



Research article

Feature fusion–based preprocessing for steel plate surface defect recognition

Yong Tian*, Tian Zhang, Qingchao Zhang, Yong Li, Zhaodong Wang

State Key Laboratory of Rolling and Automation, Northeastern University, Shenyang 110819, China

* **Correspondence:** Email: tianyong@ral.neu.edu.cn.

Abstract: To address the problem of steel strip surface defect detection, a feature fusion–based preprocessing strategy is proposed based on machine vision technology. This strategy can increase the feature dimension of the image, highlight the pixel features of the image, and improve the recognition accuracy of the convolutional neural network. This method is based on commonly used image feature extraction operators (e.g., Sobel, Laplace, Prewitt, Robert, and local binary pattern) to process the defect image data, extract the edges and texture features of the defect image, and fuse the grayscale image processed by the feature operator with the original grayscale image by using three channels. To consider also computational efficiency and reduce the number of calculation parameters, the three channels are converted into a single channel according to a certain weight ratio. With this strategy, the steel plate surface defect database of NEU is processed, and fusion schemes with different operator combinations and different weight ratios for conversion to the single channel are explored. The test results show that, under the same network framework and with the same computational cost, the fusion scheme of Sobel:image:Laplace and the single-channel conversion weight ratio of 0.2:0.6:0.2 can improve the recognition rate of a previously unprocessed image by 3% and can achieve a final accuracy rate of 99.77%, thereby demonstrating the effectiveness of the proposed strategy.

Keywords: convolutional neural network; image fusion; feature extraction; steel plate surface defect; image recognition

1. Introduction

Surface defects are an important factor affecting the quality of steel plates and strips. More than 60% of the quality objection incidents by users of steel plate and strip products are prompted by surface

defects, thus causing enormous economic losses to the steel companies [1]. With the technological advancement of optical instruments, image processing-based steel plate surface defect recognition has become a research focus of scholars worldwide [2–4]. High-quality images of steel plates can be captured by coordinating the camera, light source, and laser line. Additionally, the surface defects of the steel plate can be detected and classified by a custom algorithm. The automatic surface defect detection system can perform online detection of surface defects and provide timely feedback. Detection and timely feedback are key to improving the surface quality of steel plates and strips. With the improvement of the production line speed and the increasingly stringent requirements that users impose regarding product quality, it is urgent to improve the accuracy, speed, and efficiency of defect detection and recognition algorithms in surface detection systems.

The machine vision detection algorithm consists of two steps: Feature extraction and classification. Multiple sets of features from different aspects, such as gray level, shape, and texture, can be extracted from the defect image and are conducive to the correct classification of defects. However, too many features affect the complexity and performance of the classifier. To achieve the best possible classification result without losing features, feature selection methods are generally used to process the features. Commonly used feature extraction algorithms include the gray-level cooccurrence matrix (GLCM) [5] and scale-invariant feature transformation (SIFT) [6]. With these methods, a high-dimensional space is mapped to a low-dimensional space to generate a linear combination of the original features and to reduce the feature dimension. Defect classification falls under the scope of pattern recognition. Commonly used classification algorithms include support vector machines [7], naïve Bayes [8], K-nearest neighbors [9], and random forests (RFs) [10]. Relevant classification algorithms have received increasing attention and have achieved good results in practice. However, the abovementioned traditional detection methods have poor generalization ability and rely on the personal experience of researchers to design the feature engineering. Hence, it is very difficult to apply these methods in large-scale industrial production.

In recent years, deep convolutional neural networks (CNNs) [11–13] have sparked a resurgence of visual research because of the benefit of self-learning image features. Inspired by the biological natural visual perception mechanism, CNNs can substantially reduce the number of training parameters by using the weight sharing mechanism mainly through supervised learning and backpropagation training. However, in practical applications, because specific parameters, such as weight bias in training, cannot be set, deep CNNs have certain shortcomings. Therefore, only the hyperparameters during learning can be adjusted to control the fitting. Then, the weight bias parameters needed for recognition are generated. This black-box algorithm has been criticized by researchers. Especially in steel plate surface defect recognition, due to the low contrast of image data, CNNs learn image features through randomly generated convolution kernels. By using such convolution kernels with uncertain parameters, important features may be unminable. The unminability of important features easily leads to misjudgment and affects the further improvement of the recognition rate.

To detect cracks in nuclear power plant components, Chen [14] and Jahanshahi et al. proposed a deep learning framework based on a CNN and naïve Bayes data fusion scheme to analyze single video frames for crack detection. The authors presented a new data fusion scheme to aggregate the information extracted from each video frame to enhance the overall performance and robustness of the system. A method to extract the potential features of steel plate defects by fusing multiple convolutional network feature layers was proposed by He [15]. However, this method relies on the network to perform the fusion, and there is a certain randomness in the learning process; additionally, this method

needs a huge amount of data. Wang et al. [16] proposed an improved RF algorithm with optimal multi-feature-set fusion for distributed defect recognition. This algorithm fuses the histogram of oriented gradients (HOG) feature set and GLCM feature set by using the multi-feature-set fusion factor to change the number of decision trees corresponding to each feature set in the RF algorithm. In this paper, a feature fusion preprocessing method is designed. This method can achieve a high recognition rate for six categories of defects by using a small amount of data, help the CNN to mine the deeper features of steel plate defect images, and give full play to the learning advantage of neural networks.

Since the matrices and defects of steel plates are mostly gray and black, the overall contrast between the defects and their defects is extremely low. Even if an advanced color camera is used to collect images in a red–green–blue format, the color features of steel plate defects are not obvious. Therefore, the images acquired by industrial cameras are mostly single-channel grayscale images. In this study, specific contour and texture detection operators are combined to process the original grayscale image acquired by an industrial camera to extract the image features, fuse the image features with the original image based on the channel model, and finally convert the image to a single channel based on a certain weight ratio. By using the above processing, not only can the feature dimension of the image be increased and the image be more easily activated by the network to extract features, but the same pixel level as that of the original grayscale image can also be guaranteed, without wasting extra computational cost. Additionally, artificial guidance helps the CNN learn image features purposefully and avoid the drawbacks of black-box algorithms, thereby improving the recognition accuracy of low-contrast images.

