for the Central Universities of Central South University (Grant No. 2017zzts488).

References

1. Agarwal K, Shivpuri R, Zhu Y, Chang T-S, Huang H (2011) Process knowledge based multi-class support vector classification (PK-MSVM) approach for surface defects in hot rolling. *Expert Syst Appl* 38:7251–7262. <https://doi.org/10.1016/j.eswa.2010.12.026>
2. Ameer A, Kumar KS (2015) Efficient automatic image annotation using optimized weighted complementary feature fusion using genetic algorithm. *Procedia Comput Sci* 58:731–739. <https://doi.org/10.1016/j.procs.2015.08.094>

3. Azhar R, Tuwohingide D, Kamudi D, Sarimuddin, Suciati N (2015) Batik image classification using SIFT feature extraction, bag of features and support vector machine. *Procedia Comput Sci* 72:24–30. <https://doi.org/10.1016/j.procs.2015.12.101>
4. Basson E (2016) World steel in figures 2016. World Steel Association. <http://www.worldsteel.org/dms/internetDocumentList/bookshop/2016/World-Steel-in-Figures-2016/document/World%20Steel%20in%20Figures%202016.pdf>. Accessed 27 May 2016
5. Blackledge JM, Dubovitskiy DA (2008) A surface inspection machine vision system that includes fractal texture analysis. *ISAST Trans Electron Sig Process* 3:76–89
6. Borselli A, Colla V, Vannucci M, Veroli M (2010) A fuzzy inference system applied to defect detection in flat steel production. In: International conference on fuzzy systems. IEEE, Piscataway, NJ, pp 21–26
7. Breiman L (2001) Random forest. *Mach Learn* 45:5–32. <https://doi.org/10.1023/a:1010933404324>
8. Caleb P, Steuer M (2000) Classification of surface defects on hot rolled steel using adaptive learning methods. In: Proceedings KBIESAT fourth international conference on knowledge-based intelligent engineering systems and allied technologies. IEEE, Piscataway, NJ, pp 103–108
9. Catemi S, Colla V, Vannucci M, Borselli A (2012) Fuzzy inference systems applied to image classification in the industrial field. In: Azeem MF (ed) Fuzzy inference system - theory and applications. InTech, Rijeka, pp 243–270
10. Chen Y, Chen L, Liu X, Ding S, Zhang H (2011) Real-time steel inspection system based on support vector machine and multiple kernel learning. In: Advances in intelligent and soft computing, 2011. Springer, Berlin, pp 185–190
11. Choi K, Koo K, Lee J (2006) Development of defect classification algorithm for POSCO rolling strip surface inspection system. In: 2006 SICE-ICASE international joint conference. IEEE, Piscataway, NJ, pp 2499–2502
12. Choi D-C, Jeon Y-J, Lee SJ, Yun JP, Kim SW (2014) Algorithm for detecting seam cracks in steel plates using a gabor filter combination method. *Appl Opt* 53:4865. <https://doi.org/10.1364/ao.53.004865>
13. Cord A, Bach F, Jeulin D (2010) Texture classification by statistical learning from morphological image processing: application to metallic surfaces. *J Microsc* 239:159–166. <https://doi.org/10.1111/j.1365-2818.2010.03365.x>
14. Dutta S, Das A, Barat K, Roy H (2012) Automatic characterization of fracture surfaces of AISI 304LN stainless steel using image texture analysis. *Measurement* 45:1140–1150. <https://doi.org/10.1016/j.measurement.2012.01.026>
15. Gao LL, Song JK, Nie FP, Yan Y, Sebe N, Shen HT (2015) Optimal graph learning with partial tags and multiple features for image and video annotation. In: 2015 I.E. Conference on Computer Vision and Pattern Recognition. IEEE, Piscataway, NJ, pp 4371–4379
16. Ghorai S, Mukherjee A, Gangadaran M, Dutta PK (2013) Automatic defect detection on hot-rolled flat steel products. *IEEE Trans Instrum Meas* 62:612–621. <https://doi.org/10.1109/tim.2012.2218677>
17. Gray KR, Aljabar P, Heckemann RA, Hammers A, Rueckert D (2015) Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *NeuroImage* 65:167–175. <https://doi.org/10.1016/j.neuroimage.2012.09.065>
18. Haralick RM (1979) Statistical and structural approaches to texture. *Proc IEEE* 67:786–804. <https://doi.org/10.1109/proc.1979.11328>
19. Hemalatha G, Sumathi CP (2016) Preprocessing techniques of facial image with Median and Gabor filters. In: 2016 international conference on information communication and embedded systems (ICICES). IEEE, Piscataway, NJ, pp 1–6
20. Hu H, Liu Y, Liu M, Nie L (2016) Surface defect classification in large-scale strip steel image collection via hybrid chromosome genetic algorithm. *Neurocomputing* 181:86–95. <https://doi.org/10.1016/j.neucom.2015.05.134>
21. Yeon YJ, Choi DC, Yun JP, Park C, Kim SW (2011) Detection of scratch defects on slab surface. In: 2011 11th international conference on control, automation and systems (ICCAS). IEEE, Piscataway, NJ, pp 1274–1278
22. Jia H, Murphey YL, Shi J, Chang T-S (2004) An intelligent real-time vision system for surface defect detection. In: Proceedings of the 17th international conference on pattern recognition, 2004. ICPR 2004. IEEE, Piscataway, NJ, pp 239–242
23. Kang G-W, Liu H-B (2005) Surface defects inspection of cold rolled strips based on neural network. In: 2005 international conference on machine learning and cybernetics. IEEE, Piscataway, NJ, pp 5034–5037
24. Kumudham R, Dhanalakshmi, Swaminathan A, Geetha, Rajendran V (2016) Comparison of the performance metrics of median filter and wavelet filter when applied on SONAR images for denoising. In: 2016 international conference on computation of power, energy information and communication (ICCPDIC). IEEE, Piscataway, NJ, pp 288–290

