

# X-Ray Security Checkpoint System Using Storage Media Detection Method Based on Deep Learning for Information Security

Han-Sung, Lee<sup>†</sup>, Kang-San Kim<sup>††</sup>, Won-Chan, Kim<sup>†††</sup>,  
Tea-Kun, Woo<sup>††††</sup>, Se-Hoon, Jung<sup>†††††</sup>

## ABSTRACT

Recently, as the demand for physical security technology to prevent leakage of technical and business information of companies and public institutions increases, the high tech companies are operating X-ray security checkpoints at building entrances to protect their intellectual property and technology. X-ray security checkpoints are operated to detect cameras and storage media that may store or leak important technologies in the bags of people entering and leaving the building. In this study, we propose an X-ray security checkpoint system that automatically detects a storage medium in an X-ray image using a deep learning based object detection method. The proposed system consists of an edge computing unit and a cloud-computing unit. We employ the RetinaNet for automatic storage media detection in the X-ray security checkpoint images. The proposed approach achieved mAP of 95.92% on private dataset.

**Key words:** Artificial Intelligence, Deep Learning, Information Security System, Object Detection, Retinanet, Storage Media Detection, X-Ray Security Screen

## 1. INTRODUCTION

X-ray security checkpoints are widely being used for homeland security at airports and ports during entry and exit. X-ray security checkpoints used in airport and port immigration offices are used to search for weapons, explosives, etc., including firearms, which may be used for terrorism in passengers' bags or pose a risk to aircraft or ships [1-5]. However, finding dangerous items in passengers' bags is done manually by security guard. Since many passengers' bags must be inspected one by one, there is always the risk of missing or passing dangerous items as the fatigue of security guard increases. In addition, the accu-

racy of finding dangerous objects in a passenger's bag varies greatly depending on the skill level of the security guard. In order to solve the aforementioned problem, various studies are being conducted to automatically detect dangerous objects in the image of the X-ray security checkpoint by means of object detection within X-ray Images [6-9]. Among them, many studies have adopted deep learning based object de-tection algorithms.

Meanwhile, companies with advanced technology are operating X-ray security checkpoints at the entrance and exits of buildings to protect the intellectual property rights and technologies of companies recently. When a company's core technical information and confidential information are

※ Corresponding Author: Se-Hoon, Jung, Address: (57922) 255, Jungang-ro, Suncheon-si, Jeollanam-do, Republic of Korea, TEL: +82-61-750-3622, FAX: +82-61-750-3620, E-mail: shjung@snu.ac.kr

Receipt date: Sep. 14, 2022, Approval date: Sep. 19, 2022  
† School of Creative Convergence, Andong National University (E-mail: mohan@anu.ac.kr)

†† Co., Ltd TNMTECH(E-mail: kskim@tnmtech.com)

††† Co., Ltd TNMTECH(E-mail: wck0326@tnmtech.com)

†††† Co., Ltd TNMTECH(E-mail: tkwoo@tnmtech.com)

††††† Dept. of Computer Engineering, Suncheon National University

※ This work was supported by a Research Grant of Andong National University (2021)

leaked, not only the direct damage, but also the estimable potential loss is truly enormous. For a long time in the industrial field, research and development of technologies such as Data Rights Management (DRM) and Data Loss Protection (DLP) for preventing security accidents have been conducted. The security technologies such as DRM and DLP prevent data from being copied from a computer or data server. This technology is mainly in the form of managing user rights or controlling I/O that can be connected externally, such as USB memory sticks and Internet. However, in the case of data leaked by USB memory sticks or external hard drive, it is no longer possible to control it with cyber security. In order to solve the aforementioned problem, many private research institutes manage to report storage media such as laptop computer, external hard drives, and USB memory sticks that members or guests have when entering and exiting the building, or operate X-ray security checkpoints. The main purpose of the operation of the X-ray security checkpoint is to detect cameras and storage media that may store or leak important technologies in the bags of people entering and leaving the building. As with homeland security applications, problems caused by increased fatigue and skill level of security guard appear the same. As a result, there is an urgent need for an automatic storage medium detection technology at the X-ray security checkpoint.

