

Defect Diagnosis and Classification of Machine Parts Based on Deep Learning

Hyun-Tae Kim¹, Sang-Hyeop Lee², Sheilla Wesonga³, Jang-Sik Park^{4*}

〈Abstract〉

The automatic defect sorting function of machinery parts is being introduced to the automation of the manufacturing process. In the final stage of automation of the manufacturing process, it is necessary to apply computer vision rather than human visual judgment to determine whether there is a defect. In this paper, we introduce a deep learning method to improve the classification performance of typical mechanical parts, such as welding parts, galvanized round plugs, and electro galvanized nuts, based on the results of experiments. In the case of poor welding, the method to further increase the depth of layer of the basic deep learning model was effective, and in the case of a circular plug, the surrounding data outside the defective target area affected it, so it could be solved through an appropriate pre-processing technique. Finally, in the case of a nut plated with zinc, since it receives data from multiple cameras due to its three-dimensional structure, it is greatly affected by lighting and has a problem in that it also affects the background image. To solve this problem, methods such as two-dimensional connectivity were applied in the object segmentation preprocessing process. Although the experiments suggested that the proposed methods are effective, most of the provided good/defective images data sets are relatively small, which may cause a learning balance problem of the deep learning model, so we plan to secure more data in the future.

Keywords : Machinery Parts, Defect of Welding, Circular Plug, Electro Galvanized Nuts, Deep Neural Network

1 Dept. of Applied Software Eng., Dongeui University, Professor
E-mail: htaekim@deu.ac.kr

2 Dept. of Electronic Eng., Kyungsung University, Student
E-mail: soundslock05@naver.com

3 Dept. of Electronic Eng., Kyungsung University, Student
E-mail: sheilla2@ks.ac.kr

4* Corresponding Author, Dept. of Electronic Eng., Kyungsung University, Professor
E-mail: jsipark@ks.ac.k

1. Introduction

Recently, automation of the manufacturing process is rapidly spreading centering on smart factories, and in the manufacturing field, the automatic selection of abnormal conditions, especially for mechanical parts, is included in the automation of the manufacturing process. Most of the mechanical parts have been classified as abnormal conditions by hand, but the rate of misclassification of normal conditions into abnormal conditions or vice versa was not low. In this paper, there are numerous types of product groups of mechanical parts and various types of defects appearing in these product groups, but among them, focusing on the examples directly implemented through experiments on the method of determining the metal parts welding defect, the circular plug plating defect in the automobile steering system, and the plating defect of hexagon nuts in connection with the image-based and deep learning methods will be introduced.

2. Defective Diagnosis of Machinery Parts

Defects occurring in the welding process are largely divided into three types. These are dimensional defects, structural defects, and property defects. Among them, the most problematic part is a structural defect, which

mainly occurs due to mechanical and metallurgical causes. Radiation inspection, ultrasonic inspection, magnetic inspection, and fluorescence inspection have been applied as non-destructive inspection methods for welding defects, but optical-based imaging inspection methods are still insufficiently applied. However, thanks to the recent spread of deep learning technology, it can respond to various variables (lighting, color change, curvature, field of view, etc.) that were not possible with existing machine vision systems, and a possible stage has been reached in assessing the variability and deviation between parts that are visually very similar[1-2].

Recently, many electroplating products are being produced to extend the durability and lifespan of mechanical parts for vehicles. Typical mechanical parts include bolts and nuts. The types of defects in nuts include; scratches, non-plating, foreign matter contamination, and discoloration. Electroplating products have a high surface reflectivity due to plating, so camera image quality deteriorates due to the surrounding environment, including reflections from ambient lighting. It is necessary to acquire an input image that reduces the influence of the external environment by adjusting the camera and lighting. Even if the camera and lighting are adjusted, the highly reflective part of the nut surface appears in various areas depending on the location of the camera lighting, and if it is binarized, it is indistinguishable from the background area. Therefore, an appropriate preprocessing

process is required for deep learning analysis[3]. Various studies have been conducted on the preprocessing technique, and it was mainly applied to solve the problem of reflection by light or luminance, and interference of other objects such as people and other objects [4-5].