25. Lenc L, Král P (2015) Automatic face recognition system based on the SIFT features. *Comput Electr Eng* 46:256–272. <https://doi.org/10.1016/j.compeleceng.2015.01.014>
26. Liang C-W, Juang C-F (2015) Moving object classification using local shape and HOG features in wavelet-transformed space with hierarchical SVM classifiers. *Appl Soft Comput* 28:483–497. <https://doi.org/10.1016/j.asoc.2014.09.051>
27. Liu W, Yan Y, Li J, Zhang Y, Sun H (2008) Automated on-line fast detection for surface defect of steel strip based on multivariate discriminant function. In: 2008 second international symposium on intelligent information technology application. IEEE, Piscataway, NJ, pp 493–497
28. Liu YC, Hsu YL, Sun YN, Tsai SJ, Ho CY, Chen CM (2010) A computer vision system for automatic steel surface inspection. In: 2010 5th IEEE conference on industrial electronics and applications. IEEE, Piscataway, NJ, pp 1667–1670
29. Liu M, Liu Y, Hu H, Nie L (2016) Genetic algorithm and mathematical morphology based binarization method for strip steel defect image with non-uniform illumination. *J Vis Commun Image Represent* 37:70–77. <https://doi.org/10.1016/j.jvcir.2015.04.005>
30. Martins LAO, Padua FLC, Almeida PEM (2010) Automatic detection of surface defects on rolled steel using computer vision and artificial neural networks. In: IECON 2010 - 36th annual conference on IEEE industrial electronics society. IEEE, Piscataway, NJ, pp 1081–1086
31. Luo Q, He Y (2016) A cost-effective and automatic surface defect inspection system for hot-rolled flat steel. *Robot Comput Integr Manuf* 38:16–30. <https://doi.org/10.1016/j.rcim.2015.09.008>
32. Mäenpää T (2006) Surface quality assessment with advanced texture analysis techniques. In: Proceedings of the international surface inspection summit (ISIS 2006). International Surface Inspection Summit, p 7
33. Masci J, Meier U, Ciresan D, Schmidhuber J, Fricout G (2012) Steel defect classification with max-pooling convolutional neural networks. In: The 2012 international joint conference on neural networks (IJCNN). IEEE, Piscataway, NJ, pp 1–6
34. Neogi N, Mohanta DK, Dutta PK (2014) Review of vision-based steel surface inspection systems. *EURASIP J Image Video Process* 2014:1–19. <https://doi.org/10.1186/1687-5281-2014-50>
35. Pang Y, Yuan Y, Li X, Pan J (2011) Efficient HOG human detection. *Signal Process* 91:773–781. <https://doi.org/10.1016/j.sigpro.2010.08.010>
36. Park C, Won S (2009) An automated web surface inspection for hot wire rod using undecimated wavelet transform and support vector machine. In: 2009 35th annual conference of IEEE industrial electronics. IEEE, Piscataway, NJ, pp 2411–2415
37. Peng K, Zhang X (2009) Classification technology for automatic surface defects detection of steel strip based on improved BP algorithm. In: 2009 fifth international conference on natural computation. IEEE, Piscataway, NJ, pp 110–114
38. Sasaki T, Takada H, Tomura Y (2007) Automatic surface inspection system for tin mill black plate (TMBP). *JPE Technical Report* 9:60–63
39. Song JK, Yang Y, Huang Z, Shen HT, Luo JB (2013) Effective multiple feature hashing for large-scale near-duplicate video retrieval. *IEEE Trans Multimedia* 15:1997–2008. <https://doi.org/10.1109/TMM.2013.2271746>
40. Suvdaa B, Ahn J, Ko J (2012) Steel surface defects detection and classification using SIFT and voting strategy. *Int J Softw Eng Appl* 2:161–165
41. Tang B, Kong J-Y, Wang X-D, Chen L (2009) Surface inspection system of steel strip based on machine vision. In: 2009 first international workshop on database technology and applications. IEEE, Piscataway, NJ, pp 359–362
42. Wang XH, Gao LL, Song JK, Shen HT (2017) Beyond frame-level CNN: saliency-aware 3-D CNN with LSTM for video action recognition. *IEEE Signal Proc Let* 24(4):510–514. <https://doi.org/10.1109/LSP.2016.2611485>
43. Wu G, Zhang H, Sun X, Xu J, Xu K (2007) A bran-new feature extraction method and its application to surface defect recognition of hot rolled strips. In: 2007 I.E. international conference on automation and logistics. IEEE, Piscataway, NJ, pp 2069–2074
44. Wu G, Kwak H, Jang S, Xu K, Xu J (2008) Design of online surface inspection system of hot rolled strips. In: 2008 I.E. international conference on automation and logistics. IEEE, Piscataway, NJ, pp 2291–2295
45. Yan Y, Song K, Xing Z, Feng X (2011) The strip steel surface defects classification method based on weak classifier adaptive enhancement. In: 2011 third international conference on measuring technology and mechatronics automation. IEEE, Piscataway, NJ, pp 958–961
46. Yang Y, Song JK, Huang Z, Ma ZG, Sebe N, Hauptmann AG (2013) Multi-feature fusion via hierarchical regression for multimedia analysis. *IEEE Trans Multimedia* 15:572–581. <https://doi.org/10.1109/TMM.2012.2234731>

