Evolution of Image-Processing Technologies Centered on X-ray Inspection

Hirotaka Nakajima, Ryohei Ishikawa, Kengo Oda, Takeshi Yamazaki

[Summary] This article introduces the latest trends in image-processing technologies centered on X-ray inspection machines and explains some examples of Anritsu's image processing. It describes image analysis methods focusing mainly on the latest trends in computer vision as well as efficient use of computing resources with practical examples of Anritsu's deep learning and rules-based processing. The deep-learning examples cover foodstuffs with large in-dividual differences and difficulties in extracting quantifiable characteristics. Examples of sealing check using rules-based processing to specify the package detection region as well as dual-energy methods to reduce the product effect on contaminant detection caused by product irregularities and piles are described. Last, future problems and prospects for X-ray inspectors are discussed.

1 Introduction

Detecting contaminants in food is an important theme in food safety and reliability underpinning consumer and social trust.

The market has high hopes for X-ray inspectors because they cannot only detect contaminants but can also detect malformed product shape and missing quantity. Anritsu has developed a number of signal processing and image-analysis algorithms based on its own research to accurately capture contaminant data only.

With the recent appearance of deep learning, image-processing technology is making remarkable progress. Consequently, similar progress is required in X-ray image-processing technology. Anritsu has been playing catch-up with the latest technologies while also improving conventional techniques and is now offering new processes for detecting contaminants and abnormalities that were previously undetectable to help improve the efficiency of food production.

This article introduces the latest trends in computer vision technologies and our company's image-processing technology.

2 Trends in Image-Processing Technology

2.1 General Computer Vision

Conventional image-processing technologies are mostly rule-based in which an engineer designs a valid quantification feature and image analysis is performed using statistical characterization processing. Recent deep-learning based computer vision has produced excellent results.

The leading technology in deep learning for image processing is called convolutional neural networks (CNN). CNN creates a powerful model based on unique small quantifiable product differences, or, in other words, automatically classifies an image as a combination of smaller parts. It is mainly used for tasks such as object detection and segmentation, which take an image as input and estimate the type and area of the object in the image, and some models achieve better-than-human recognition rates.

On the other hand, the last few years have seen the rise of Transformer[1] technology that does not use CNN. Transformer technology is a branch from the machine translation field which predicts separate affiliated data from other data. The intermediate expression obtained by the prediction process is said to "understand" the affiliated data expressed in the data, and the technology with high general applicability has become popular not only for natural language processing (NLP) but also in the image and audio processing fields. In particular, Vision Transformer[2] has shocked engineers in the imaging field by producing far better results than CNN.

Current research is focused on reducing computing costs as well. Although on-device computation is required in situation that inference processing cannot be assigned to the cloud network, securing computation resources is still an issue depending on the device size limitations. The main approaches to minimizing the amount of data while assuring the functionality of machine-learning models are pruning, quantization, and distillation. The pruning method removes parameters (branches) of low importance. Quantization methods reduce memory and computational cost of using neural network with lower bit floating-point format.

Distillation method is the process of transferring knowledge from large model to small model. By training with original data and large model output, small model learns concise knowledge representations.

The recent development of deep-learning technologies is remarkable and the results of the devised techniques are appearing in fields other than image processing. We'll continue to research wide field.

2.2 X-ray Image Processing

Although the size of X-ray monochrome image data is much smaller than that of color images, deep learning is already an indispensable technology in enhancing performance as well as in computer vision overall. The number of instances where deep learning solves challenging task is increasing. Known examples of this is detection of nonstandard product that is easily distinguished from normal products by human, but failed by traditional image processing.

On the other hand, rules-based processing is still required as well. Deep learning requires large amounts of image data for each inspection item which in turn requires high-speed hardware, resulting in much higher costs. Introduction to production lines requires considering hardware performance and inference time and currently uses algorithms according to the situation.

3 Anritsu Approach (Deep Learning Technology)

3.1 Outline–Individual Differences in Large Food Shapes

Production and shape inspection of sausages sees many customer complaints caused by smaller than normal size and bent shapes, requiring automated inspection to solve. The degree of difficulty in using X-ray image processing to inspect sausages was unknown for a long time until our company realized this by using object recognition using deep learning technology.

Two factors cause difficulties at X-ray image inspection. One thing is big individual differences in size and shape. The other thing is sausages overlapping each other on image. This an unavoidable problem with bagged products due to the transmission property of X-rays. Requires processing to separate each item but if the product pile is dense, the increased pattern combination is difficult to process. In the case of sausage shape, use of natural sausage casings makes it difficult to assure consistent thickness and length. As a result, conventional techniques focusing on details such as mass and shape are unable to distinguish normal and abnormal products with high accuracy.

