Development of Convolutional Neural Network Architecture for Detecting Dangerous Goods for X-ray Aviation Security in Artificial Intelligence

Woong Kim, Chulung Lee

Development of Convolutional Neural Network Architecture for Detecting Dangerous Goods for X-ray Aviation Security in Artificial Intelligence

Woong Kim[1], Chulung Lee[2]

[1] Department of Industrial Management Engineering, Korea University, Korea, Republic of (South Korea)
[2] School of Industrial Management Engineering, Korea University, Korea, Republic of (South Korea)
woong2241@naver.com

Abstract. Aviation-security X-ray equipment is used to screen objects, while human screeners re-examine baggage and travelers to detect prohibited objects. Artificial Intelligence technology is applied to increase the accuracy in searching guns and knives, considered the most dangerous in X-ray images at baggage and aviation security screening. Artificial intelligence aviation security X-ray detects objects, finds them rapidly, reducing screeners' labor, thereby providing better service to passengers. In this regard, neural networks based on machine learning have been continuously updated to develop such advanced equipment. In this study, the neural network O-Net is developed to improve object detection. O-Net is developed based on U-Net. The developed O-Net is tested for various neural networks, providing a wide range of experimental results.

Keywords: Artificial Intelligence, Machine Learning, Aviation security, X-ray detection.

1 Introduction and Related Research

The main reason for improving aviation security was the 9/11 terrorist attack. As the attack carried the largest number of fatalities and shocks worldwide, countries began to increase their aviation security [1]. In order to build an intelligent security system, numerous data collection and computing technology concerning aviation security systems have been developed. Safe air freight transportation by restricting the transport of dangerous baggage using X-rays and preparing countermeasures against aircraft terrorism in various ways is being promoted [2]. Passenger safety has increased due to such reinforcement of aviation security. However, a thorough baggage inspection makes travel overseas difficult and causes losses to the airline industry, such as airport entry/exit costs and flight schedule change costs [3][4]. Particularly, although the security control system has been developed at a professional level, in practice, errors when identifying dangerous substances have been increased due to the increase in work stress and fatigue of aviation security personnel caused by exhaustive immigration controls [5]. In this regard, artificial intelligence has been developed to enable X-rays to reduce the workload of the aviation security personnel as much as possible [6]. In this study, we construct an algorithm for the automatic detection of dangerous goods from X-ray inspection images by applying an artificial neural network.

Aviation security equipment improves aviation services by detecting passengers' carry-on items, checked baggage, oversized baggage, and dangerous or hazardous substances. Artificial intelligence aviation security systems able to detect dangerous objects are constantly evolving, depending on the type of security service. While most equipment detects dangerous goods, X-ray scanning equipment can capture images of the contents inside the luggage. The screener can check for the presence of dangerous goods. Because mistakes occur during the human identification process, screeners have a high probability of error identification. Therefore, image recognition algorithms can help improve the aviation security process and detect dangerous objects from improved X-ray images [7].

X-ray images are expressed by the X-ray transmittance, which is related to density. When no object is present, or the density is low, the image displays white. In contrast, when an object is present, or the density is high, the image displays blue or red with high saturation are [8]. In a situation where objects overlap, the image displays a diagram according to the degree of transmission of X-rays. In addition, all information of the overlapping object is displayed as a visual picture through X-ray. Therefore, visual difficulty, complexity, and overlapping problems are characteristics of X-rays [9] [10] [11]. The "overlapping" phenomenon increases the stress of the screener. Thus, various studies considering U-Net structural changes have been developed. For instance, various studies represented the modification and preprocessing of the input image-processing step [12]. In this study, deep learning is implemented to detect dangerous objects such as guns and knives. Moreover, X-ray images that can accurately identify target objects are obtaining by overcoming the limitations of overlapping phenomena in X-ray images. That is, we designed two U-Nets in an O-Net to study the characteristics of the X-ray image.