In this study, we propose an X-ray security checkpoint system for automatically detecting the storage medium using deep learning to prevent the leakage of technical information of high tech companies. The proposed system consists of an edge computing unit and a cloud computing unit. The Edge computing unit detects the storage medium by receiving the image from the X-ray security checkpoint in real time. The cloud computer is connected to multiple edge computers to manage and process security items transmitted from the edge computers. Artificial intelligence technology

developed for individual purposes is managed in the form of nodes and nodes are managed on the cloud platform and constitute required services. In order to automatically detect storage media in X-ray security checkpoint images, RetinaNet is employed, which is one of popular object detection algorithms based on deep learning. In this study, the storage media were divided into 8 classes as follows: Hard disk drive (HDD), Solid state drive (SSD), mobile phone, USB memory stick, pentype USB memory, Secure digital (SD) card, laptop computer, and tablet PC, and the result of the experiment was 95.92% of a mean average precision (mAP).

The main contributions of this study are as follows. First, we defined the problem of automatically detecting storage media at the X-ray security checkpoint from an information protection perspective. Second, this study proposed a cloud based architecture that enables organizations to simultaneously manage multiple X-ray security checkpoints. Finally, the results of basic research and performance evaluation on storage medium detection are presented so that subsequent researchers can refer to them.

The rest of this paper is organized as follows. In section 2, we provide the related works. In section 3, we then propose the X-ray security checkpoint system for storage media detection. Experimental results and explanations are provided in section 4. In section 5, we make discussions of the proposed approach. Finally, some concluding remarks are given in section 6.

## 2. RELATED RESEARCH

### 2.1 Object Detection using X-ray Image

According to the literature review, there are ongoing researches about the threat detection for the security checkpoints using X-ray images. Initial methodologies for object detection in X-ray images adopted the method of extracting traditional hand-

craft features and using them as input for machine learning [10–15]. As deep learning achieves excellent results in general image analysis, deep learning is being actively used in X-ray security image analysis recently [6–9, 16–18]. A. Petrozziello et al. [6] compared the algorithms for firearm part detection in the X-ray images of the baggage. They focused on identifying steel barrel bores as threat objects, being the main part of the weapon needed for deflagration. They reported the comparison of the detection results between CNN, Auto-encoder, Shallow Neural Network with the histograms of oriented Basic Image Features, and Random Forest algorithm with the histograms of oriented Basic Image Features. K. Liang et al. [7] provided the research results of a deep learning based automatic threat detection system for the airports security checkpoints of the transportation security administration in the USA. They collected the data using the Rapiscan 620DV scanner. The threat categories are divided in to firearms (e.g. pistols), sharps (e.g. knives), blunts (e.g. hammers), and LAGs (e.g. liquid-filled bottles). They employed the object detection approach for the threat detection system. This study reported the results of Single Shot MultiBox Detector with Inception V2, Faster R-CNN with ResNet101, Faster R-CNN with ResNet152, Faster R-CNN with Inception ResNetV2. S. Akcay et al. [8] provided the literature review about the automated security systems and its algorithms for security checkpoints with X-ray image. The X-ray Security Imaging is categorized into Machine Learning based Algorithm and Deep Learning based Algorithm. These approaches are divided into object classification, object detection, and object segmentation. The Accuracy, Mean Average Precision (mAP), Area Under Curve (AUC) are used for the evaluation criteria. S. Akcay et al. [9] provided the X-ray baggage image classification approach employing the convolutional neural networks with transfer learning. They tried to classify the firearm, firearm components,

knives ceramic knives, camera, laptop.

## 2.2 Machine Learning and Deep Learning for Object Detection

An object detection algorithm is generally comprised of three processes: first, object candidate regions are selected in a video; second, the features of each candidate region are extracted; and third, the object candidate regions are subjected to multi class classification with a classifier applied to the extracted features. Localization follows classification through the regression of bound boxes according to object detection methods. In deep learning based object detection methods, the essence is the extraction of features with CNN. Deep learning-based object detection techniques are mainly divided into one-stage and two-stage object detection models according to structural features and methods [19–20].

Research was conducted on one-stage object detection models to supplement the low processing speed of two-stage object detection models, which cause high calculation costs in mobile or wearable systems with limited storage and calculation memory. A single integrated pipeline is used to predict directly class classification and the regression learning of bound boxes, which used to happen in two stages. A one-stage object detection model uses a predefined anchor box with no region proposals at the stage of finding object candidate regions. An anchor box helps to perform one-stage object detection based on a single deep learning model to classify and carry out bound box regression. Good examples of one-stage object detection models are YOLO and SSD [8, 21–23].

Two-stage object detection models conducted the first research based on CNN as an object detection algorithm, which processes region proposals with input videos. Region proposals find regions where input objects might be. This approach can detect objects faster than the old approaches including the sliding window method thanks to re-

duced computation. It selectively explores regions with the high possibilities of containing objects or utilizes a computer vision technology such as an edge box. It sequentially performs localization with a deep learning based network of region proposals. Two-stage object detection models then perform classification and more detailed localization in the selected candidate regions, thus recording higher accuracy than one-stage ones. Good examples include R-CNN, Fast R-CNN, and Faster R-CNN [5-7, 24-27].