Deep learning technology solves the above problems because it can focus on fine details and uses huge numbers of pattern combinations to make a judgement in a realistic time.

3.2 Technology Introduction

In this example, we used CNN-based technology in consideration of required hardware resources and inference speed. CNN creates a powerful model based on unique small quantifiable product differences, or, in other words, automatically classifies an image as a combination of smaller parts. The CNN technology is explained below.

The relatively early R-CNN[3] model is divided into two stages: extraction of region proposals and classify regions. The object candidate domain is extracted by selective searching and features of images in the domain are converted to vectors by CNN and classified by SVM. There may be several thousand numerated candidate domains per image with the disadvantage of SVN's requiring a very long time to train and make inferences. Moreover, since both deep learning and other machine learning models are combined, each requires repeated training and tuning.

The later Fast R-CNN[4] and Faster R-CNN[5] models are improvements on the earlier version with faster speeds and precision expressing the whole in one deep learning model. The first-stage enumeration of candidate domains replaces inference about whether or not there is an object in the domain using an anchor box with recurrence of domain size, and the second class-identification and domain-range correction stage uses pooling and full-layer combination, solving both the output precision and speed issues to assure major progress in introduction of deep learning on production lines.

Recent progress has speeded-up the NN model with the appearance of Residual Block[6] technology that creates large-scale models preventing vanishing gradient, as well as a key technology called Feature Pyramid Network[7] featuring robust multi-scaleup for creating multiple-size feature maps using an Encoder-Decoder. Two resulting high-speed and high-performance object recognition technologies are SSD[8] and YOLO[9]. Anritsu has achieved CNN-based high-speed, high-precision inspection by adopting all these techniques.

3.3 Function Evaluation Test

3.3.1 Test Preparation

Mis-shaped sausages were created mostly by ripping filled sausage casing. Figure 1 shows X-ray images of both normal and two typical non-standard sausage shapes. The bent non-standard shape was created by folding during production and the spherical shape was 1/8th normal size with a shape like a rugby ball.

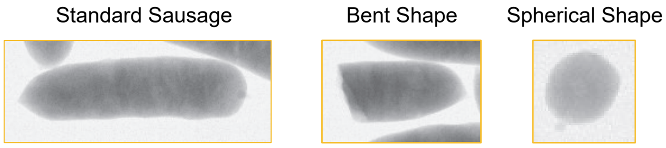

Figure 1 Sausage X-Ray Image Types

Both standard and non-standard shaped sausages were packed in a single bag as a sample for performance evaluation using our X-ray inspector to create a data set for image evaluation. The pile of sausages in the sample bags were the same as at X-ray inspection on normal production lines.

3.3.2 Test Results

5-Fold Cross Validation was used to evaluate performance. Five subsets of each type of non-standard sausage were created from the data sets created as described in section 3.3.1. At this, time, care was taken to prevent arrangement of the same samples across subsets. Five trainings and evaluations were performed while replacing subsets so that $train: eval: test = 3:1:1$ and the mean value was used as the final evaluation result. In addition, images of normal sausages were evaluated in the same manner without using training.

Detection of non-standard products showed performance of better than 99% at the calculated recall (Eq. (1)). In addition, precision (Eq. 2)) for normal and faulty products was also confirmed as meeting the required standard. The inference time per image was less than about 100 ms, which is sufficient in terms of speed for introduction of this technology to production lines.

$$Recall = \frac{True\ Postive}{True\ Positive + False\ Negativee} \qquad Eq. (1)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \qquad Eq. (2)$$

Figure 2 shows the detected image result. Although the sausages were piled in the bag, non-standard spherical sausages (indicated by blue rectangle) could still be detected.

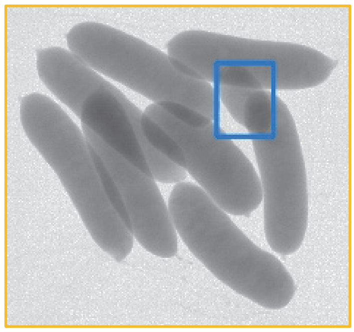

Figure 2 Spherical Sausage Detection Result

4 Anritsu Approach (Traditional Image Processing)

4.1 Sealing check

4.1.1 Outline

Abnormalities such as contents trapped in package seals are commonly called ‘contaminant defects’. Not only do these contaminant products create a bad impression about the package but there is also risk of lowered content quality due to reduced package air and moisture tightness, requiring full inspection of all packages to prevent occurrence.