The main contributions of this study are as follows:

(1) This study promotes the purposeful learning of a CNN, improves the classification capability of the network, and proposes a multichannel fusion strategy based on the combination of feature operators. This strategy not only increases the feature dimension but also retains the feature information of the original image.

(2) This study proposes converting multiple channels into a single channel for network training according to a certain weight ratio to reduce computational cost. Such a conversion not only ensures the classification accuracy but also does not affect the computational speed.

(3) This study obtains the optimal weight ratio of fusion and conversion by using the traversal method, which involves comparing the impacts of different fusion and conversion schemes on the classification results.

2. Proposed method

When a CNN extracts images, the shallow convolution kernel is mainly used to extract the edge features, and the deep convolution kernel is mainly used to extract the high-level abstract features, such as texture. Then, these features are integrated through the full connection layer to make the classification. A convolution kernel with a specific template can extract the directional features of the image by using, e.g., the Sobel [17], Laplace [18], Prewitt [19], and Roberts operators [20]; the texture features of image data can be encoded by contrast and integration of pixels, such as with the local binary pattern (LBP) [21]. In this paper, a feature extraction operator is used to extract the edge and texture information of an image to make it is easier for the shallow convolution kernel to learn edge features and for the deep convolution kernel to learn texture features. Additionally, the influence of different fusion schemes on the steel plate surface defect recognition rate is investigated.

Currently, five feature extraction operators are commonly used: Sobel, Laplace, Prewitt, Robert, and LBP. In this paper, we choose each pair of these operators to process the original grayscale image, and the resulting feature matrix is fused with the original grayscale matrix. The original grayscale matrix is placed in the middle channel, and the image processed by the feature operators is randomly placed on the remaining two channels. Fifteen combination schemes of the same feature operator and different feature operators form a three-channel color-effect image. The three-channel data after fusion results in the data calculation burden for two additional channels in the model. In contrast, the model of the original grayscale image requires only the data calculation for a single channel. To mitigate the increase in computational load after multichannel fusion, we use a traversal method to explore the optimal weight ratio and convert the three-channel image data into single-channel image data according to a certain weight ratio. The flowchart is shown in Figure 1.

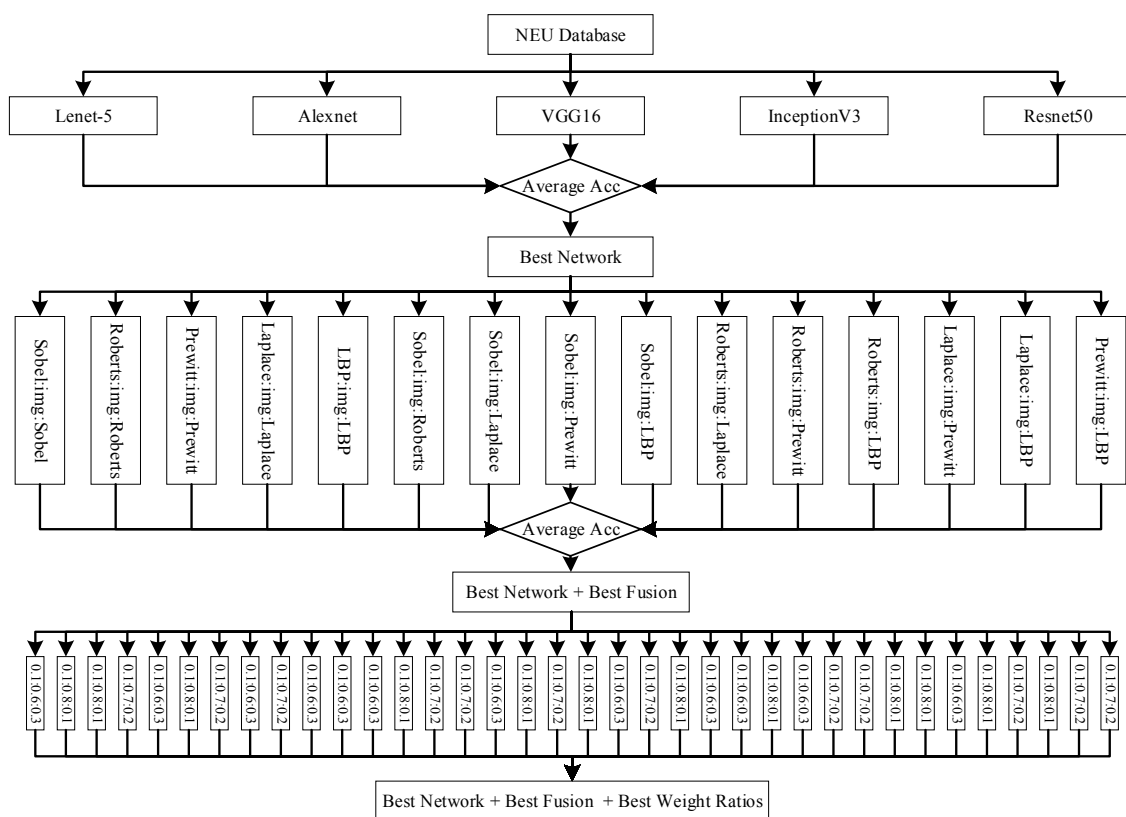


Figure 1. Overall flowchart of feature fusion–based preprocessing.

3. Experiments and analysis

In this study, the same CNN framework and hyperparameters are used for the training. The sensitivity of the feature fusion preprocessing model at the defect area is verified using a heat map and a feature activation map of the visual learning area. The accuracy (acc)–loss curve is used to observe the convergence of the model before and after feature preprocessing. The confusion matrix is used to assess the accuracy of network recognition before and after feature fusion preprocessing. The goal of the experiment is to find the optimal fusion scheme and conversion weight.

3.1. Experimental data

The NEU steel surface defect database created by Song et al. [22] of Northeastern University, China, was selected for use in this study. There are six categories of defects, namely, crazing, inclusion, patches, pitted, rolled-in, and scratches. There are 300 sample data points for each defect category, thus, resulting in a total of 1800 data points. The image resolution is 200×200 pixels. Figure 2 shows the defect samples.