### 3. X-RAY SECURITY CHECKPOINT SYSTEM FOR STORAGE MEDIUM DETECTION

In this study, we propose an X-ray security checkpoint system for automatically detecting the storage medium to prevent the leakage of important information of companies and institutes. The proposed system acquires the X-ray image from the X-ray security checkpoint, then applies a deep learning algorithm to detect storage medium in X-ray images at the edge computing device. The detection results of the storage medium at the edge computing unit are transmitted to the cloud com-

puting unit. The cloud computing unit analyzes the data received from the edge computing unit more closely and continuously retrain the AI algorithm. Overall architecture of the proposed system is presented in Fig. 1. The structure of the proposed system includes: 1) the edge computing unit that acquires images from X-ray security checkpoint and detects storage medium. 2) A cloud computing unit that analyzes the storage medium detection result of the edge computing unit and retrains the AI node, that is, the storage medium detection algorithm. 3) AI Node, an artificial intelligence algorithm that detects storage medium in X-ray images.

The storage medium detection algorithm, AI Node, is trained on cloud computing units and deployed to each edge computing unit. The edge computing unit detects the storage medium in the X-ray image acquired from the X-ray security checkpoint using the pretrained AI Node. The edge computing unit transmits the storage medium detection result to the cloud computing unit. The cloud-computing unit stores and analyzes the storage medium detection results received from multiple edge computing units. The storage me-

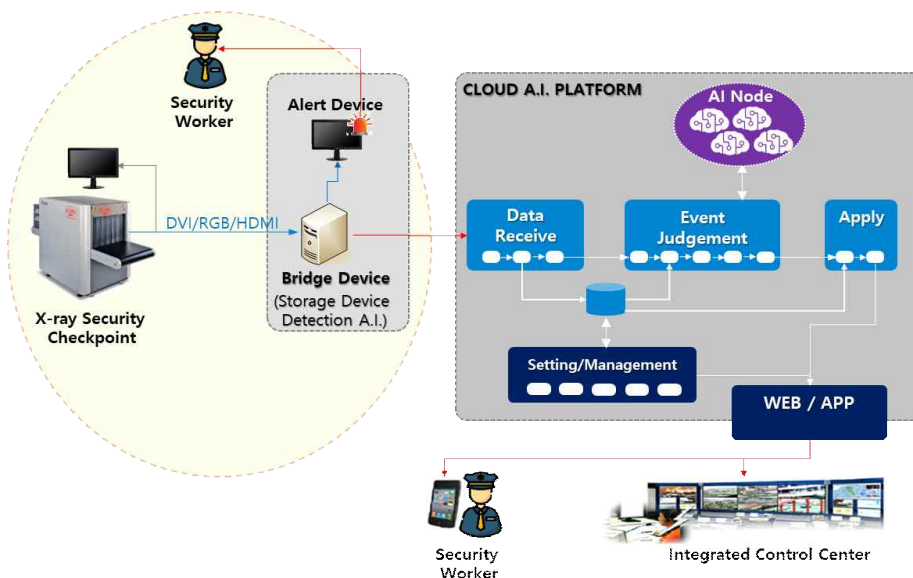


Fig. 1. Overall of architecture of the proposed X-ray security checkpoint system for storage medium detection.

dium detection results are divide into 3 classes, i.e., True Positive instances, False Positive instances and False Negative instances. The cloud-computing unit continuously performs re-training the AI Node using the analyzed data to increase the performance of the AI Node. The cloud-computing unit distributes the retrained AI Node to the edge-computing units.

### 3.1 Edge Computing Unit

The edge computing unit consists of a legacy X-ray security checkpoint and a bridge device equipped with an AI Node, a storage medium detection module. The bridge device receives the updated/trained AI Node from the cloud computing unit and uses it to detect storage medium. The bridge device receives the image acquired from the X-ray scanner as input and detects the storage medium. In addition, the bridge device outputs the alarm and the storage medium detection image together so that the security guard can make a final decision. Security guard look at the results of the storage medium detection at the edge computing unit and, if there is a False Negative or False Positive, input it into the system and generate metadata. The edge computing unit transmits the results of storage media detection and metadata to the cloud computing unit for secondary image

analysis and AI Node performance enhancement. Fig. 2 shows the operational flow of the edge computing unit.

