Classification of the Steel Surface Defects via Machine Learning and Deep Learning

S. Y. Lee†, C. Park† and S. Lee*

Department of Mechanical Engineering, Pohang University of Science and Technology, Pohang, Korea

†Both authors contributed equally.
*Corresponding author: seunglee@postech.ac.kr

1. Introduction

In the steel industry, reducing surface defects is essential for securing the quality of steel plates and increasing reliability of the steel products. Detecting the defects of the steel surface must be preceded to analyze the causes of the defects and thus a reliable classification algorithm needs to be developed.

In this work, we applied traditional machine learning algorithms such as SVM (Support Vector Machine), Logistic Regression, Decision Tree to develop classification model and also utilized two kinds of ensemble algorithms, Random Forest and XGBoost (Extreme Gradient Boosting). We additionally built a model based on deep learning algorithm especially CNN (Convolutional Neural Network) in the expectation of better performance on accuracy and visualization of where the defects are. We validated all of our models through extensive experiments on NEU surface defect database where six kinds of typical surface defects of the hot-rolled steel strip were collected, i.e., rolled-in scale, patches, crazing, pitted surface, inclusion and scratches [1-2]. NEU dataset has 300 images for each class and we used 180 images for training, 60 images for validation and 60 images for test in each class.

2. Methodology

In carrying out traditional machine learning algorithms, extracting appropriate features from images directly affects model performance. Instead of using a single type of feature extraction method, two different approaches were conducted through GLCM (Gray Level Co-occurrence Matrix) and HOG (Histogram of Oriented Gradients) for building models, respectively. One is the approach that is widely used in extracting texture features from an image [3]. For example, GLCMs of defect images in 'crazing' can be distinguished from those in 'inclusion' as in Fig. 1. As a common usage of GLCM, we extracted six kinds of statistical features ($\mathbf{F_G}$) from the matrix, i.e., ASM (Angular Second Moment), energy, contrast, dissimilarity, homogeneity and correlation. The other is the feature descriptor yielding histograms that contain high-dimensional texture information and can be visualized distinctively among classes as shown in Fig. 2. While using the entire histograms that are extracted by HOG, optimal 60 components of the features ($\mathbf{F_H}$) were determined via PCA (Principal Component Analysis), one of the most frequently used for dimension reduction.

Moreover, we concatenated the two kinds of features into a new feature ($\mathbf{F_C} = \mathbf{F_G} \cup \mathbf{F_H}$), expecting better performance than both models using only one of the features, $\mathbf{F_G}$ and $\mathbf{F_H}$.

Using CNN, we built a VGG-like CNN model which had three convolutional blocks and each block consisted of two convolutional layers and a max-pooling layer. For the classification, we used two fully connected layers with dropout ratio of 0.3 and six output nodes. For the model comprehension, we also visualized an attention map to find out the location of the defects and to interpret how deep learning model discriminates classes using Grad-CAM [4].

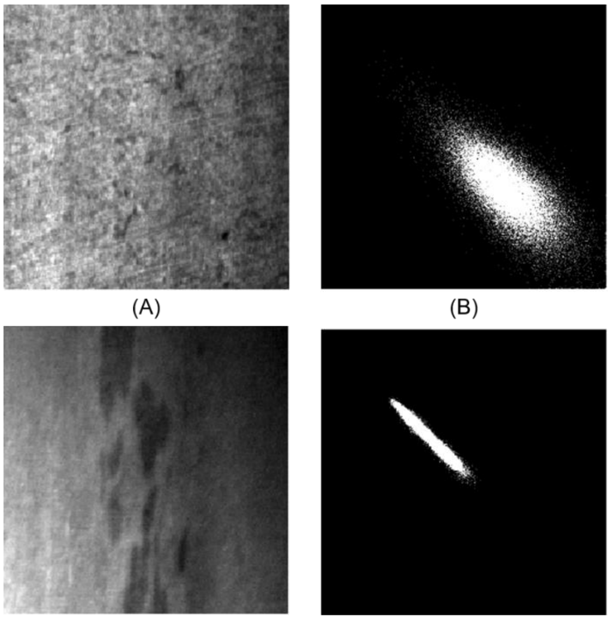

Fig.1 Example of GLCM on defect images. (A) a defect image in 'crazing'; (B) GLCM of 'crazing' example; (C) a defect image in 'inclusion'; (D) GLCM of 'inclusion' example.

