

Real-Time Inspection of Multi-sided Surface Defects Based on PANet Model

Yohan Han and Jongpil Jeong(✉)

Department of Smart Factory Convergence, Sungkyunkwan University, Suwon, Gyeonggi-do 16419, Republic of Korea
{coco0416, jpjeong}@skku.edu

Abstract. Quality of products is the most important factor in manufacturing. Machine vision is a technique that mainly performs human cognitive judgment in the industrial field or performs a task that is generally difficult for a human. However the detection of traditional methods of scanning with human eyes has many difficulties due to repetitive tasks. Recently, an artificial intelligence machine vision has been studied to improve these problems. Using the vision inspection system, it is possible to collect information such as the number of products, defect detection, and types without human intervention, which maximizes the operation-al efficiency of a company such as productivity improvement, quality improvement, and cost reduction. Most of the vision inspection systems currently in use are single-sided images, which collect and inspect one image of the product. However, in the actual manufacturing industry, products that are valid for single-sided image inspection are limited to some product groups, and most require multi-sided image inspection. In addition, the inspection system used in the field must meet the production speed required by the actual manufacturing site and inspect the defects of the product. In this paper, we propose a deep neural network-based vision inspection system that satisfies the multi-sided image inspection and fast production speed of products. By implementing seven cameras and optical technology, multi-sided images of the product are collected simultaneously, and a defect in the product can be quickly detected in real time using a PANet (Path Aggregation Network) model. Through the proposed system, it is possible to inspect product defects at the level required at the manufacturing site, and the information obtained in the inspection process will be used as a very important data to evaluate and improve product quality.

Keywords: AI · Deep learning · Machine vision · Defect inspection · PANet

1 Introduction

With the advent of Industry 4.0, the development of IT technology has brought new changes to the manufacturing industry. The smart factory is an intelligent factory that maximizes the operational efficiency of companies such as productivity improvement, quality improvement, and cost reduction by applying Information and Communication Technology (ICT) to all production processes from product planning to sales. The smart factory is pursuing multi-volume, small-scale production method that meets

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12250, pp. 623–633, 2020.
https://doi.org/10.1007/978-3-030-58802-1_45

various requirements of customers and rapidly produces, away from the existing mass production method. It is meaningful to improve production efficiency for products by real-time identification.

In particular, quality is a very important factor in manufacturing. Product quality has a significant impact on the company's growth and securing market competitiveness, and is considered as the first priority in implementing a smart factory. As the manufacturing environment changes, the production system to satisfy this needs to change. Accordingly, many studies are being conducted to improve the quality, which is representative of product defect inspection. Traditionally, product defects were visually inspected by the operator. However, it is difficult to detect defects due to reduced productivity and poor accuracy due to repetitive work. Recently, artificial intelligence (AI) testing systems have been studied to improve these problems. The artificial intelligence inspection system refers to the detection of product defects by collecting, processing, and analyzing product images by replacing the role of the operator with a vision camera and artificial intelligence. By using this method, it is possible to improve productivity and increase detection accuracy and reduce the operating cost of the company.

Most machine vision currently in use collects and inspects a single image of a product by single-sided image. However, in the real manufacturing industry, products that are valid for single-sided inspection are limited to some product groups such as printed circuit board (PCB), semiconductor, display and most of them want multi-sided inspection. Multi-sided inspection requires a high level of hardware and software technology and in particular, the product must be inspected while meeting the production speed required by the actual production site.

In this paper, we propose a system that satisfies the product's multi-sided inspection and fast production speed. By implementing 7 cameras and optical technology, multi-sided images of the product are collected at the same time. In addition, defects of products can be accurately inspected in real-time by using a PANet (Path Aggregation Network). Through the data training model, it is possible to inspect multi-sided defects of products with minimum time and improve the quality of multi-sided products. We classified 3,024 defect databases into 4 defect types and trained a total of 16,000 times. As a result, 98.12% mAP, 98.00% Recall and 98.06% F1 Score were achieved. Through this study, information such as product type, defect location and size can be checked and this can be used as a very important data for evaluating and improving the quality of production products.