### 3.2 Cloud Computing Unit

Fig. 3 is the system configuration diagram of the cloud computing unit, which is consists of Data Management Module, Event Management Module, Data Annotation Module, Continuous Learning Module and AI Node Management Module. The data management module receives X-ray images, storage medium detection results, and metadata from the edge computing unit, and classifies the data into True Positive, False Positive, and False Negative data instances based on the metadata. The classified data is stored in the data storage. The event management module sends an alert to the inte-grated control center through the API Layer based on the True Positive data among the information received from the data management module. The data annotation module performs annotation on False Positive data instances and False Negative data instances among data stored in data storage. Humans manually perform annotations for False Positive data instances and False Negative data instances using the annotation interface that is provided by the data annotation module. False Positive data instances are labeled with the others

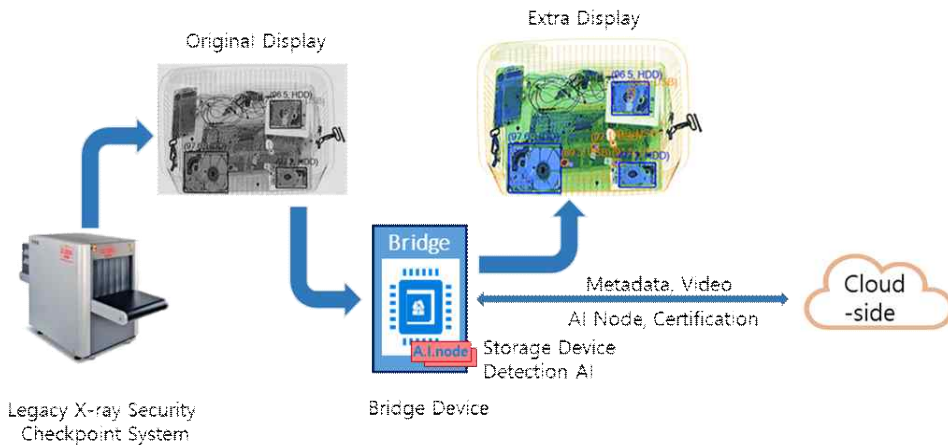


Fig. 2. The operational flow of the edge-computing unit.

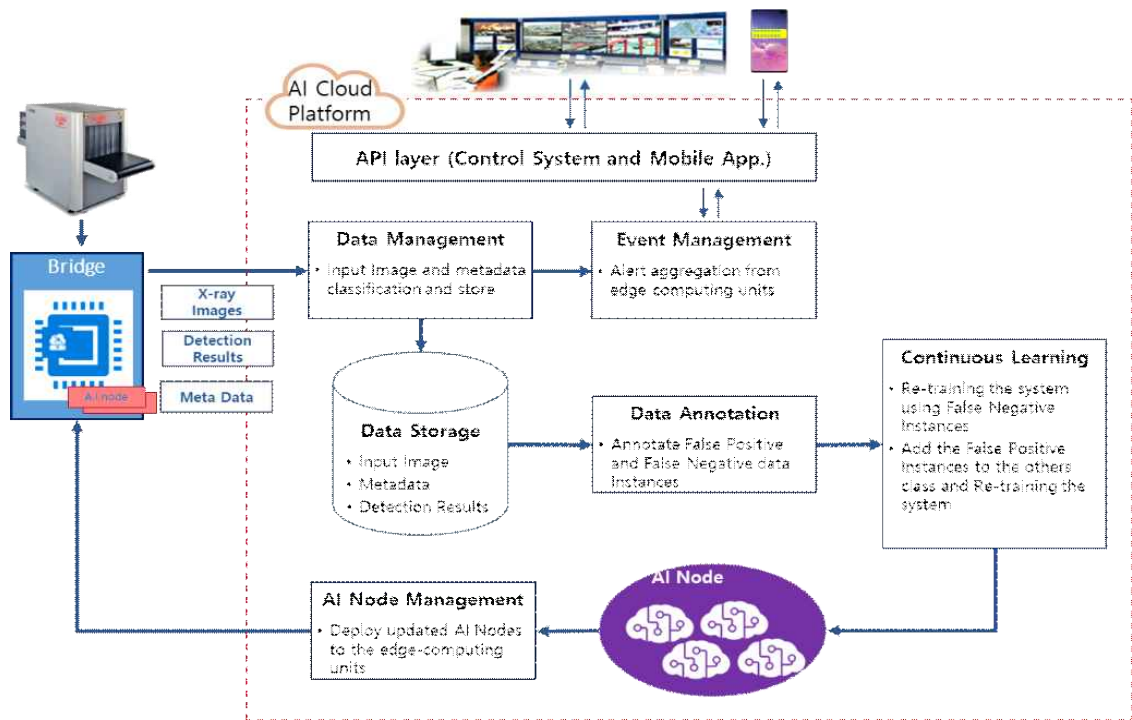


Fig. 3. The system configuration diagram of the cloud-computing unit.