In this paper, we propose a method for discriminating between normal and abnormal nuts through deep learning, a nut object is separated and detected from an input image through segmentation pre-processing, which sequentially processes two-dimensional connectivity and mathematical behavioral filtering for camera images. In the object segmentation preprocessing process, a nut is segmented and detected using the connectivity information of 2D pixels in the steps of converting a gray scale image to a binary image and in the step of extracting the object[6]. Since the separately detected nut is a divided nut object that does not include a background, performance can be improved by reflecting only on the nut region that does not include a background in the training and inference. It compares and analyzes performance by applying various deep learning to the segmented and detected objects. For the split- detected nut image, VGG-16 increases the non-linearity three times as the hidden layer deepens compared to the comparison target AlexNet[7], decision boundaries with non-linear features[8], which are the strengths of deep learning, can be classified better, and thus have better classification performance.

3. Proposed Method of Defective Diagnosis of Machinery Parts

3.1 Object Segmentation Preprocessing

The deep learning classification performance is improved by performing object segmentation preprocessing that detects a nut from an image taken by a camera and removes a background image from the detection area[9]. The object segmentation preprocessing proposed in this paper sequentially processes 2D connectivity and mathematical behavioral filtering. The object segmentation preprocessing process sequentially converts the input RGB color image into a gray scale image and a binary image so as not to be affected by lighting vegetation. In the process of converting a grayscale image to a binary image, connectivity is processed to extract only the nut region to obtain a mask image. Fig. 1 shows the object segmentation preprocessing process to acquire the color image, grayscale image, binary image, and mask image proposed in this paper.

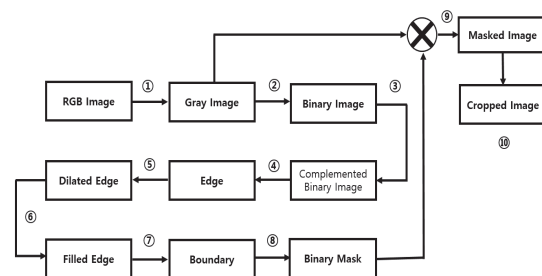


Fig. 1 The preprocessing for getting masked image

3.2 Residual Error Based Convolutional Neural Network (Resnet-50)

ResNet-50 is a convolutional neural network announced in 2015 and consists of learning about the difference between output and input, unlike existing learning methods. A ResNet structure usually consists of several convolutional layers and several fully-connected layers, between which a pooling layer, an activation layer, and a batch normalization layer is included[10]. In the convolution layer, a 3D convolution operation is performed on input activation values called an input feature map to generate an output feature map, that is, output activation values. The fully connected layer views the input activation values as one vector and performs a matrix-vector multiplication operation. In the integration layer, a single output activation value is generated by using several input activation values. The integration layer has the effect of reducing the number of activation values and coping with changes in neural network input size, etc. The activation layer is a layer that calculates a non-linear function, and serves to impart non-linear characteristics to the neural network, to get one output activation, apply a nonlinear function to one input activation. As a nonlinear function, a sigmoid function or a tanh function has been used a lot in the past, and the following rectified linear unit (ReLU) function is frequently used in recent convolutional neural networks. Finally, the

batch normalization layer plays a role in stabilizing learning by normalizing the input activation values of the convolutional layer or the fully connected layer[11].