Seal check process takes a two-phased approach. Phase 1 is extracted seal regions. Phase 2 is detect the presence or absence of a product caught in the seal. Since there are many types of package seals on the market, general-purpose processing is required. This article introduces detection of single-edge seals.

4.1.2 Technology Introduction

The basic procedure for determining the seal area pinpoints the vertical and horizontal side from the package external shape and the inspection target is specified as a large area inside the packet edge. To capture data from thin packages, noise filter processing is performed to distinguish the package and the production-line belt surface and then the edge is highlighted by edge tuning before other processing such as creation of a bounding rectangle. Next, the package part is binarized (labeling) to extract the package.

Although the above-described method can extract seal location from the overall product shape data, sometimes – depending on the product – it may be better to use data on the contents. Furthermore, since it is necessary to infer data about the relative position even when products are carried at an angle on the production line, the angle is calculated from data on the product center of gravity, and long and short axes to perform correction processing.

The methods described so far perform pre-processing such as differential and averaging processes on the extracted location and then finally detect the contaminant product part by binarizing. The above-described functions are used widely to detect contaminant product in the various types of packaging on the market.

4.1.3 Application Examples

・ Side Sealed Package

Figure 3 shows inspection images of side packaging. This method is used to specify the area inside the package edges by detecting specific outside edges (indicated by yellow lines).

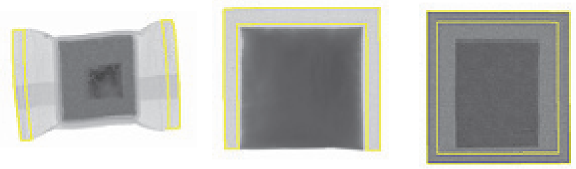

Figure 3 Example of Side Sealed Package Inspection

・ Tray Package

This method is used to specify an area inside the product edges. As shown in Figure 4, because the tray corners are rounded, there is a function for settings different curvatures for each corner to set a matching inspection area.

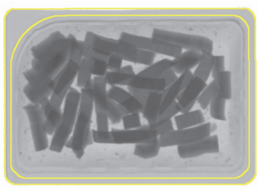

Figure 4 Example of Tray Package Inspection

・ Serial Package

Products linked by a series of packages are called serial packaged products. As shown in Figure 5, as well as specifying the region inside the product edges, a central inspection area is created that is separate from the periphery.

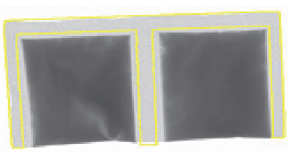

Figure 5 Example of Serial Package Inspection

・ Pillow Bag Package

A bagged product means a bag that is created by making a film loop and sticking a bag on the back side. If the contaminant product area overlaps the attached back part, the region with different density can prevent detection.

As shown in Figure 6, this method specifies the region inside the product edge. We add a function to set an inspection area for back-lining parts that can be set for a separate detection limit.

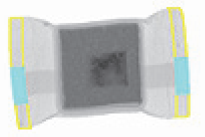

Figure 6 Example of Bagged Product Inspection

・ Vacuum Sealed Package

Deep-drawing is a sealing method that creates a concavity matching the package contents. It is commonly used for blocks of sliced ham and bacon but since the content position is not constant there may be problems with variations in the region to be inspected.

In this case, this method is used to specify the outside edges of the contents. Only the area of the contents of the entire product are binarized and used as a mask to inspect the seal region other than the content region. Figure 7 shows an inspection image. The green part indicates the contents and the gray part is the inspected area.

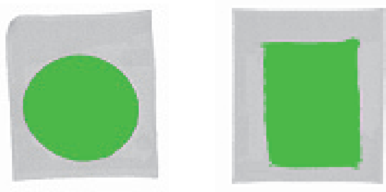

Figure 7 Example of Vacuum Sealed Package Inspection

4.2 Dual-Energy Processing Technology

4.2.1 Outline

A standard X-ray inspector can sometimes receive a signal of the same level from the inspected product itself as from the contaminant (due to product shape irregularities or piles). To prevent misdetection of the contaminant, the contaminant detection sensitivity can be adjusted but detection of the target contaminant tends to become more difficult. Consequently, dual-energy X-ray inspectors have been developed as a technique for reducing the inspected product effect.

4.2.2 Technology Introduction

The dual-energy method captures images using a dual-energy sensor that is sensitive to two types of energy radiated from a single X-ray source. The product and contaminant in the captured image have different transmission properties and this difference reduces the product effect. As a result, although there is some dependence on the size and type of contaminant, the contaminant signal is emphasized, which widens the range of detectable contaminants.