class. The others class is a special class composed of object instances that have a shape and color similar to the target objects. In this study, we introduced the others class to reduce the False Positive rate of the proposed system. The continuous learning module continuously trains AI Node to improve the performance of the system by using the new training dataset generated by the data annotation module. The AI Node management module periodically distributes the newly updated AI Node to the edge computing unit to improve and maintain the best performance of the system.

### 3.3 AI Node: the Storage Medium Detection Module

In this paper, a deep learning based object detection algorithm is adopted as a method to automatically detect storage media in x-ray security checkpoint images to prevent information leakage from Hi-tech companies. The object detection method based on deep learning consists of finding

a candidate region of an object and classifying the image of the candidate region. Object detection algorithms based on deep learning can be largely divided into one stage object detector and two-stage object detector. In the case of a one-stage object detector such as the YOLO series, the speed is very fast and the accuracy is low compared to the two-stage object detector such as the R-CNN series. This is due to the extreme class imbalance problem. The R-CNN series algorithm classifies and filters the foreground and background to some extent through the Region Proposal Network, but in the case of a one-stage object detector, a relatively large number of background areas are included because the grid of the feature map is used. This problem causes the class imbalance problem of the data.

To solve this problem, The RetinaNet [20] uses the Focal Loss, which slightly modified the cross entropy loss function commonly used in class classification. Focal Loss gives small weights to

well categorized examples, while large weights are assigned to some examples that are difficult to classify, thus concentrating learning on difficult examples. Consequently, the RetinaNet is able to solve the problem that learning is overwhelmed by most of the negative samples that are easily classified. In addition, if Focal Loss is employed, we can get more than a two-stage network in a one-stage network. The RetinaNet is as fast as a one-stage detector, yet has an accuracy that surpasses all existing top performing detectors [20]. In this study, we employed the RetinaNet as the storage media detector in X-ray security checkpoint images. We define the storage media detection as object detection problem of 8 classes of object, i.e., HDD, SSD, mobile phone, USB memory stick, pentype USB memory, SDCARD, laptop computer and tablet PC. In addition, others class includes auxiliary battery and portable fan. The others class is a class to reduce False Positives and the system is trained using False Positive data instances.

## 4. EXPERIMENTAL RESULTS

### 4.1 Experimental Setup

In this study, the performance was evaluated using private data set to validate the proposed system. The target storage medium was placed in

various types of bags, and images were acquired while passing the bags through an x-ray inspection machine at security checkpoint. We collected a total 10 types of storage media including HDD, SSD, mobile phone, USB memory stick, pentype USB memory, SDCARD, laptop computer, and tablet PC. 80% of the acquired data was used as training data, and 20% was used as test data. Table 1 provides detailed information about the collected dataset.

The dataset for initial training was collected in two ways. First, each type of storage medium was collected while passing through the x-ray security checkpoint alone. Second, images were acquired by putting each type of storage medium in various bags and passing through the x-ray security checkpoint. The examples of training dataset is shown in Fig. 4. The number of storage media that can be used when generating training data is limited, but a very large amount of training data is required to learn a deep learning model. In order to solve the above problem, a large number of training data was generated several times by arranging usable storage media in various locations and angles. In the case of the storage medium contained in the bag, images were acquired by varying the position and angle of the bag.

For test dataset, images were acquired by putting each type of storage medium in various bags

Table 1. The description of data set for storage medium detection.

Categories	No. of Image	No. of Objects	Ratio
HDD	4,191	23,192	19.17%
SSD	1,932	10,692	8.84%
Mobile Phone	2,918	16,146	13.35%
USB Memory Stick	7,578	41,937	34.67%
Pen-Type USB Memory	632	3,497	2.89%
SDCARD	2,194	12,141	10.04%
Laptop Computer	417	2,310	1.91%
Tablet PC	514	2,846	2.35%
Auxiliary Battery	1,121	6,203	5.13%
Portale Fan	362	2,004	1.66%
Total	21,860	120,968	100.00%