2 Development of Convolutional Neural Network Architecture

2.1 Neural Network Structure Development: O-Net Structure

O-Net networks are usually composed of fully convolutional networks (FCN) based on the U-Net of semantic divisions. The structure of the encoder-decoder for image segmentation of the O-Net structure is shown in Fig. 1. The numerous layers are multi-channel feature maps. As the neural network layers deepen, the number of parameters can be significantly reduced, considering the characteristics of the layers. The most important step in the network is to copy and crop the 3 x 3 convolution kernel computed from the multi-channel functional map of the encoder part and connect it from the top to avoid loss of boundary pixels in each convolution process. The reason is that reducing and stretching the input image through the neural network prevents the loss of pixel information. Encoder and decoder reasons can specify the exact location of spatial information. The first image of the input value was a color image, and the second image was a grayscale image. The two images were each trained on a neural network, and the output image was a segmentation map representing the predicted class of each pixel.

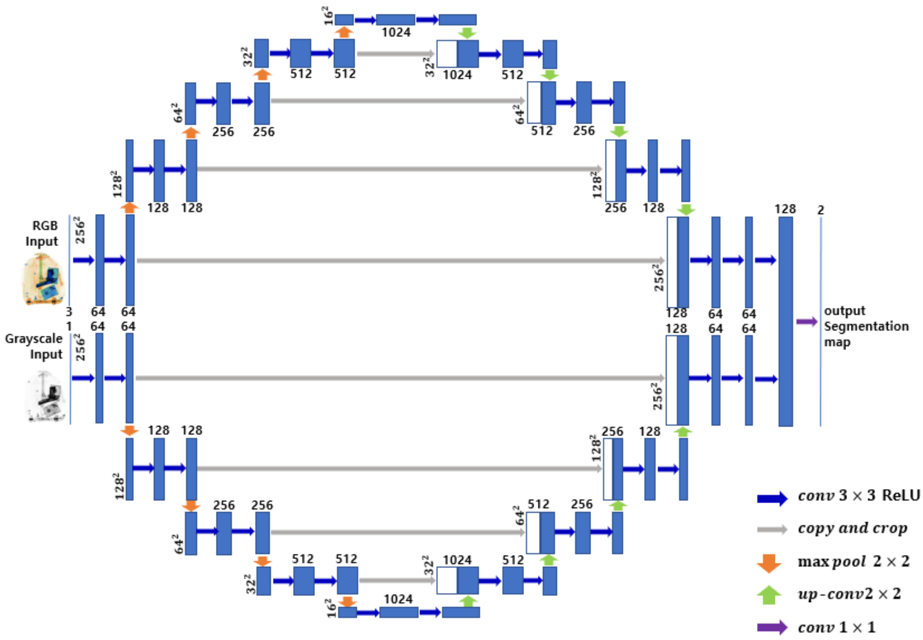


Fig. 1. The first image of the input value was a color image, and the second image was a grayscale image in O-Net structure.

2.2 Performance Measure

When evaluating a model, the *Confusion Matrix* is used to evaluate the precision of the model. How practical and accurate the model classified the image? The confusion matrix is shown in Table 1. Four values define the confusion matrix: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). TP predicts a correct answer as true, FP predicts a false answer as true, FN predicts a correct answer as false, and TN predicts a false answer as false [13].

Table 1. Confusion Matrix

		Actual	
		Positive	Negative
Predicted	Positive	TP	FP
	Negative	FN	TN

In particular, because the proposed model quantifies the value for each pixel of the semantic segmentation, *Pixel Accuracy* and *m-IoU* evaluation scales are necessary. *Pixel Accuracy* refers to the number of pixels predicting successfully among all pixel classes as follows (1). The model's evaluation index evaluates the pixel-wise predicted values of *Intersection-over-Union (IoU)* as follows (2). *IoU metric*, also known as *Jaccard index*, is basically a method to quantify the percent overlap between the target and the prediction. Therefore, IoU_i are denoted by $TP + FP + FN = Ground\ truth\ \cup Prediction$ and $TP = Ground\ truth\ \cap Prediction$. *m-IoU* represents IoU as the arithmetic mean of several test images, as shown in Eq. (3). *Precision* and *Recall* are pattern recognition and information retrieval fields using binary classification. *Precision* is the proportion of results classified as relevant among the search results, as shown in Eq. (4). *Recall* is the percentage of items actually searched among items classified as relevant, as follows (5). Both *Precision* and *Recall* rely on measures of relevance.