The composition of this paper is as follows. Section 2 describes related work. Section 3 introduces the components of the proposed inspection system and Sect. 3.2 describes the PANet model used for data analysis. Section 4 conducts experiments on the proposed system. Finally, Sect. 5 presents conclusion and presents future research directions.

2 Related Work

A machine vision system is a system that inspects products produced at a manufacturing site through a camera [1]. Machine vision enables the number of products, product specs and perfect inspection without human intervention, which maximizes the operational efficiency of the enterprise, including productivity improvement, quality improvement and cost reduction [2]. Traditional machine vision uses a rule-based machine learning algorithm. These algorithms require manual design of product feature extraction and take a lot of time [3]. Due to the development of technology, artificial intelligence has been introduced, attracting attention in many fields. Artificial intelligence is limited because it mimics the structure of a human brain's neural network, but thinking and judgment are possible.

In general, artificial neural networks are implemented as Deep Neural Network (DNN) [4]. Convolution Neural Network (CNN) is the most powerful and effective DNN for recognizing image objects [5]. CNN's deep multi-layers architecture can extract more powerful features than manually built functions and all features are automatically extracted from training data using a back-propagation algorithm [6]. The convolution network provides an end-to-end solution ranging from low defect images to predictions, alleviating the requirement to manually extract suitable features [7]. In addition, the convolutional detection network allows the object to be detected in a few milliseconds using the exact location and size information of the object [8]. However, CNN has a disadvantage that processing speed is relatively slow.

With the advent of PANet, which recently improved the CNN, it is possible to perform defect inspection quickly and accurately. The PANet uses the existing pyramid network, Feature Pyramid Networks (FPN), as a backbone to add an additional three-stage model to improve accuracy [9]. The study of [10] conducted real-time steel strip defect inspection through the improved YOLO detection network. With the advent of the improved deep learning model, the machine vision system is continuously evolving as the processing speed increases and the accuracy increases [11]. Thanks to advances in technology that meet the speed of production of real products, machine vision is available in almost all production facilities. The study of [12] suggested a solution that can easily detect various defects of products through one or more cameras and machine vision software. The study of [13] proposed a study that inspects the number of holes in the PCB through the machine vision system. According to this study, accuracy of more than 95% can be guaranteed, and the study of [14] inspected the odd number defects of the PCB through various image processing technologies such as Otsu Thresholding, Canny Edge Detection, Morphological Process and Hough Circle Transform. [15] studied defects in conveyor belt. In addition, [16] studied cracks and defects occurring in railways.

In this paper, we propose a system that satisfies the multi-faceted inspection and fast production speed of products. The next section describes the Inspection System for PANet Model.

3 Inspection System for PANet Model

3.1 Components of the Inspection System

The components and considerations for designing and implementing the AI vision inspection system are as follows [17].

- Camera: Resolution, Frames Per Second (FPS), Interface, Type, Shutter, Mount
- Lens: Image Size, Focal Length, Field of View (FoV), Working Distance (WD)
- Illumination: Light source types, Intensity, Number of light sources and Angles
- Software: Image acquisition, Image pre-processing, Image processing, Analytics

The inspection system consists of product recognition device, image acquisition device, image processing device and product classification device. Product recognition devices mainly consist of sensors, triggers, etc., and image acquisition devices consist of cameras, lenses, lighting and industrial PC (IPC). The collected images are processed and analyzed through the PANet model at the image processing unit stage, and finally, the product classification unit stage classifies the products according to the results.

Figure 1 shows the structural diagram of an image acquisition device that includes 7 cameras, and each camera has an angle of 60°. When the product moves on the conveyor belt, the optical sensor recognizes the product and the camera and illumination are activated through hardware trigger and ethernet to capture the product image. The collected images are processed through the deep learning model and classify products according to the analysis results.