For example, although a contaminant in an irregular product might be seen in the X-ray transmission image, it can sometimes be difficult to evaluate and confirm using the signal values. In this case, using the dual-energy method can emphasize the signal.

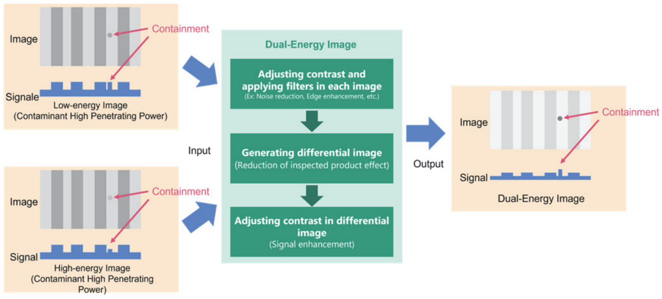

Figure 8 Outline of Sensitivity Improvement using Dual-Energy Method

Figure 8 outlines dual-energy image creation. A contaminant low-transmission-energy image (low-energy image) and a contaminant high-transmission-energy image (high-energy image) are used as the input. First, to reduce the product effect, contrast is adjusted (gamma correction) so that changes in the pixel level caused by product irregularities and materials are about the same. Next, pre-processing, such as noise filtering and signal emphasis, are performed for each image so as to maintain the contaminant pixel level for low-energy images as much as possible as well as to lower it for high-energy images. After obtaining the differences between each image in this way, the contrast is adjusted to emphasize the contaminant signal and output the dual-energy image.

If contaminant detection processing is applied to this dual-energy image and is then compared with a standard type, the contaminant can be detected with high sensitivity. In this procedure, a dual-energy image is designed by selecting the best preprocessing and contrast to create multiple images matching the target product and contaminant. This is combined with contaminant detection processing to configure a dual-energy contaminant-detection algorithm matching the application. For example, this method is very useful in commercial food processing where mostly metals and plastics are detected and the products have large differences and are in piles, as well as in meat production (chicken cuts and packs, minced meat, meat joints, etc.) where the main contaminant is bone splinters.

4.2.3 Contaminant Targets

Figure 9 shows the different contaminant targets for each type of X-ray inspector. While the standard X-ray inspector can detect non-metallic contaminants, such as stone and glass as well as metals with high sensitivity, the dual-energy method increases the detected range of contaminant materials to include low-density contaminants, such as plastics and bone in meat.

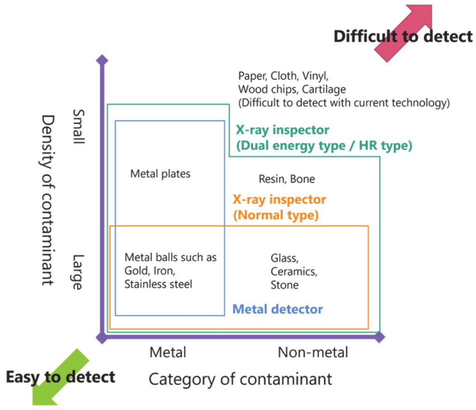

Figure 9 Target Contaminants for Each Inspection Type

4.2.4 Application Examples

This section introduces some easy-to-see examples of dual-energy inspection results.

The first example uses the inspection results for packaged macaroni. Figure 10 is an actual inspection image; the inspected macaroni package contains aluminum discs of six different radii and four thicknesses arranged in order of smaller to larger radius. While the standard X-ray inspector had difficulty detecting discs with a thickness of 0.5 mm or less, the dual-energy X-ray inspector could clearly detect discs with a thickness of 0.3 to 0.5 mm, irrespective of the radius. Table 1 compares the detection results for a 0.3-mm thick disc.

