

# Development of Visual Inspection System for Assembly Machine

Jeonghong Kim  
School of Computer Science and Engineering,  
Kyungpook National University,  
Daegu, Korea,  
jhk@knu.ac.kr

**Abstract**—In a factory that assembles various devices, functional inspection and visual inspection are carried out before the product is shipped out. Most functional tests are done automatically through the system, whereas visual inspection is done by the human eyes. Human remember what they have just seen before, and it is difficult to detect the failure of a similar assembling device. Even if the function of the assembling device is satisfied, there is a problem in the reliability of the quality control due to the appearance error. Various image processing techniques have been developed to solve this problem. In this paper, we proposed a method to improve the accuracy of error detection by applying the CNN model to the external inspection of the assembling device.

**Keywords**—Convolution Neural Network, Image processing, Image classification, Machine Learning

## I. INTRODUCTION

By applying information processing and image processing technology to the assembly plant, productivity and quality have been improved. The visual inspection of the assembling device is carried out after the functional inspection is completed. Functional testing of the assembly device is mostly automatic, while visual inspection depends on the human vision. However, human beings have difficulty in detecting a defect due to the effect of the afterimage that the brain has memorized what it had just seen. For this reason, researches on error detection have been in the field of computer vision for a long time.[1,2] Methods for detecting defects through various processes on the image include a method of using information such as contrast, brightness, distribution of images, and pixel groups by considering surrounding pixels. However, these methods require a program change according to the shape in order to be applied to various types of assembling apparatuses.

Recently, a CNN model has been developed, which has the advantage of recognizing various types of patterns by imitating the visual processing of animals. The CNN model consists of two core structures, convolutions and sub sampling. It decreases the complexity of the entire model by effectively reducing the number of parameters each stage.[3, 4, 5] In this paper, we propose a method to improve the accuracy of error judgment by applying the CNN model to the external inspection of the assembling device and to use it without changing the same program on various assembling devices.

---

This research is supported by Korea Technology and Information Promotion Agency for SMEs.

## II. RELATED STUDIES

In order to detect the error in the image, the template matching and the rule-based method were used for the vision inspection by using the information such as the brightness of the pixel or edge of the image. Template matching uses image subtraction and correlation between error-free and test images. The rule-based method extracts feature values from images, constructs description rules using these feature points, and applies them to test images. The features used in the rule description include the position of the object center point, the slope of the contour line, and the spacing between objects. The template matching method is simple because it compares the values of pixels between two images, but preprocessing such as direction and size adjustment is necessary before comparing the test image with the template.

Since the rule-based image processing must detect feature points in an object and compare them with the rules, the algorithm is somewhat complicated and takes a long processing time. However, the rule-based image processing is advantageous in that it is relatively accurate for images of various sizes and little noise. The figure below shows a typical template matching algorithm.

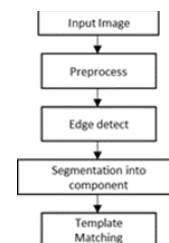


Fig. 1. A typical template matching algorithm

In recent years, with the development of artificial intelligence technology, image processing technology through machine learning has been developed. Convolution Neural Networks (CNN) is the neural network method most commonly used for image recognition.[6] CNN is an integrated model that combines the two stages of feature extraction and classification in existing pattern recognition methods. CNN is composed of different functional layer structure, convolutional layer, subsampling layer and fully-connected layer. We can add many

new layers to CNN for the process of evolution and improvement. Figure 2 shows typical CNN.