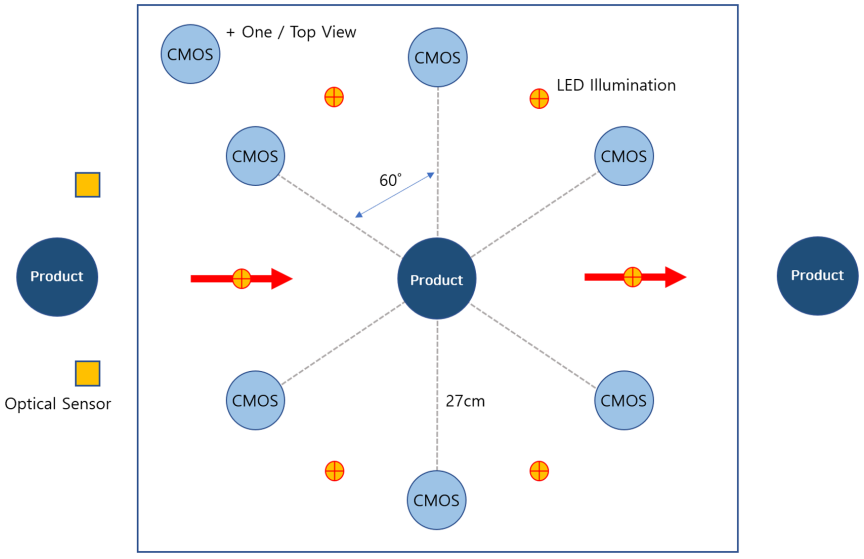

Fig. 1. Structure of image acquisition device.

(a) **Product Recognition Device**

The product recognition device is used to detect a product produced through a conveyor belt and trigger a camera module so that the camera can be photographed at a predetermined time and location. The proposed system uses a radar sensor to detect the product because it is suitable for accuracy and quick response to product detection. The signal detected by the radar sensor triggers the camera and lights for image acquisition. Since the product produced on the conveyor belt is a moving subject, images are collected in a moving state rather than in a stationary state. Therefore, it is a trigger signal that calculates the production speed of a product and gives a signal so that the camera can be photographed at a predetermined time and location.

The trigger is mainly used when a user shoots an image in an aperiodic manner such as a constant frame rate. The trigger mode can operate in two ways: hardware trigger and software trigger. A hardware trigger is an electrical signal input and acts as a trigger, and a software trigger acts as a trigger by instruction. Hardware triggers are ideal for ultra-precise applications that are not suitable for software triggers inherent latency, so it is common to apply hardware triggers to vision inspection systems.

(b) **Image Acquisition Device**

The image acquisition device is a process of collecting an image of an inspection product through an optical system, and is mainly composed of a camera, a lens, lighting, and IPC. A total of seven complementary metal-oxide semiconductor (CMOS) cameras are used, which consist of one Top View and six Around Views. Based on the depth of focus (DoF) of the product, the number of cameras that do not generate image blind spots across the product is 6 and the angle of each camera is 60°. Also, when calculating the distance from the product, a focal length of 27 cm was set.

For accurate and fast inspection, an image-taking camera is very important. Depending on the type of product and production speed, appropriate cameras and lenses are required. Therefore, based on the collected information, the necessary components for selecting cameras and lights are obtained. The components required for camera and lighting selection are as follows.

- Camera: Sensors, Effective Pixels, Frame Rate, Color, Shutter, Interface
- Lens: Imager Size, Focal Length, Aperture Range, Mount

The most important thing in a camera is an image sensor. The image sensor works in conjunction with the lens to receive light data. The lens also accepts light data through a focus ring that focuses and an iris that controls the amount of light. Therefore, it is essential to consider Focal Length (FL), Frames Per Second (FPS), Field of View (FoV), and Working Distance (WD) to select cameras and lenses.

Field of View (FoV) refers to the horizontal *(H)* or vertical *(V)* area that can be obtained with the configured optical system. The formula is as follows.

- $FoV\,(H)\,or\,(V) = \frac{Sensor\,Size\,(H)\,or\,(V)}{Magnification}$
- $Magnification = \frac{Sensor\,Size\,(H)\,or\,(V)}{FoV\,(H)\,or\,(V)}$

The light source is also very important to collect quality images from the product. This is because, in the process of illuminating the product through lighting, if the product has a shadow, the shaded part cannot be inspected for defects. Therefore, when the product to be inspected is selected, the distance and angle between the camera and the light source should be set through experiments. Whether it is to reflect light directly to an object, indirectly, or whether to use coaxial, transmissive, or backlit lighting at a certain angle, and at what angle, and which color of lighting is used, it greatly affects the inspection results. Therefore, when selecting a light source, it is most important to find the optimal condition after sufficiently experimenting with various conditions.