4. Experimental Results and Discussions

4.1 Experiment Environment

Fig. 2 is sample photos of normal and abnormal conditions related to welding. The provided images were 53 normal models and 47 abnormal models. Since the amount of data is relatively small, these images were rotated by 90, 180, and 270 degrees respectively,

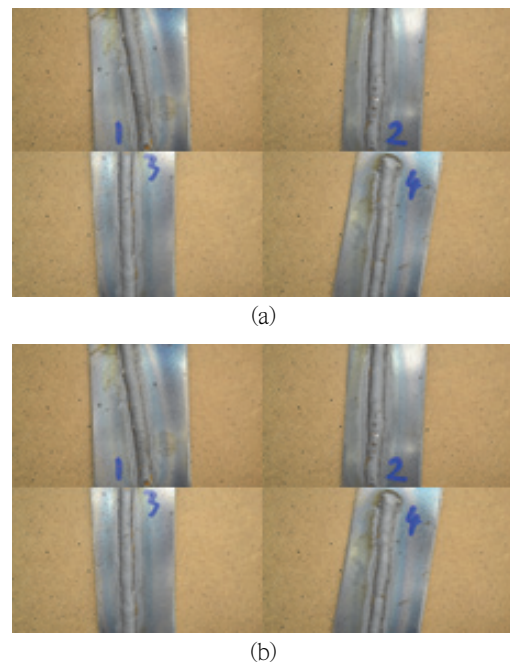


Fig. 2 The example of welding model image
(a) normal condition, (b) abnormal condition

and the results including the original image were symmetrical and 424 normal models expanded to 376 abnormal models and built a data set of 800 in total. Of these, 400 were used for deep learning training, 200 were used for verification, and the remaining 200 for testing, of the 200, 94 are normal and 106 are abnormal. Based on these data, a comparison experiment was performed with a different kernel size affecting the defect detection performance targeting “VGG-16”, a representative deep learning model[12]. As is well known, the kernel size of VGG-16 is fixed to 3×3 , but the depth of the layer may be insufficient to determine welding defects thus a comparative experiment was carried out by expanding to 7×7 , 9×9 , 11×11 [13].

The system configured for image-based abnormal condition classification for hexagon nuts. There are a total of 4 cameras installed in the sorting system, and cameras 1 to 3 are installed at an interval of 120° from each other, camera 4 is installed in the vertical direction from under the glass plate,

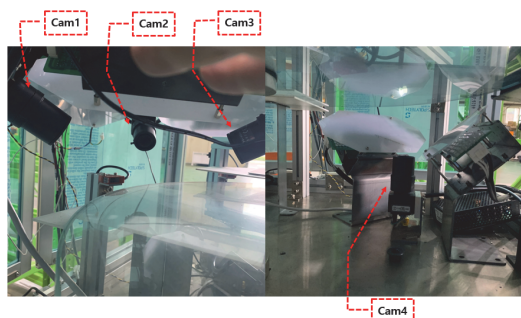
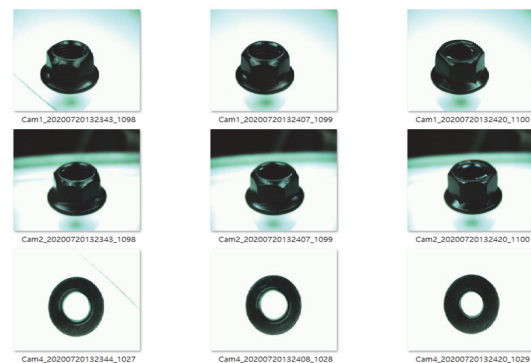


Fig. 3 The location of the mounted cameras in the discrimination system

and acquires an image of the bottom surface of the nut. Fig. 3 shows the camera installation. Images are acquired from 3 cameras out of 4 cameras, and camera numbers are set for each in order. The location of the camera is shown in Fig. 3.

Fig. 9 is an image obtained from cameras 1, 2, and 4 in Fig. 3. Fig. 4(a) is an example of a normal condition and (b) is an example of abnormal condition. The abnormal condition in Fig. 4(b) mainly show defects due to non-diameter plating, color abnormality, scratches, and foreign substances.