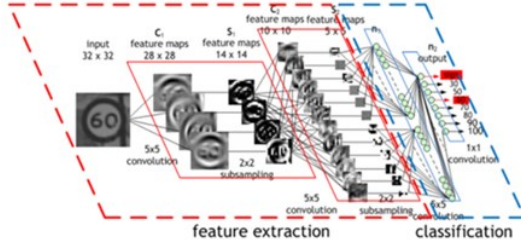


Fig. 2. Typical CNN structure

The subsampling layer reduces the complexity of the problem by reducing the data for a particular map created. Operations performed in the subsampling layer include a max subsampling operation taking a maximum value and an average subsampling having an average value. CNN increases the concept of receptive field and shared weights [8], which reduces the parameters of training and also reduces the complexity of network model. The features of each layer are generated from previous layer's local area (receptive field) by sharing the weight of the convolutional kernel. These characteristics make the CNN more suitable for the learning and represent of image features than other neural networks, and it can also keep the translation and scale invariance to a certain extent.

Image Classification groups and divides the region having similar feature values by classifying the pixel values of the image into similar classes. One of the image classification methods is supervised classification. [7, 8] It is a method to classify image data according to the characteristics of data by analyzer designating classification items and training data for each classification item. In the supervisory classification, classification items and training data for each classification item are selected by the analyst in advance, and based on the characteristics of the selected data, by discriminating which data is most similar to each pixel. Therefore, for the classification, the image data should be analyzed and the classification item should be set first.

On the other hand, unsupervised classification is a method to classify only image data by wavelength and spatial characteristics without prior information such as sample data. This method, often called clustering, requires only minimal input data and allows the computer to autonomously select the mean and covariance of the classes to be used for classification. K-means algorithm and isodata algorithm are representative methods of such grouping.

### III. DEFECT DETECTION SYSTEM FOR ASSEMBLY MACHINE USING CNN MODEL

Appearance inspection of the assembling device is whether or not the screw and the label are properly attached. The system for detecting assembly defects is largely divided into image learning and error detection. In the image learning process, the collected data is trained through CNN. After the training process, the test data to be newly recognized was inputted and the detection defect was judged.

The input image is some area of the assembling device image that includes screw in the assembly machine. For defect detection, CNN model is designed through a fully connected layer multi-layer network after two convolutions and subsampling.

The first convolution uses six  $7 \times 7$  convolution matrices to construct six  $30 \times 30$  feature maps. And then half-subsampled from this map to construct six  $15 \times 15$  subsampling maps. The  $8 \times 8$  convolution matrix is used for these six subsampling maps to obtain  $20 \times 4 \times 4$  feature maps. Through this process, the output was finally calculated through two RBF (Radial Basis Function) neurons. The cost function for hypothesis for binary classification in defect detection has values of 0 and 1. In Equation 1,  $W$  is a weight value,  $H(x)$  is a hypothesis function for input  $x$ , and  $y$  is a value for input  $x$ .

$$\text{cost}(H(x), y) = \begin{cases} -\log(H(x)) & : y = 1 \\ -\log(1 - H(x)) & : y = 0 \end{cases} \quad (1)$$

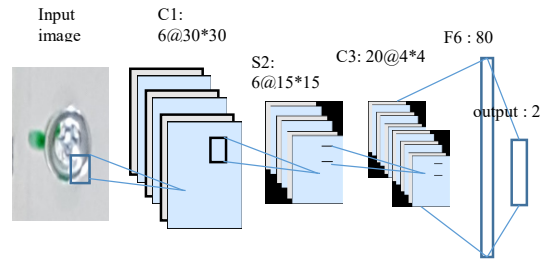


Fig. 3. CNN configuration diagram for defect detection.

### IV. EXPERIMENTS AND RESULT

CNN, which is proposed for defect detection, is implemented using tensor flow, an open source software library for machine learning used in Google products. The system uses the Python language in windows 10. The figure below shows the training and test data for the assembly device and the screw. Fig. 4 shows the photo of assembly machine for error detection. The training data and test data are in Fig. 5.