(c) **Image Processing Device**

Through the image processing and analysis process collected through the image acquisition device, it is determined as a normal product or a defective product. Typically, a camera manufacturer provides a library of cameras. This is called software development kit (SDK), and the most commonly used SDK compilers and languages are Visual Studio and C/C#.

(d) **Product classification Device**

The product classification device classifies the product as a normal product or a defective product based on the results obtained through the image processing device. The image detected through inspection sends a signal to the actuator according to the result, and the actuator is located in the result part of the inspection system and operates as a motor at the center point of the conveyor belt. The result value received through inspection is transferred to the actuator, so that the normal product flows as it is, and the defective product is classified as a separate left by the actuator to reach the defective parts.

3.2 Inspection System for PANet Model

Recently, due to the development of artificial intelligence (AI) technology, image processing technology through machine learning has been studied a lot. PANet (Path Aggregation Network) is a model optimized for real-time image recognition. PANet is an algorithm that uses the existing pyramid network, FPN as a backbone and adds an additional three-level model. In general, the more low-level features that are first extracted, the greater the distance to the output, the result is not reflected sufficiently in the result due to the long distance. To compensate for this, a method in which the extracted features are reflected in the result is introduced. Figure 2 shows the structure of a typical PANet [18].

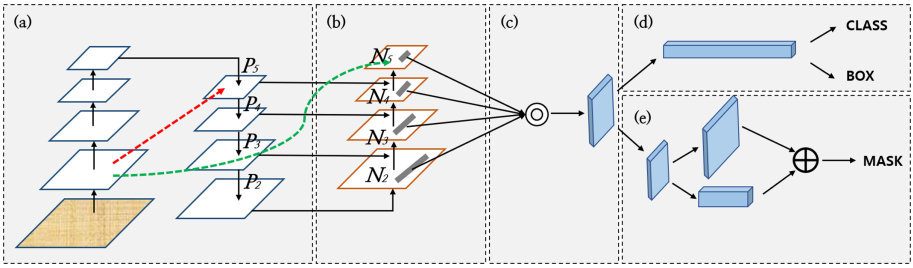

Fig. 2. Architecture of PANet detection network [18]

(a) **FPN Backbone**

Feature Pyramid Networks (FPN) Backbone is an existing pyramid type network. Because the low-level layer does not contribute much to the final prediction, the prediction result is incorrect, so each black layer of this pyramid is like a layer in the structure diagram, but it is actually composed of ResNet-50. To reach the top level, it passes a very large number of layers, and as a result, the Propagate does not work smoothly.

(b) **Bottom-up Path Augmentation**

PANet constructs a new *Nn* pyramid of simple convolution layers. The new pyramid only takes one shortcut and a few convolution layers, like the green dotted line, in order for the low-level layer to reach the high-level layer. This means that smooth information flow is possible compared to the existing Backbone Model. Although all *Nn* layers have different resolutions, they have the same 256 filters and are implemented by *Pn* and Concatenate the result of down sampling *Nn-1* with 3×3 Convolution operation of Stride 2.

(c) **Adaptive Feature Pooling**

The region of interest is extracted for each *Nn* through the ROI Align technique. Since the feature map has a difference in resolution of the final output, if you set the ROI with Naive, there may be a problem of inconsistent with the exact pixel corresponding to the output. The technique to correct this is ROI Align. Then, ROI of all *Nn* layers is pooled element-wise.

(d) **Box Branch**

If you want to simply classify the instances into boxes, you can connect the multi-sided pooling results to the Fully-connected Layer as shown in the above structure diagram, and apply the following method to perform segmentation.

(e) **Fully-Connected Fusion**

PANet separates pooling results into a convolution layer and a fully-connected layer for accurate prediction and then integrates them. Since PANet has spatial information for the convolution layer and no spatial information for the fully-connected layer it has a rich amount of information. The specific implementation method is shown in Fig. 3.