$$Pixel\ Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{TP+TN}{Total} \tag{1}$$

$$IoU_i = \frac{TP_i}{TP_i+FP_i+FN_i} = \frac{Ground\ truth\ \cap\ Prediction}{Ground\ truth\ \cup\ Prediction} \tag{2}$$

$$mIoU = \frac{1}{n}\sum_{i=1}^{n} IoU_i \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

3 Experiments

3.1 Dataset

The dataset considered image data generated by a large hub airport in Northeast Asia and an international hub airport in Asia. The datasets included dangerous goods and images of the baggage of ordinary passengers. Dataset images were acquired using a HI-SCAN 6040i X-ray machine and a HI-SCAN 6040-2is HR X-ray machine. The X-ray machine was manufactured by Smiths Detection GmbH (Germany). We also checked Realize, Comprehensive, and Randomize to ensure the accuracy of the data. The dataset used 2,000 RGB image data, and the aviation security process data in our study has a relatively large amount of data compared to other studies. The experiment was performed with a training set of 700 images and a validation set of 300 images, with 1,000 images of "Gun" and 1,000 images of "Knife." The experiment was conducted considering epoch 100 and batch size 8.

3.2 Experiment Results

This experiment is restricted to U-Net and O-Net, and the study analysis considers the *Pixel Accuracy, Accuracy, Loss, Precision, Recall*, and *m-IOU* values. The proposed O-Net model experiments related to the gun show a better *Pixel Accuracy* and *m-IoU*, 95.23% and 98.60%, respectively than the U-Net. The “Knife” experiment shows 97.92% pixel accuracy and 90.86% *m-IoU*.

Table 2. Comparative values of U-Net and O-Net in the Gun scenario

	Gun		
Base Model	Pixel Accuracy	Loss	m-IOU
U-Net	0.9678	0.0172	0.8389
O-Net (proposed model)	**0.9860**	0.0165	**0.9523**

Table 3. Comparative values of U-Net and O-Net in the Knife scenario

	Knife		
Base Model	Pixel Accuracy	Loss	m-IOU
U-Net	0.9522	0.0072	0.8251
O-Net (proposed model)	**0.9792**	0.0054	**0.9086**

Conversely, the O-Net performance index was improved compared to U-Net. In Gun detection, *Pixel Accuracy* and *Recall* increased by approximately 6% and 8%. In the knife detection scenario, the proposed method improved by approximately 7%, and 10%, respectively.

Table 4. Performance Measure of U-Net and O-Net for the Gun scenario

	Gun		
Base Model	Accuracy	Precision	Recall
U-Net	0.9080	0.9578	0.8853
O-Net (proposed model)	**0.9692**	**0.9802**	**0.9671**

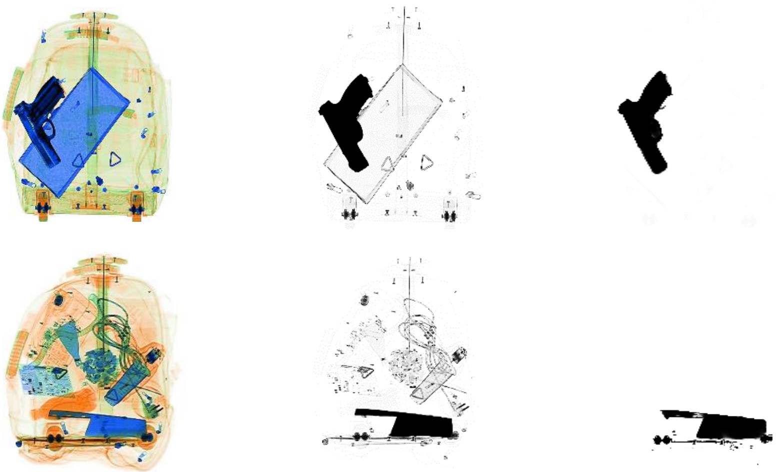
Table 5. Performance Measure of U-Net and O-Net for the Knife scenario