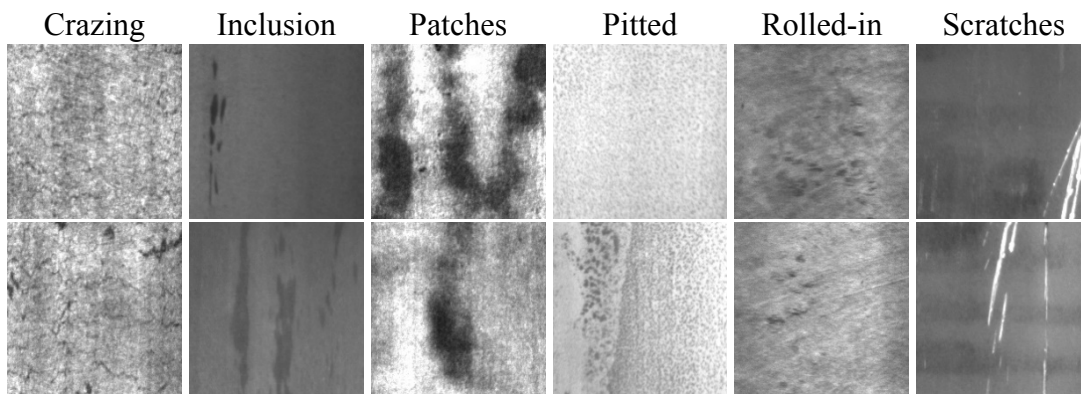


Figure 2. NEU steel plate defect data set.

Defect recognition based on the NEU steel plate surface defect database [23] faces three difficulties: (1) Defects of different classes have similar features; (2) illumination and material changes can affect the gray value of the acquired defect image; and (3) there are effects of various noises. To prevent overfitting, data augmentation is used in this study to expand the existing database capacity to 5000 samples, to increase the data volume of the experiment set, and to improve the generalization ability of the model. The experimental set in this paper is divided into three categories: Training set, validation set and test set, which are distributed according to the ratio of 7:2:1. So there are 3500 samples in training set, 1000 samples in validation set and 500 samples in test set. In order to ensure the consistency of the samples, the number of samples for each category in each sample set is roughly the same.

3.2. Experimental parameters

To fairly compare the data results before and after feature fusion processing, the exact same hyperparameters are used as follows: Adam, which is used as the optimizer; ReLU, which is used as the activation function; random initialization of weights; and use of the learning rate decay strategy with an initial learning rate of 0.00001, an epoch of 100, and a batch size of 32.

3.3. Experimental environment

All experiments were run on a graphics workstation with two 10-core Intel Xeon E5-260 Wv4 central processing units (CPUs), an NVIDIA Titan 1080Ti graphics processing unit (GPU), and 128 GB memory. Windows 10-based Python was the development environment, and Keras was used as the learning framework.

3.4. Experimental program and analysis

3.4.1. Determination of the network framework

Five classical models, Lenet-5 [24], Alexnet [25], VGG16 [26], InceptionV3 [27], and Resnet50 [28], are selected as the main framework of the CNN. Taking the original image data as samples, the recognition accuracies of these classical network models are compared using identical hyperparameters. The results of this comparison are shown in Figure 3. VGG16 has the highest recognition rate, up to 95.55%. Therefore, this model is the most suitable for learning the features of the NEU data. Therefore, VGG16 is chosen as the main framework.

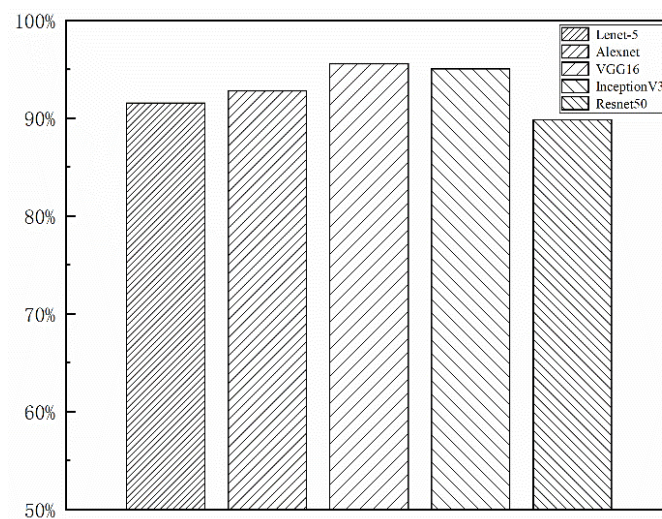


Figure 3. Recognition accuracies of different classical CNN frameworks.

3.4.2. Fusion of the same feature operators

To study the influence of the fusion of the same operators and the original image on the model, the image processed by each operator is selected and combined with the original grayscale image. The original grayscale data are placed in the middle channel, and the grayscale matrix is placed in the other two channels. The result is five different fusion schemes, as shown in Figure 4. Five sets of experiments (Nos. 1 to 5) are carried out for each fusion scheme by using the above hyperparameters for training. Hence, there are a total of 25 sets of experiments. The results are shown in Table 1. The fusion scheme of Sobel:image:Sobel has the highest average accuracy, reaching 96.22%.

3.4.3. Fusion of different feature operators

To study the influence of the feature operator combination scheme on the recognition accuracy of the model, the image processed by each pair of operators is selected and combined with the original grayscale image, and the original grayscale data are placed in the middle channel, while the grayscale matrix processed by the two operators is randomly placed in the other two channels. Ten combination schemes are based on channel fusion, as shown in Figure 5. The fused data are all in color, and the

inclusion areas of the images with edge operator fusion tend to be darker. The inclusion area of the image data processed by the LBP operator and the texture of the steel plate background are more prominent. Using the aforementioned model and hyperparameters, five sets of experiments (Nos. 1 to 5) were carried out for each fusion scheme; thus, a total of 50 experiments were carried out. The accuracy of the fusion results obtained using different operators is shown in Table 2.