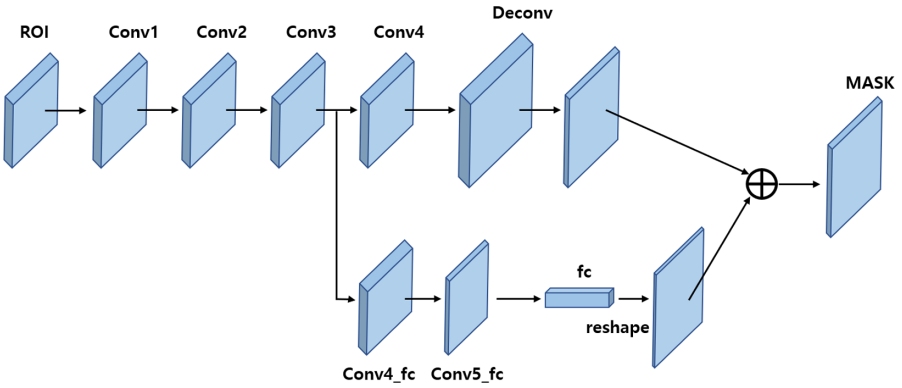

Fig. 3. Mask prediction branch with fully-connected fusion [18]

4 Experiments

In this paper, we conducted an experiment on a cosmetic case produced in a real factory to examine defects in the product. Cosmetic cases of PE, PP and ABS material produce various effects from injection, painting, and assembly processes such as scratches, scars, pollution, unassembled. The defect image is shown in Fig. 4. These defects not only affect the appearance of the product, but also the durability of the product. Therefore, it is very important to improve production quality by inspecting defects in products.



Fig. 4. Several types of surface defects on cosmetic case

4.1 Defect Database

We collect the defect database of products in the assembly process, which is the final stage of product production. In this paper, a total of four types of defect images were collected. Each image is cropped to 608 * 608 before being transmitted over Ethernet. The defect data base collected for the experiment has 3,024 images of 4 types. Details of the data set are given in Table 1. Bulleted lists may be included and should look like this:

Table 1. Dataset of cosmetic case surface defect images

	Scratch	Scar	Pollution	Unassembled	Total 2
Training	779	642	481	322	2224
Test	200	200	200	200	800
Total 1	979	842	681	522	3024

4.2 Training

The proposed system consists of seven complementary metal oxide semiconductor (CMOS) cameras, shadow-less diffused ring lights, optical sensor, hardware trigger and electric precipitator to acquire multi-sided image of the product. In addition, PANet algorithms were training from Window10 pro equipped with i7-9700F, DDR4-32 GB and GTX 2080 Ti, and four types of defect database were training a total of 16,000 times. During the training process, batch 64, resolution 608 * 608, learning rate 0.00261 and activation function Leaky ReLU. The time taken to detect defects in 800 test images was 8 s, mAP 98.12%, recall 98.00% and F1 score 98.06%.

4.3 Real-Time Analysis

The time to inference the multi-sided image of the product through the proposed system is 0.224 s on average and the time to inspect per image is only 0.032 s on average. The production speed at the actual production site is 0.09 m/s and the distance of the camera is 22–35 cm. In order to meet the production speed required at the production site, the speed of 32–45 FPS must be guaranteed. As a result of analysis using the PANet algorithm proposed in this paper, the inspection speed required at the production site was fully satisfied by achieving a speed of 35 FPS. Finally, normal and defective products were classified through separate classifiers. The detection details are shown in Table 2, and the detection results are shown in Fig. 5.