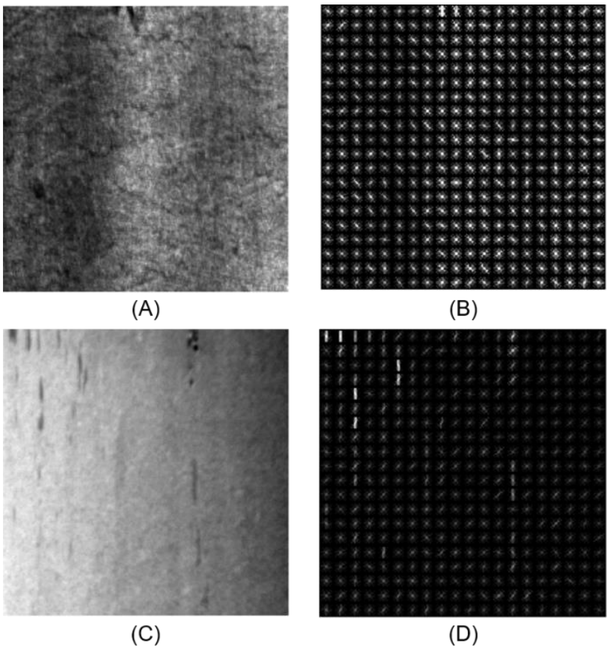

Fig.2 Example of HOG visualization on defect images. (A) a defect image in ‘crazing’; (B) HOG visualization of 'crazing’ example; (C) a defect image in ‘inclusion’; (D) HOG visualization of ‘inclusion’ example.

3. Results

Among the traditional machine learning models, we found that that a XGBoost model using the combined feature ($\mathbf{F_C}$) showed the best performance with 95.56 % accuracy as shown in Table 1. Using the combined feature ($\mathbf{F_C}$) resulted in an improvement of approximately 3 percent point in accuracy across the entire machine learning models.

As shown in Table 1, the CNN model showed remarkable performance on accuracy compared to the traditional machine learning models. Even though the model misclassified few images, the predicted probability of belonging to each class had a second-highest value in the true class whereas it had extremely low values in the other classes. Furthermore, as in Fig. 3, we found that the attention map which was obtained by Grad-CAM activated the defective part of the image in red, which means that the CNN model can not only classify detects, but also automatically identify their locations.

Table 1 Test accuracy (%) of each model

	$\mathbf{F_G}$	$\mathbf{F_H}$	$\mathbf{F_C}$
SVM	88.06	78.61	92.22
Log. Regression	85.83	73.89	89.44
Decision Tree	90.83	80.28	91.94
Random Forest	92.22	88.06	95.28
XGBoost	92.50	87.78	95.56
CNN		99.44	

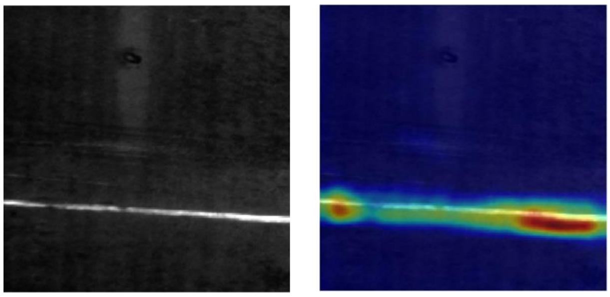

Fig.3 Example of attention map on a defect image. (left) a defect image in ‘scratches’; (right) attention map of 'scratches’ using Grad-CAM.

Acknowledgment

This research is funded by Korea Institute of Machinery & Materials, Ministry of Trade, Industry and Energy, Korea Evaluation Institute of Industrial Technology (KEIT, 10080729).

References

[1] K. Song and Y. Yan, A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects, *Applied Surface Science*, 285 (B) (2013) 858-864.

[2] Y. He, K. Song, Q. Meng, and Y. Yan, An end-to-end steel surface defect detection approach via fusing multiple hierarchical features, *IEEE Transactions on Instrumentation and Measurement*, (2019).

[3] G. Saurabh, Use machine learning to detect defects on the steel surface, *Intel AI Developer Program Documentation*, (2018).

[4] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, Grad-cam: visual explanations from deep networks via gradient-based localization, *IEEE International Conference on Computer Vision*, (2017).