47. Yao S, Pan S, Wang T, Zheng C, Shen W, Chong Y (2015) A new pedestrian detection method based on combined HOG and LSS features. *Neurocomputing* 151:1006–1014. <https://doi.org/10.1016/j.neucom.2014.08.080>
48. Yazdchi MR, Mahyari AG, Nazeri A (2008) Detection and classification of surface defects of cold rolling mill steel using morphology and neural network. In: 2008 international conference on computational intelligence for modelling control & automation. IEEE, Piscataway, NJ, pp 1071–1076
49. Yazdchi M, Yazdi M, Mahyari AG (2009) Steel surface defect detection using texture segmentation based on multifractal dimension. In: 2009 international conference on digital image processing. IEEE, Piscataway, NJ, pp 346–350
50. Yun JP, Park C, Bae H, Hwang H, Choi S (2010) Vertical scratch detection algorithm for high-speed scale-covered steel BIC(Bar in Coil). In: 2010 international conference on control automation and systems (ICCAS). IEEE, Piscataway, NJ, pp 342–345
51. Yun JP, Choi D-C, Jeon Y-J, Park C, Kim SW (2014) Defect inspection system for steel wire rods produced by hot rolling process. *Int J Adv Manuf Technol* 70:1625–1634. <https://doi.org/10.1007/s00170-013-5397-8>
52. Zhao J, Yang Y, Li G (2010) The cold rolling strip surface defect on-line inspection system based on machine vision. In: 2010 second Pacific-Asia conference on circuits, communications and system. IEEE, Piscataway, NJ, pp 402–405



Yalin Wang is currently a full professor at the School of Information Science and Engineering, Central South University, Changsha, People's Republic of China. She received her B.Eng. degree in Industrial Electrical Automation and Ph.D. degree in Control Science and Engineering from Central South University in 1995 and 2001, respectively. She visited the University of Alberta, Canada, from 2011 to 2012. Her research interests include the modeling, optimization and control of complex industrial processes; pattern recognition; and machine learning.



Haibing Xia is a postgraduate student at the School of Information Science and Engineering, Central South University, Changsha, People's Republic of China. He received his B.Eng. degree in Automation from Central South University in 2015. His research interests include image processing, machine learning, and data mining.



Xiaofeng Yuan is currently an associate professor with the School of Information Science and Engineering, Central south University. He received the B.Eng. and Ph.D. degrees from the Department of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 2011 and 2016, respectively. He was a visiting scholar with the Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB, Canada, from November 2014 to May 2015. His research interests include pattern recognition and machine learning, big data and deep learning, industrial process soft sensor modeling, and process data analysis.



Ling Li is a Ph.D. candidate at the School of Information Science and Engineering, Central South University, Changsha, People's Republic of China. She received her B.Eng. degree in Automation from Xiangtan University in 2010 and her M.Eng. degree in Control Theory and Control Engineering from Nanjing University of Science and Technology in 2013. Her research interests include process performance assessment, machine learning, and data mining.



Bei Sun is currently a lecturer at the School of Information Science and Engineering, Central South University, Changsha, People's Republic of China. He received his Ph.D. in Control Science and Engineering from Central South University in 2015 and was with the Department of Electrical and Computer Engineering, Polytechnic School of Engineering, New York University, United States from 2012 until 2014. His research interests include the data-driven modeling, optimization and control of nonferrous metallurgical processes.