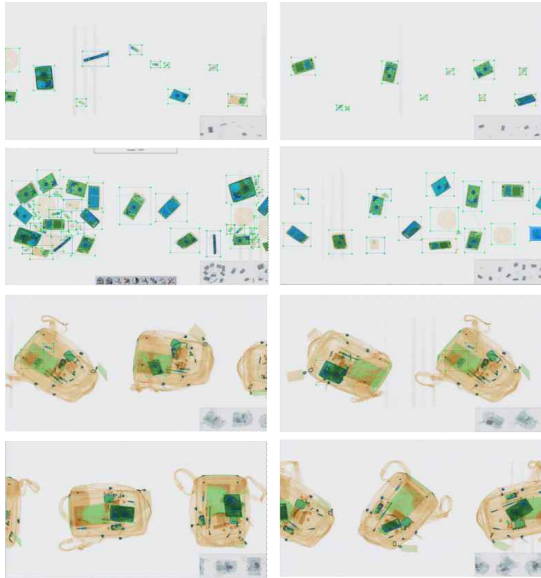


Fig. 4. Example images of initial training dataset.

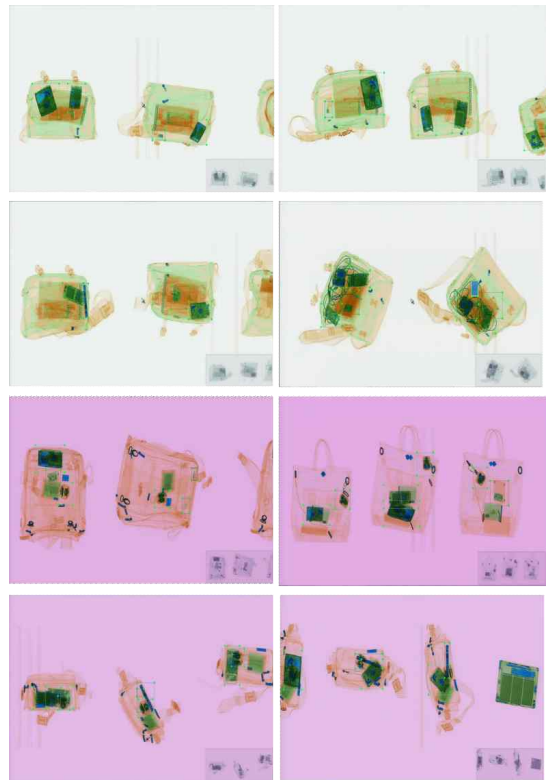


Fig. 5. Example images of test dataset.

and passing through the x-ray security checkpoint. In order to acquire test data in an environment similar to the real world, images were acquired by additionally putting books, notebooks, writing instruments, clothes, and wallets in the bag in addition to the storage medium. The examples of test dataset is shown in Fig. 5.

#### 4.2 Evaluation Results

Table 2 shows the evaluation results for 8 object classes: HDD, SSD, mobile phone, USB memory stick, pen-type USB memory, SDCARD, laptop computer and tablet PC. The average precision of the HDD, SSD, mobile phone, USB memory stick, laptop computer and tablet PC are over 99% and

the average precision of the pen-type USB memory and SDCARD are 65.63% and 96.39% respectively. The proposed approach achieved a mean average precision (mAP) of 95.92% on our dataset in overall.

Fig. 6 below is an example of a storage medium detected by the proposed system. It shows that storage media are detected relatively well in the X-ray images taken at the X-ray security checkpoints in various environments.

Fig. 7 shows example images of a falsely detected object by the proposed system. (a) shows that the metal decoration of the bag was mistakenly recognized as a mobile phone. In (b), the metal decoration of the bag was recognized as a

Table 2. Evaluation Results (Average Precision).