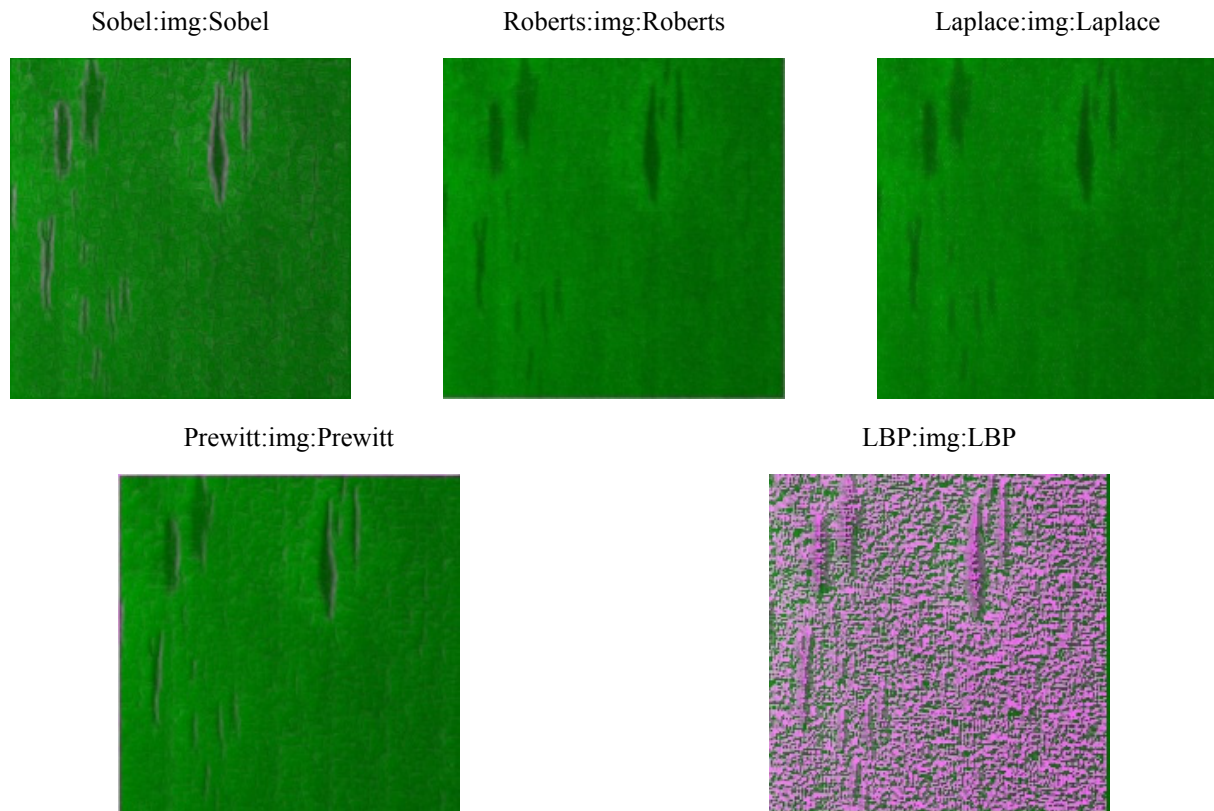


Figure 4. Fusion results with the same feature operators.

Table 1. Accuracy of fusion results with the same operators (percentage, %).

Conversion weights	Accuracy					Average
	No. 1	No. 2	No. 3	No. 4	No. 5	Accuracy
Sobel:img:Sobel	97.24	96.54	94.27	93.28	96.77	96.22
Roberts:img:Roberts	95.34	95.49	94.65	96.97	95.30	95.55
Prewitt:img:Prewitt	95.25	93.94	95.21	94.24	94.71	94.67
Laplace:img:Laplace	88.13	87.12	86.44	87.63	87.08	87.28
LBP:img:LBP	92.11	92.48	92.26	93.32	94.43	92.92

Table 2 shows that among the five sets of experiments, the fusion result of Sobel:image:Laplace has the highest average accuracy, reaching 98.61%. The fused images with the LBP operator generally have poor performance. In this case, the CNN performs poorly because although the CNN is more sensitive to the visual feature information of the image, the CNN is not good at analyzing the intrinsic meaning of the abstract texture features encoded by LBP.

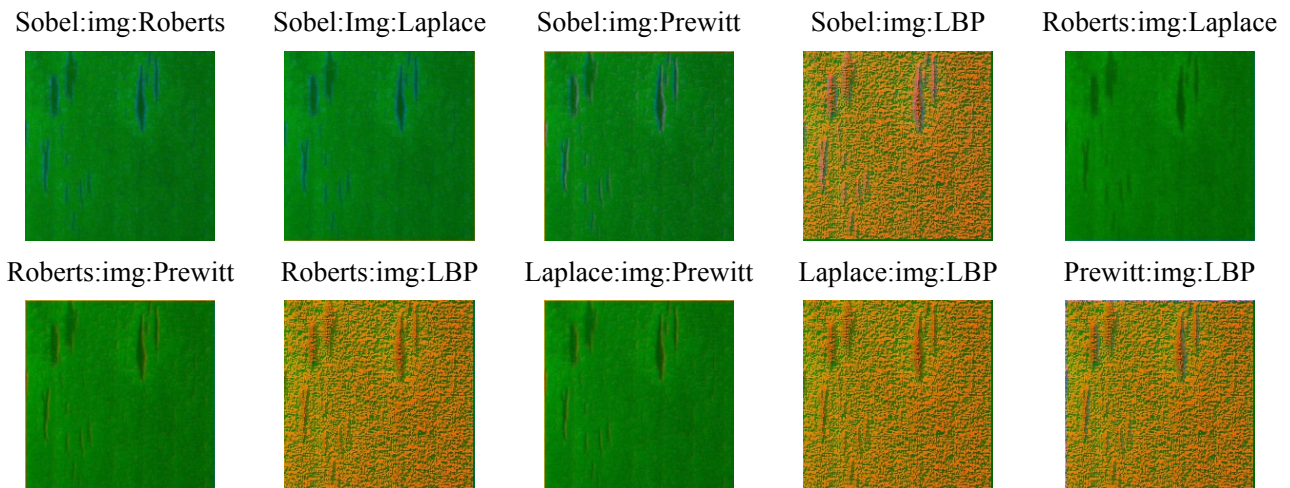


Figure 5. Fusion results by using different feature operators.