(a)



(b)

Fig. 4 The example of hexagon nuts image (a) normal condition, (b) abnormal condition

4.2 Experiment Results

The images used in the experiment are images of electroplated nuts used as parts of mechanical devices, and there are a total of 367 images, of which 180 are of normal conditions and 187 are of abnormal conditions. The ratios of images used for training, validation, and testing (inference) are shown in Table 1 as 60%, 20%, and 20%, respectively. “Cam1, Cam2, Cam3” and “Bottom” mentioned in Table 1 refer to the camera that acquired the image and the location of the normal and abnormal conditions classification system. The hardware platform used for training is Intel Xeon Gold 5118, 128GB memory, NVIDIA TITAN V 12GB GPU, and utilizes the Pytorch deep learning platform. The computer specifications used for the test are Intel i7-6700HQ, 16GB memory, NVIDIA GTX 1060 6GB GPU.

The size of the learning input image of deep learning is $224 \times 224 \times 3$, the learning model is VGG-16, the epoch is 1,000, the batch size is 20, the optimization algorithm is

Adam, and the initial learning rate is It is 0.001. The accuracy is calculated as an indicator for performance evaluation of the classification model, and the accuracy calculation formula is as Equation (1).

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100 \quad (1)$$

Where, TP (True Positive) is the number of images detected or recognized as True for the correct answer that is actually True, FN (False Negative) is the number of images detected or recognized as False for an actual True answer, FP (False Positive) is the number of images detected or recognized as True for a correct answer that is actually False, TN (True Negative) is the number of images detected or recognized as False for a correct answer that is actually False[14].

The kernel size was 11×11 , we found it to be the best in the determination of abnormal conditions, we also confirmed that there was little difference compared to the case in which the kernel size was small. The reason why the kernel size had an effect in the case of abnormal condition determination is that the expansion to a deeper layer through an increase in the kernel size has some effect in extracting the features of fine defects.

The experimental results are shown in Table 2. In Table 2, in the case where the preprocessing process was not applied and in the comparative experiment, the abnormal condition showed similar performance. the performance of normal condition is lower than when pre-processing is applied, especially when all four cameras including the floor are

Table 1. Configuration of image dataset

Class		Train (60%)	Validation (20%)	Test (20%)	Sum
Bottom	NG	55	18	18	91
	OK	59	20	20	99
	Sum	114	38	38	190
Cam1, 2, 3	NG	58	19	19	96
	OK	49	16	16	81
	Sum	107	35	35	177
Bottom+Cam1, 2, 3	NG	113	37	37	187
	OK	108	36	36	180
	Sum	221	73	73	367

Table 2. Classification results

Dataset	Per class accuracy		Remark
	NG	OK	
Camera Position	NG	OK	-
Bottom_original	94.44	100	No Pre-processing
cam1,2,3_original	94.74	87.5	
Bottom+cam1,2,3_original	100	0	
Bottom_alpha Off	100	90	Proposed Pre-processing
Bottom_white BG	94.44	90	
cam1,2,3_alpha Off	89.47	100	
cam1,2,3_white BG	78.947	93.75	
Bottom+cam1,2,3_white BG	91.89	91.67	

used, the result is not decisive at all. And in the case of using all four images with a white background in the proposed method, the accuracy of both abnormal and normal condition identification is more than 90%, and the accuracy is similar to each other. And the best performance is when only the floor image excluding the alpha value is determined as a target, but since defects may occur on surfaces other than the floor surface, determining only the floor surface has a problem in that it is not realistically valid.