HDD	SSD	PHONE	USB	USB_PEN	SDCARD	LAPTOP	TABLET	mAP
0.9991	0.9989	0.9979	0.9962	0.6562	0.9639	0.9995	0.9925	0.9592



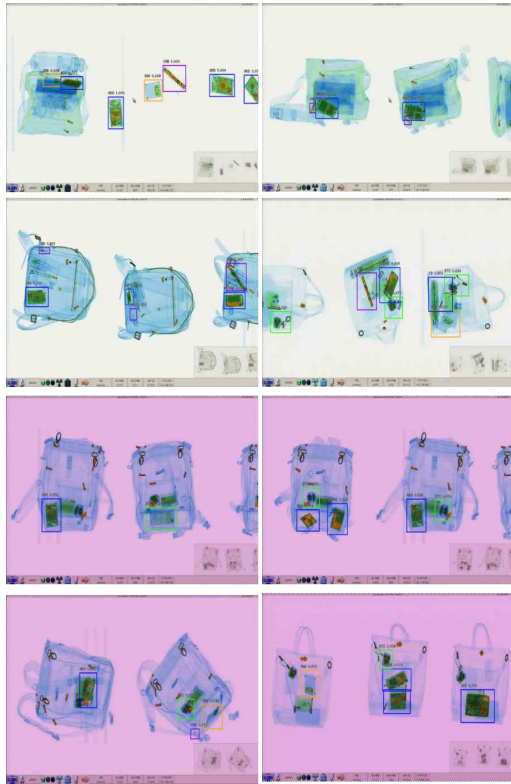


Fig. 6. Example of a storage medium detected by the proposed method.

USB memory stick. (c) and (g) are examples of recognizing the glasses case as a mobile phone. (d) and (h) are examples of incorrectly recognizing the metal zipper of the bag as a USB memory stick. In (e), the tip of the umbrella was mistakenly recognized as a pen type USB memory. In (e) and (f), a specific part of the umbrella was mistakenly recognized as a pen type USB memory.

### 5. DISCUSSIONS

In the field of deep learning based object detection, researchers have investigated a variety of methods to enhance model performance including model network structures, extraction methods of candidate regions, and class classification methods in terms of speed as well as accuracy. They are currently focusing attention on performance enhancement with one-stage object detection models with faster speed. The present study reviewed previous studies on the old object detection algorithms to make use of RetinaNet. The review covered one-stage object detection techniques such as YOLO and SSD and

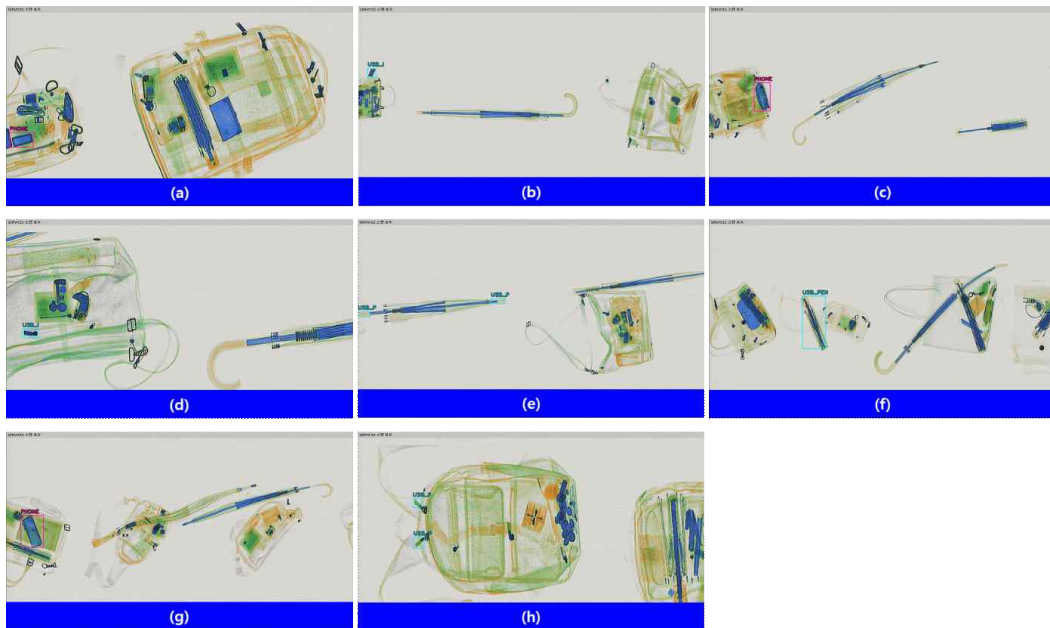


Fig. 7. Example images of a falsely detected object by the proposed system.

two-stage ones such as R-CNN, Fast R-CNN, and Faster R-CNN [5-8, 19-26].

Proposed to detect objects rather than classification like the old CNN, R-CNN can detect multiple objects instead of a single one in one image. The SVM classifier and the bound box regression learning require a lot of disc space and time. CNN is applied to the features of each image, which increases calculation costs. R-CNN has approximately 2,000 proposed regions for a single image, thus presenting a couple of disadvantages including many calculations, low speed, and complicated learning processes comprised of multiple stages. Fig. 8 shows the structure map of R-CNN, which defines the region proposal of an input image and classifies classes by calculating its features [5, 24].