Table 2. Accuracy of fusion results by using different operators (percentage, %).

Fusion scheme	Accuracy					Average Accuracy
	No. 1	No. 2	No. 3	No. 4	No. 5	
Sobel:img:Roberts	98.88	97.36	98.05	99.09	97.97	98.27
Sobel:img:Laplace	97.51	98.14	98.98	98.91	99.45	98.61
Sobel:img:Prewitt	98.51	97.04	98.17	99.01	97.27	98.00
Sobel:img:LBP	76.11	76.25	75.22	74.89	75.58	75.61
Roberts:img:Laplace	98.02	96.94	96.01	98.57	99.61	97.83
Roberts:img:Prewitt	95.17	96.37	95.25	96.44	97.32	96.11
Roberts:img:LBP	81.21	81.37	83.27	81.65	80.25	81.55
Laplace:img:Prewitt	97.41	97.11	97.02	96.51	95.55	96.72
Laplace:img:LBP	70.52	67.61	65.84	64.32	67.26	67.11
Prewitt:img:LBP	19.84	14.28	15.77	16.19	17.22	16.66

3.4.4. Conversion to a single channel

Processing the three-channel data after fusion will double the computational cost, while the original grayscale image model only needs to process the single-channel data. In this study, to avoid this drawback after multichannel fusion, the three-channel data are converted into a single-channel grayscale image according to a certain weight ratio, as shown in Eq (1):

$$\text{Final image} = \alpha \text{ (1st channel)} + \beta \text{ (2nd channel)} + \gamma \text{ (3rd channel)} \quad (1)$$

where α , β , and γ are the weight ratio coefficients of different channels, and they sum up to 1. The traversal method is used with a step size of 0.1. And there is a total of 36 combinations. Based on the original model, the average accuracy is tested, as shown in Figure 6.

The results show that when the weight coefficients of α , β , and γ are 0.2, 0.6, and 0.2, respectively, the computational cost is the same as with the original data. However, the accuracy is 1.16% higher than when using the three-channel fusion. Figure 7 compares the single-channel image converted using a weight ratio of 0.2:0.6:0.2 with the original image. The figure shows that the defect feature edge of

the feature fusion image of the Sobel:image:Sobel algorithm is more obvious. The weight of 0.6 assigned to the original grayscale image can properly retain the information of the original image. The ability to retain this information helps the convolution kernel to learn weights, and the weight of 0.2 assigned to the edge extraction operator can well highlight the defect features on the steel plate image. Consequently, these defect features are conducive to CNN capturing. The same calculation parameters as those of the original data can be retained, while the recognition rate can reach as high as 99.77%.

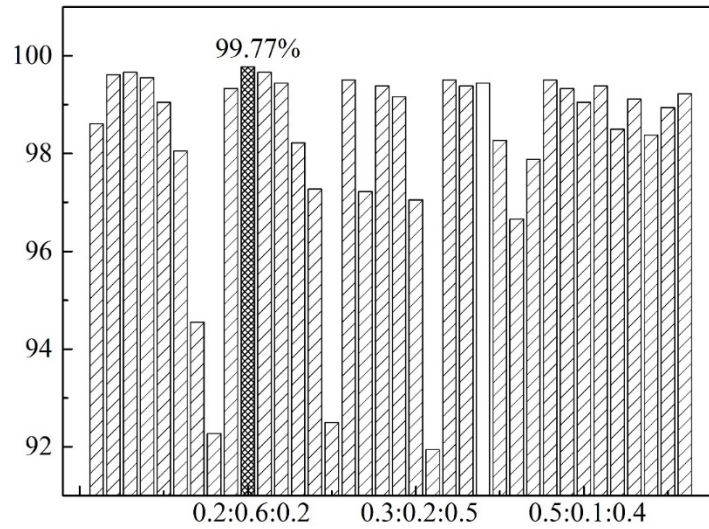


Figure 6. Accuracy using coefficients with different conversion weight ratios (percentage, %).

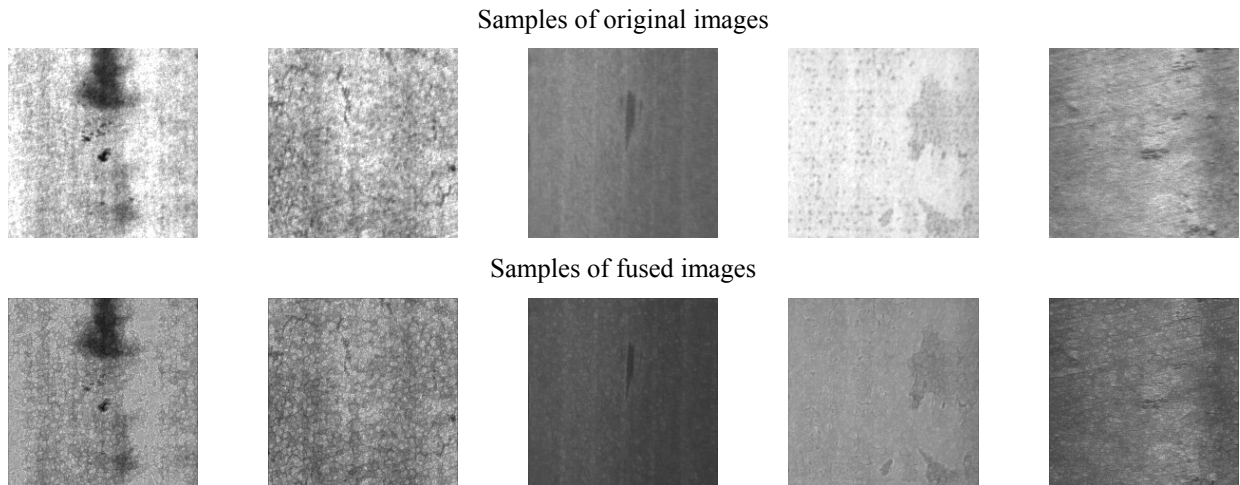


Figure 7. Grayscale images of the fusion results.