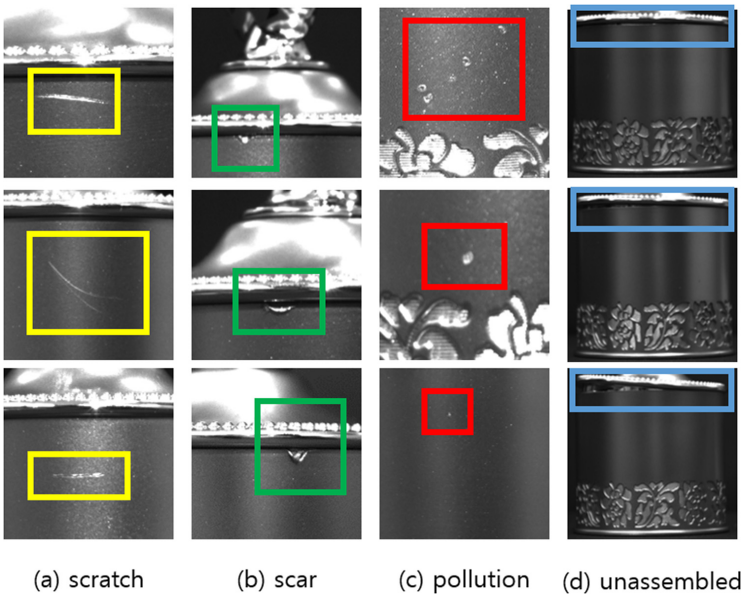

Fig. 5. Detection results of 4 types of surface defects on cosmetic cases

Table 2. Detection results (mAP & recall)

	Scratch	Scar	Pollution	Unassembled	Average
mAP	98.47%	97.03%	97.01%	100.00%	98.13%
Recall	97.00%	98.00%	97.51%	99.55%	98.02%

5 Conclusion

We have established a system that meets multi-sided inspection of products and fast production speed. Defects were inspected for four types that occurred during the production of cosmetic cases and the results show that the proposed system achieves 98.12% mAP, 98.00% Recall and 98.06% F1 score. Using a PANet defects in products can be inspected in real time and information such as product type, defect location and size can be checked, which can be used as a very important data to evaluate and improve the quality of production products. Through the proposed system, defect inspection is possible in the environment required by the actual manufacturing site, the more defect types and sample data are obtained, the better the results will be.

References

1. Nashat, S.: Machine vision for crack inspection of biscuits featuring pyramid detection scheme. J. Food Eng. **120**, 233–247 (2014)
2. Khude, P.: Object detection, tracking and counting using enhanced BMA on static background videos. In: 2013 IEEE International Conference on Computational Intelligence and Computing Research (2013)
3. LeCun, Y.: Deep learning. Nature **521**, 436–444 (2015)
4. Demuth, H.: Neural Network Design. Martin Hagan (2014)
5. Krizhevsky, A.: ImageNet classification with deep convolutional neural networks. In: 2012 Advances in Neural Information Processing Systems (2012)
6. Bengio, Y.: Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1798–1828 (2013)
7. Sermanet, P.: Overfeat: integrated recognition, localization and detection using convolutional networks. 2013 arXiv preprint arXiv:1312.6229 (2013)
8. Redmon, J.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016)
9. Liu, S.: Path aggregation network for instance segmentation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, no. 116, pp. 8759–8768 (2018)
10. Li, J.: Real-time detection of steel strip surface defects based on improved YOLO detection network. IFAC-Papers OnLine **51**(21), 76–81 (2018)
11. Baygin, M.: A new image stitching approach for resolution enhancement in camera arrays. In: 2015 9th International Conference on Electrical and Electronics Engineering (ELECO), pp. 1186–1190 (2015)
12. Yaman, O.: A new approach based on image processing for detection of wear of guide-rail surface in elevator systems. Int. J. Appl. Math. Electron. Comput. **4**, 296–300 (2016)
13. Wang, W.: A machine vision based automatic optical inspection system for measuring drilling quality of printed circuit boards. IEEE Access **5**, 10817–10833 (2017)
14. Baygin, M.: Machine vision based defect detection approach using image processing. In: 2017 International Artificial Intelligence and Data Processing Symposium (2017)
15. Yang, Y.: On-line conveyor belts inspection based on machine vision. Optik-Int. J. Light Electron Opt. **125**(19), 5803–5807 (2014)
16. Yaman, O.: PSO based diagnosis approach for surface and components faults in railways. Int. J. Comput. Sci. Softw. Eng. (IJCSSE) **5**(5), 89–96 (2016)
17. Bahaghighat, M.: Designing quality control system based on vision inspection in pharmaceutical product lines. In: 2018 International Conference on Computing, Mathematics and Engineering Technologies (2018)
18. Lie, S.: Path aggregation network for instance segmentation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, no. 116, pp. 8759–8768 (2018)