	Knife		
Base Model	Accuracy	Precision	Recall
U-Net	0.8652	0.9223	0.8456
O-Net (proposed model)	**0.9352**	**0.9466**	**0.9462**

Therefore, the results shows that the proposed O-Net architecture has a very high detection rate of guns and knifes with a very high accuracy. Fig. 2 shows output results from the proposed model.

4 Conclusion

The proposed O-Net network was derived starting from the U-Net to improve its performance. The accuracy of O-Net was 6.56% higher than that of U-Net, showing the excellent performance of O-Net. As shown below, Fig. 2(a) is the original image file with a gun and knife, which are dangerous goods in the baggage, while Fig. 2(b) is the Ground Truth indicating the correct answer. Fig. 2(c) shows an experiment with the O-Net structure.



(a) (b) (c)

Fig. 2. (a) is Original image, (b) is Ground Truth, and (c) is the O-Net

References

1. The 9/11 Commission Report: The Final Report of the National Commission on Terrorist Attacks upon the United States, Barnes & Noble Publishing, Inc., New York, USA, (2004).
2. Klenka, M.: Major incidents that shaped aviation security, Journal of Transportation Security 12(1) pp. 39–56 (2019).
3. Oum, T.H., and Xiaowen F.: Air transport security user charge pricing: an investigation of flat per-passenger charge vs. ad valorem user charge schemes, Transportation Research Part E: Logistics and Transportation Review 43(3), pp. 283–293 (2007).

4. Coughlin, C.C.,.Cohen J.P., and Khan S.R.: Aviation security and terrorism: a review of the economic issues, Research Department, Federal Reserve Bank of St. Louis (2002).
5. Vagner, J., et al: Fatigue and stress factors among aviation personel, Acta Avionica 20 pp. 1335–947 (2018).
6. Demao Y., et al: An auxiliary intelligent identification system for contraband of x-ray machine, AOPC 2020: Vol. 11565. International Society for Optics and Photonics, (2020).
7. International Air Transport Association. (2019). IOSA Standards Manual. [Online]. Available, https://www.iata.org/en/iata-repository/publications/iosa-audit-documentation/iosa-standards-manual-ism-ed-132/
8. Krug K.D., Aitkenhead W.F., Eilbert R.F., Stillson, J.H., and Stein J.A.: Detecting explosives or other contraband by employing transmitted and scattered X-rays, U.S. Patent 5 600 700, Feb. 4, (1997).
9. Donnelly, N., Muhl-Richardson, A., Godwin, H., and Cave, K.: Using eye movements to understand how security screeners search for threats in X-ray baggage, Vision, 3(2) pp. 24, (2019).
10. Schwaninger, A., Hardmeier, D., and Hofer, F.: Aviation security screeners visual abilities & visual knowledge measurement, IEEE Aerospace and Electronic Systems Magazine, pp. 29–35, (2005).
11. Bolfing, A., Halbherr, T., Schwaninger, A.: How image based factors and human factors contribute to threat detection performance in X-ray aviation security screening, in Proc. Symp. Austrian HCI Usability Eng. Group. Berlin, Germany: Springer, pp. 419–438(2008).
12. Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional networks for biomedical image segmentation, in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, pp. 234–241(2015).
13. Visa, S., Ramsay, B., Ralescu, A.L., and Van Der Knaap, E. (2011). Confusion Matrix-based Feature Selection. MAICS, 710, 120–127.