3.4.5. Comparison of models before and after multichannel fusion

The feature activation map of the first layer of VGG16 before and after the feature fusion processing is visualized, and the learning results are analyzed. One such image is randomly selected for processing (Figure 8). The yellow area indicates that the defect feature is activated and contains

the black inclusion defects; hence, it can be determined that this feature produces the correct phase response. Figure 8 also shows that, in the activation map without feature fusion, 23 feature maps in the defect area generate correct activation. In contrast, in the results after feature fusion, 31 feature maps in the defect area generate correct activation. These data indicate that because the edge extraction operator captures the target defect area in advance and fuses it with the original image, the edge extraction operator can enhance the edge pixel feature of the defect area of the image and ensure that the original data are not lost. Consequently, it is easier for the convolution kernel to capture the defect.

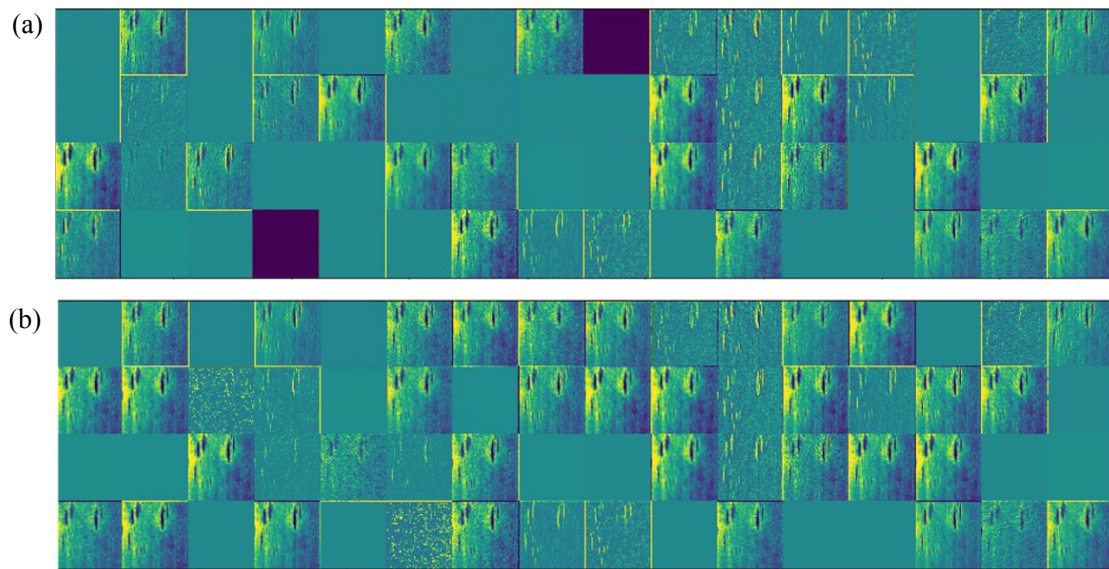


Figure 8. Visualization of the first activation layer. (a) Visualization of the feature map of the original grayscale image. (b) Visualization of the feature map of the fused image.

In this study, the gradient-weighted class activation mapping (Grad-CAM) algorithm [29] is used to further validate the target area where the model finally learns. This algorithm can display the area where the model learns in the form of a heat-affected zone to facilitate the observation of whether the final model successfully learns the features of the correct area. Given an input image, the algorithm can obtain the output feature map of a convolution layer through the trained model, which can obtain the parameters of the convolution kernel corresponding to each pair of feature maps without modifying the original model structure. The category relative to each channel of this feature map is used as a weight to generate a spatial map of the activation intensity of the input image to the category. So the heat-affected zone can be generated. In this paper, this method is used to visualize the model before and after feature fusion. Figure 9 shows the learning result of the inclusion defect. The figure shows that the defect area of the steel plate image data subjected to feature fusion process is red; i.e., the area is a “high-temperature” zone. This zone can be accurately located in the image, while the defect area of the original image data not processed by the feature operator is not very well covered. The results show that the data processed by the feature operator have an increased pixel level in the defect area; therefore, the area is more likely to be captured by the CNN. The heat map analysis further verifies that feature fusion preprocessing significantly improves the recognition ability of the CNN model.

The acc and loss curves of the model learning before and after feature fusion are plotted in Figure 10. This figure shows that the acc curve after feature operator processing has a narrow range

of fluctuation and overall larger fluctuation; in contrast, the acc curve without feature operator processing has a relatively wide range of fluctuation and overall smaller fluctuation. The loss curve after feature operator processing has a narrow range of fluctuation and overall smaller fluctuation; in contrast, the loss curve without feature operator processing has a wide range of fluctuation and overall larger fluctuation. It can be concluded that the model after feature fusion has a higher fitting ability and better convergence than the model without feature fusion.

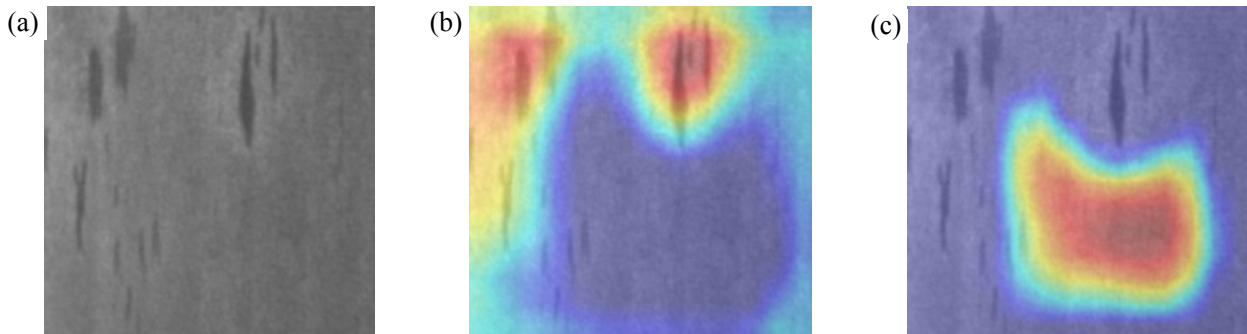


Figure 9. Visualization results of heat map analysis. (a) Original defect image. (b) Heat map after feature fusion. (c) Heat map without feature fusion.