5. Conclusion

In this paper, there are many types of product groups of machine parts and various types of defects appearing in these product groups. However, among them, we introduced based on the contents of experiments conducted in conjunction with image-based and deep learning methods, focusing on poor

welding between metal parts, poor plating of circular plugs in automobile steering system, and poor plating of hexagon nuts. Fine welding defects occurring in welding between metal parts showed better discrimination performance when the deep learning model enlarged the kernel size and gave the effect of composing a deeper layer. In the case of a circular plug, when the round border outside the minimum inner diameter is removed by a preprocessing method using a Circle Hough Transform and linked with a deep learning model, it shows better performance than when linked with a deep learning model while leaving the round border as it is. We concluded that a strategy is needed to prevent unnecessary parts from being included as feature values in the convolutional network of the deep learning model. Finally, in the case of the hexagon nut, it was concluded that the pre-processing technique that removes the effect of lighting, etc. from images acquired from multiple cameras from multiple angles was effective due to the 3D shape. However, most of the images of normal or abnormal conditions provided by demanding companies are relatively small, which can cause problems with the learning balance of the deep learning model. Based on the sufficiently secured data in the future, we intend to conduct more reliable research and more objective performance evaluation.

Acknowledgement

This research was supported by the BB21plus

funded by Busan Metropolitan City and Busan Institute for Talent & Lifelong Education(BIT)

References

- [1] Hansen, E. B., and Bøgh, Sg, “Artificial intelligence and internet of things in small and medium-sized enterprises: A survey,” *Journal of Manufacturing Systems*, Vol. 58, No. 2, pp. 368-372, 2021.
- [2] Zhao, Yong Jie, et al., “Automatic and accurate measurement of microhardness profile based on image processing,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-9, 2021.
- [3] Pang, Guansong, et al. “Deep learning for anomaly detection: A review”. *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1-38, 2021.
- [4] L. Liu, Ouyang, W. Wang, X. Fieguth, P. Chen, J. Liu, and M. Pietikainen, “Deep learning for generic object detection: A survey.” *International journal of computer vision*, Vol.128, pp. 261-318, 2019.
- [5] M. Al-Garadi, A. Mohamed, A. Al-Ali, A. Du, X. Ali, and M. Guizani, “A survey of machine and deep learning methods for internet of things (IoT) security.” *IEEE Communications Surveys & Tutorials*, 2020.
- [6] Cruttwell, Geoffrey SH, et al., “Categorical foundations of gradient-based learning,” *European Symposium on Programming*, pp. 1-28, 2022.
- [7] D. Park and K. Jun, “CHT-based Automatic Go Recording System under Illumination Change and Stone Dislocation,” *Journal of the Korean Institute of Information Scientists and Engineers*, vol. 41, no. 6, pp. 448-455, 2014.
- [8] Beyer, Lucas, et al., “Are we done with imagenet?,” *arXiv preprint arXiv:2006.07159*, 2020.
- [9] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao., “Scaled-yolov4: Scaling cross stage partial network.”. *IProceedings of the IEEE/cvf conference on computer vision and pattern recognition*. p.13029-13038, 2021.
- [10] Touvron, Hugo, et al. “Going deeper with image transformers.” *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.
- [11] H. J. Kang, “Efficient Fixed-Point Representation for ResNet-50 Convolutional Neural Network,” *Journal of the Korea Institute of Information and Communication Engineering*, Korea, vol. 22, no. 1, pp. 1-8, May 2018.
- [12] Bai, X., Wang, X., Liu, X., Liu, Q., Song, J., Sebe, N., and Kim, B., “Explainable deep learning for efficient and robust pattern recognition: A survey of recent developments.”, *Pattern Recognition*, vol. 120, 2021.
- [13] T. Jin, “Feature Extraction Using Convolutional Neural Networks for Random Translation,” *Journal of the Korean Society of Industry Convergence*, vol. 23, no. 3, pp.515-521, Jun. 2020.
- [14] S.-Y. Choi, “A Study on the Automation of Cam Heat Treatment Process using Deep Learning,” *Journal of the Korean Society of Industry Convergence*, vol. 23, no. 2_2, pp. 281-288, Apr. 2020.

(Manuscript received March 17, 2022;

revised April 4, 2022; accepted April 5, 2022)