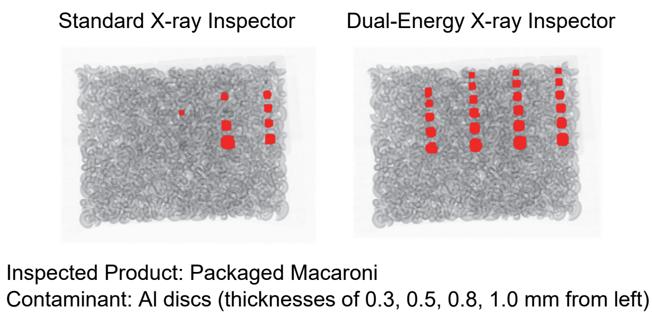

Figure 10 Dual Energy Application Examples (Packaged Macaroni)

Table 1 Comparison of 0.3-mm Thick Al Disc Detection Results

Radius	ϕ2.0	ϕ3.0	ϕ4.0	ϕ5.0	ϕ6.0	ϕ8.0
Standard	×	×	×	×	×	×
Dual	×	○	○	○	○	○

The next example shows detection of bone in chicken breast meat. Figure 11 is an actual inspection image of chicken breast with small thin slivers of bone attached. While the standard X-ray inspector had low detection sensitivity, the dual-energy inspector clearly has better sensitivity.

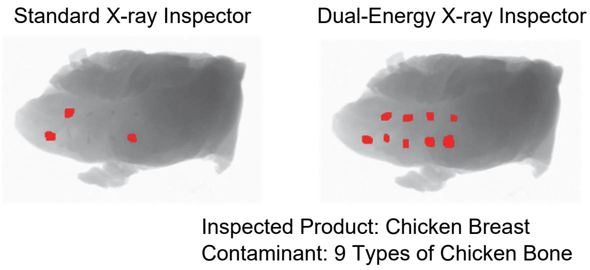

Figure 11 Dual Energy Application Examples (Chicken Breast)

5 Outlook

Deep learning is expected to become the mainstream approach to detection over the next few years. Rules-based detection has difficulties in solving current hard detection problems and use of deep-learning based detection is anticipated when large-volume pattern recognition becomes practical. There are two current issues to solve in two steps at introduction of deep learning.

First is the need and time required to train large numbers of images for each product. For example, consideration must be given to building a general-purpose system for efficiently collecting customer production-line data.

Second is the high cost of hardware to achieve step 1. Cutting data sizes using quantization as described in section 2 is required to promote efficient computation resources. Labor-saving computation will shorten inference times and is expected to increase the number of applicable lines.

Conventional rules-based systems have strengths to these issues above, while deep-learning can support objects which rules-based is unable. Since each has strengths and weaknesses, future development of X-ray inspection systems will require using both technologies as necessary.

On the other hand, both algorithms and imaging methods will be important elements in image-processing technology. Image processing is a technology for detecting contaminants that do not appear in images. For example, X-ray detectors are not useful for detecting materials that are difficult to image using X-rays. Since contaminants such as plastics, rubber, etc., with high X-ray transmissivity, and large parts such as thin-film packaging do not appear in X-ray images, contaminant detection and poor-packaging faults remain difficult even with introduction of deep learning. There are multi-modal methods for solving these problems by adding sensors to capture other data in addition to X-ray images. Because images captured by each technology have both good and bad points, in the future, it will be necessary to solve these issues by combining these characteristics according to the problem.

References

1) A. Vaswani, et al. “Attention is all you need.” In NIPS, 2017.
2) A. Dosovitskiy, et al. “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.” CoRR, Vol.abs/2010.11929 (2020) (online), available from <https://arxiv.org/abs/2010.11929>.
3) R. Girshick, et al. “Rich feature hierarchies for accurate object detection and semantic segmentation.” In CVPR, 2014.
4) R. Girshick. “Fast r-cnn.” In ICCV, 2015.
5) S. Ren, et al. “Faster r-cnn: Towards real-time object detection with region proposal networks.” In NIPS, 2015.
6) K. He, et al. “Deep residual learning for image recognition.” CVPR 2016, pp. 770-778, 2016.
7) T-Y. Lin, et al. “Feature Pyramid Networks for Object Detection,” CVPR 2017, pp. 2117-2125, 2017.
8) W. Liu, et al. “SSD: Single shot multibox detector.” In ECCV, 2016.
9) J. Redmon, et al. “YOLO9000: Better, Faster, Stronger.” In CVPR 2017, pp. 7263-7271 , 2017.
10) H. Watabiki, T. Takeda, S. Mitani, T. Yamazaki, M. Inoue, H. Koba, I, Miyazaki, N. Saito, S. Wada, T. Kanai: “Development of Dual-Energy X-ray Inspection System”, Anritsu Technical Review No.20 (Mar.2013) pp. 59-66

Authors

Hirotaka Nakajima
Product Development Dept.
Development Div.
Infivis Company

Ryohei Ishikawa
Product Development Dept.
Development Div.
Infivis Company

Kengo Oda
Product Development Dept.
Development Div.
Infivis Company

Takeshi Yamazaki
Product Development Dept.
Development Div.
Infivis Company

Publicly available