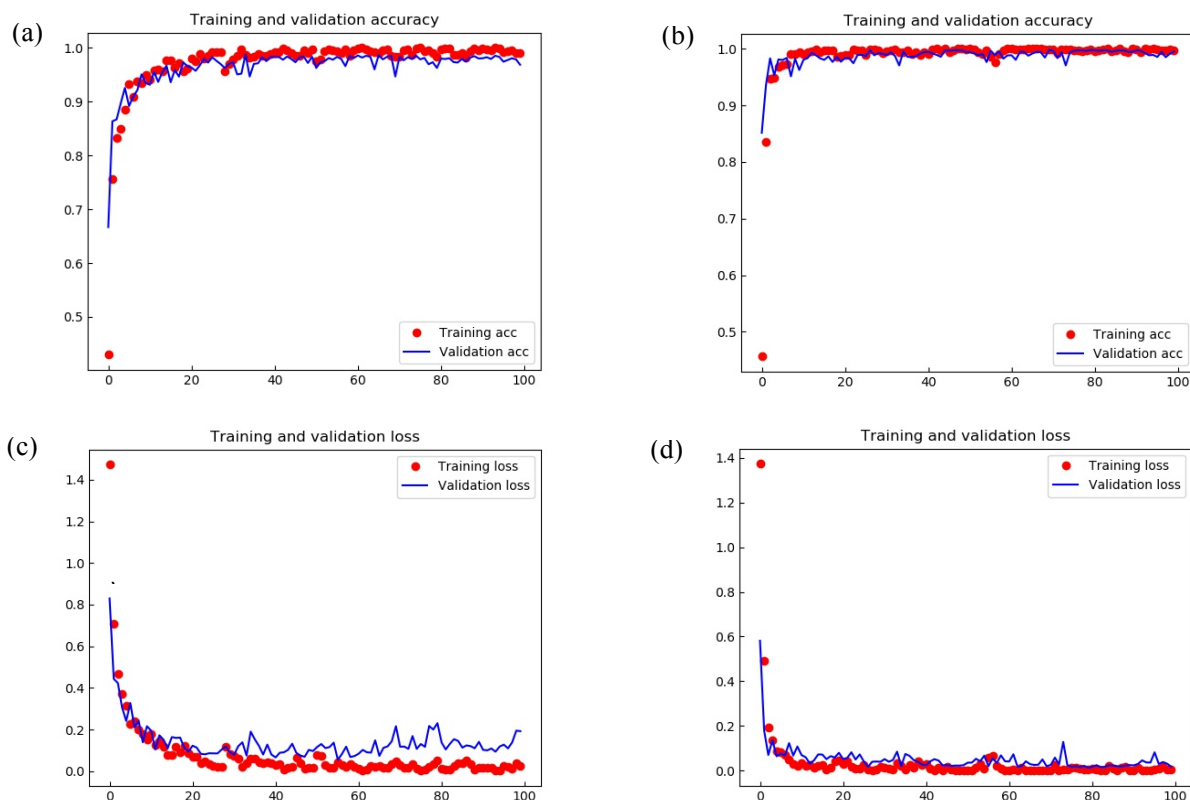


Figure 10. Comparison of results before and after fusion. (a) Acc curve before feature fusion. (b) Acc curve after feature fusion. (c) Loss curve before feature fusion. (d) Loss curve after feature fusion.

The confusion matrix is used in this study to further compare the classification results of the model before and after feature fusion. Each row of the matrix represents the results of a predicted category. The optimal results before and after feature fusion are represented by the confusion matrix, as shown in Figure 11. The figure shows that the recognition accuracy for the crazing category with a steel plate background is 100% with or without feature fusion; the recognition accuracy for the inclusion category is 99.33% with feature fusion and 97.58% without feature fusion; the recognition accuracy for the patches category feature is 100% with feature fusion and 98.48% without feature fusion; the recognition accuracy for the pitted category is 100% with feature fusion and 78.48% without feature fusion; the recognition accuracy for the rolled-in category is 100% with feature fusion and 95.76% without feature fusion; and the recognition accuracy for the scratches category is 99.33% with feature fusion and 95.15% without feature fusion. The above results show that the recognition accuracy of the model is higher in each category after preprocessing with feature fusion than without feature fusion. These results indicate the feasibility of preprocessing with feature fusion in improving the recognition accuracy.

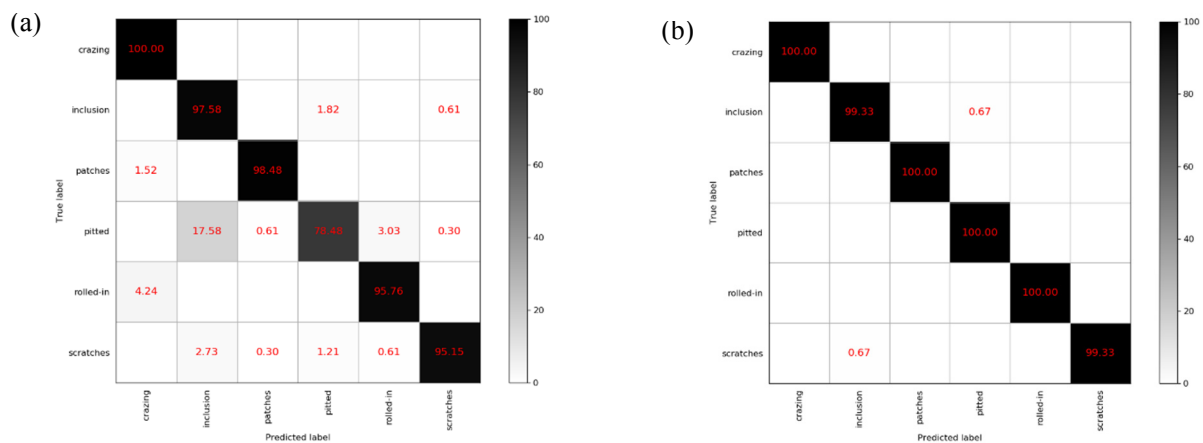


Figure 11. Results of confusion matrix accuracy. (a) Before feature fusion. (b) After feature fusion.