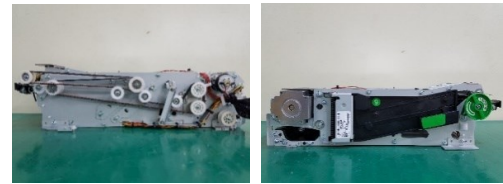


Fig. 4. Photo of assembly machine

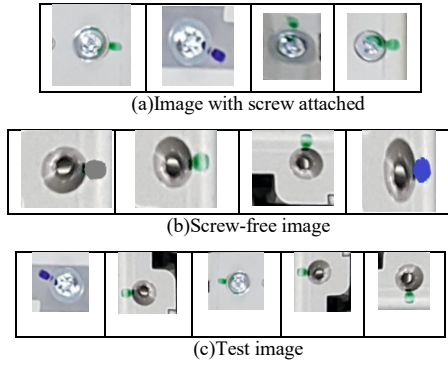


Fig. 5. Training and test data set

True Positive(TP) is the number of samples that the true is predict as positive. True Negative(TN) is the number of samples that the true is predict as negative. False Positive(FP) is the number of samples that the false is predict as positive. False Negative(FN) is the number of samples that the false is predict as negative. Precision (P) is defined as the following formula (2).

$$R = \frac{TP}{TP + FP} \quad (2)$$

Recall is defined as the following formula (3).

$$R = \frac{TN}{TN + FN} \quad (3)$$

Accuracy is defined as the following formula (4).

$$P = \frac{TP}{TP + FN + TN + FP} \quad (4)$$

The table below is the result of the proposed system for 2000 image data. 1000 images are normal and the remaining images are not normal.

TABLE I. RESULT OF EXPERIMENT

(a) Training and validation result

		Training		Validation		Total
		Success	Fail	Success	Fail	
Screw	Screw attached	1000	0	998	2	1000
	Screw-free	1000	0	997	3	1000
Label	Label attached	1000	0	993	7	1000
	Label -free	1000	0	1000	0	1000

(b) Evaluation for the classification

(%)	Training			Validation		
	P	R	A	P	R	A
Screw	100	100	100	99.8	99.7	99.75
Label	100	100	100	99.3	100	99.65

## V. CONCLUSION

Visual inspection of various types of assembling devices produced in a small assembly plant is carried out by human eyes. Human have difficulty in detecting a defect due to the effect of the afterimage that the brain has memorized what it had just seen. Therefore, many researches have been carried out to perform this automatically. However, these methods require modifications to the program if the appearance of the assembly device changes. Recently, researches are being made on techniques for handling bad judgment through image learning.

In this paper, we propose a method to improve the accuracy of error detection by applying this technique to the external inspection of the assembly device and to use it without changing the same program on various assembly devices. To improve defect detection performance, we introduced CNN. CNN has strong ability in feature extraction, it can compensate for the drawback existing in hand-crafted features. In the future, we try to reduce the loss of feature information in image by improving the structure of CNN.

## REFERENCES

- [1] Escalera, J. M. Armingol, J. M. Pastor, and F. J. Rodríguez. "Visual sign information extraction and identification by deformable models for intelligent vehicles." IEEE Transactions on Intelligent Transportation Systems, 2004: 57-68
- [2] J.R. Uijlings, K.E. Sande, T. Gevers, and A.W. Smeulders, Selective Search for Object Recognition, International Journal of Computer Vision, 104(2): 154-171, 2013.
- [3] Wang Zhiqiang, Liu Jun, A Review of Object Detection Based on Convolution Neural Network, Proceedings of the 36th Chinese Control Conference, 11104-11109, July 26-28, 2017
- [4] X. Zeng, W. Ouyang, B. Yang, J. Yan, and X. Wang. Gated Bi-directional CNN for Object Detection: Springer International Publishing, 2016.
- [5] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86(11): 2278-2324, 1998.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition, IEEE Transactions on Pattern Analysis & Machine Intelligence, 37(9): 1904-16, 2015.
- [7] K.Q. Huang, W.Q. Ren and T.N. Tan, A review on image object classification and detection, Chinese Journal of Computers, 2014.
- [8] K. Simonyan and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, Computer Science, 2014.