4. Conclusions

This study proposes an image preprocessing method based on feature operator fusion that is mainly used for defect recognition on low-contrast steel plate surface grayscale images. Through many combination tests, the original grayscale images have been processed using the Sobel and Laplace operators, then superimposed with the original grayscale image at a weight ratio of 0.2:0.6:0.2 as the input to a CNN. The experimental results show that this fusion scheme effectively improves the recognition rate of the test set. Without changing the experimental network framework, the accuracy rate can reach 99.77%, which is 4.22% higher than the accuracy rate of the CNN trained directly on the original images without fusion.

In this study, the influence of several commonly used operators on the accuracy of the model have been investigated, and a solution scheme from the perspective of optimizing the data sources has been offered to resolve the bottleneck problem of CNN learning. The results show that this feature operator-based channel fusion strategy is highly effective at improving the pixel features of

defects and making it easier for the network to capture the features of defects, thereby successfully enhancing classification accuracy.

Acknowledgments

The authors acknowledge the support from the National Key R&D Program of China (Grant No. 2018YFB1701601).

Conflict of interests

The authors declared that they have no conflicts of interest to this work.

References

1. D. He, K. Xu, P. Zhou, D. Zhou, Surface defect classification of steels with a new semi-supervised learning method, *Opt. Lasers Eng.*, **117** (2019), 40–48.
2. R. Gong, M. Chu, Y. Yang, Y. Feng, A multi-class classifier based on support vector hyper-spheres for steel plate surface defects, *Chemom. Intell. Lab. Syst.*, **188** (2019), 70–78.
3. B. Wu, J. Zhou, X. Ji, Y. Yin, X. Shen, Research on Approaches for Computer Aided Detection of Casting Defects in X-ray Images with Feature Engineering and Machine Learning, *Procedia Manuf.*, **37** (2019), 394–401.
4. X. Wen, K. Song, M. Niu, Z. Dong, Y. Yan, A three-dimensional inspection system for high temperature steel product surface sample height using stereo vision and blue encoded patterns, *Optik*, **130** (2017), 131–148
5. M. Kuffer, K. Pfeffer, R. Sliuzas, I. Baud, Extraction of slum areas from VHR imagery using GLCM variance, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **9** (2016), 1830–1840.
6. Y. Liu, S. Liu, Z. Wang, Multi-focus image fusion with dense SIFT, *Inf. Fusion*, **23** (2015), 139–155.
7. A. Zendejboudi, M. A. Baseer, R. Saidur, Application of support vector machine models for forecasting solar and wind energy resources: A review, *J. Cleaner Prod.*, **199** (2018), 272–285
8. I. Rish, *An empirical study of the naive Bayes classifier*, IJCAI 2001 workshop on empirical methods in artificial intelligence, 2001.
9. S. Zhang, X. Li, M. Zong, X. Zhu, R. Wang, Efficient knn classification with different numbers of nearest neighbors, *IEEE Trans. Neural Networks Learn. Syst.*, **29** (2018), 1774–1785.
10. G. Biau, E. Scornet, A random forest guided tour, *Test*, **25** (2016), 197–227.
11. C. He, H. Kang, T. Yao, X. Li, An effective classifier based on convolutional neural network and regularized extreme learning machine, *Math. Biosci. Eng.*, **16** (2019), 8309–8321.
12. L. Wen, Y. Dong, L. Gao, A new ensemble residual convolutional neural network for remaining useful life estimation, *Math. Biosci. Eng.*, **16** (2019), 862–880.
13. Z. Ning, Y. Feng, M. Collotta, X. Kong, X. Wang, L. Guo, et al., Deep Learning in Edge of Vehicles: Exploring Tri-relationship for Data Transmission, *IEEE Trans. Ind. Inf.*, **15** (2019), 5737–5746.
14. F. Chen, M. R. Jahanshahi, NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion, *IEEE Trans. Ind. Electron.*, **65** (2018), 4392–4400.

15. Y. He, K. Song, Q. Meng, Y. Yan, An end-to-end steel surface defect detection approach via fusing multiple hierarchical features, *IEEE Trans. Instrum. Meas.*, **69** (2020), 1493–1504.
16. Y. Wang, H. Xia, X. Yuan, L. Li, B. Sun, Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion, *Multimedia Tools Appl.*, **77** (2018), 16741–16770.
17. S. Gupta, S. G. Mazumdar, Sobel edge detection algorithm, *Int. J. Comput. Sci. Manage. Res.*, **2** (2013), 1578–1583.
18. W. Zheng, K. Liu, *Research on Edge Detection Algorithm in Digital Image Processing*, 2017 2nd International Conference on Materials Science, Machinery and Energy Engineering, 2017.
19. D. Adlakha, D. Adlakha, R. Tanwar, Analytical comparison between Sobel and Prewitt edge detection techniques, *Int. J. Sci. Eng. Res.*, **7** (2016), 1482–1485.
20. P. Vit, Comparison of various edge detection technique, *Int. J. Signal Process. Image Process. Pattern Recognit.*, **9** (2016), 143–158.
21. T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: Application to face recognition, *IEEE Trans. Pattern Anal. Mach. Intell.*, **28** (2006), 2037–2041.
22. K. Song, Y. Yan, A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects, *Appl. Surf. Sci.*, **285** (2013), 858–864.
23. K. Li, X. Wang, L. Ji, *Application of Multi-Scale Feature Fusion and Deep Learning in Detection of Steel Strip Surface Defect*, International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), 2019.
24. A. El-Sawy, H. EL-Bakry, M. Loey, *CNN for handwritten arabic digits recognition based on LeNet-5*, International conference on advanced intelligent systems and informatics (AISI), 2016.
25. A. Krizhevsky, I. Sutskever, G. E. Hinton, *ImageNet classification with deep convolutional neural networks*, Advances in neural information processing systems (NIPS), 2012.
26. A. Vedaldi, A. Zisserman, Vgg convolutional neural networks practical, *Dep. Eng. Sci. Univ. Oxford*, **2016** (2016), 66.
27. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., *Going deeper with convolutions*, Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 2015.
28. K. He, X. Zhang, S. Ren, J. Sun, *Deep residual learning for image recognition*, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
29. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, *Grad-cam: Visual explanations from deep networks via gradient-based localization*, Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.



AIMS Press

©2020 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)