Fast R-CNN is a neural network that improves the low speed of R-CNN. Fast R-CNN uses selective explorations to find region proposals just like R-CNN, but it does not classify the proposed 2,000 regions from an image right away unlike R-

CNN. It moves the entire input video through CNN, extracts a feature map, and detects region proposals through selective explorations. In addition, it combines classifier and bound box loss values and trains them at the same time, thus simplifying the training stage. It uses softmax instead of SVM for classification, improving its performance. ROI pooling was applied to extract the features of distance fixed between the last convolution layer and the first complete connection layer. Unlike R-CNN, Fast R-CNN moves CNN through the entire input video just one time instead of moving each proposed region through CNN to detect objects, thus reducing computation relatively from R-CNN and improving the processing speed. Fig. 9 shows the structure map of Fast R-CNN applied in the present study [6, 25].

Faster R-CNN is a neural network proposed to improve the bottleneck phenomenon of region proposals, which is an issue with Fast R-CNN. It eliminates the method of extracting regions

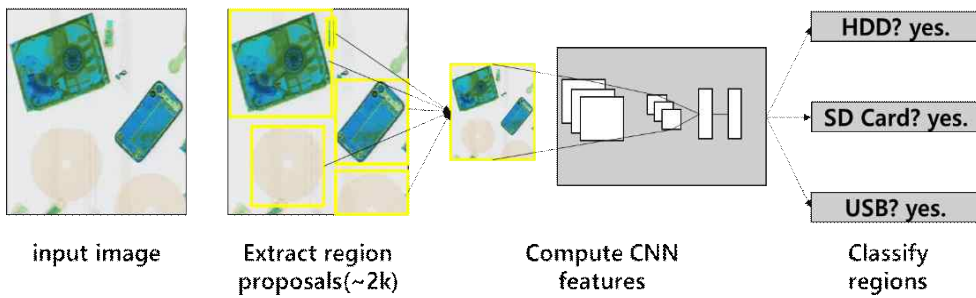


Fig. 8. The structure of class classification using R-CNN.

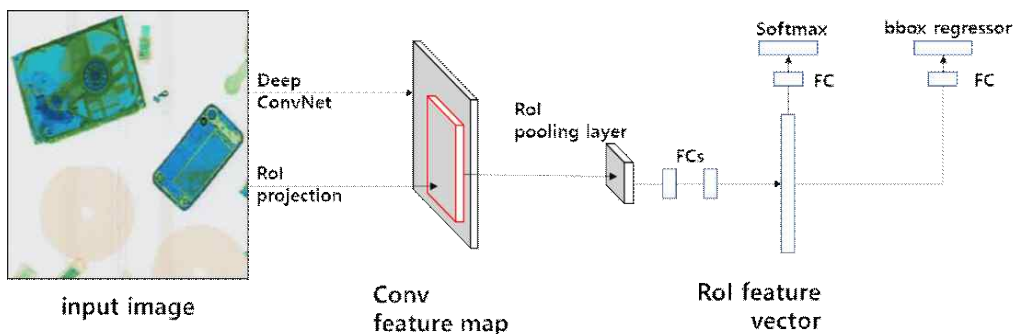


Fig. 9. The structure of class classification using Fast R-CNN.

through a selective exploration algorithm that causes a bottleneck in Fast R-CNN and instead uses a region proposal network(RPN). The convolution process was shared between RPN and Fast R-CNN. The features that passed the last shared convolution layer were used in object class and region classification in separate regions. Fig. 10 shows the structure map of RPN, which is a network to extract proposed regions by estimating object locations based on the input of feature maps as the output of CNN in Fast R-CNN. RPN is divided into object proposals to signify bound box locations and object scores to tell whether there are objects or not. RPN introduces the anchor concept to detect objects of various sizes and proportions in each slide. An anchor is a bound box of many different predefined sizes and proportions. When a sliding window passes through a feature map, each slide creates bound box candidates in k number, 2k scores for classification, and 4k coordinates for regression learning. Faster R-CNN has improved from the old Fast R-CNN in terms of detection accuracy and processing speed, but it would face the old issues again without using the high performance hardware of X-ray security checkpoints [7, 26].

As a one-stage object detection technique, YOLO uses the grid method to detect the bound

boxes of objects. Input images are divided in a grid of  $S \times S$ , and each grid cell has a B number of bound boxes and reliability scores of bound boxes. Reliability scores are the probability values of objects present inside a bound box. When there are no objects in a grid cell, it will be 0. Each grid cell will have a C number of conditional probability values, which indicate the class of an object inside each bound box. It increases the speed by approximately five times by omitting the RPN process. YOLO checks the entire images at once for prediction, encoding implicitly the context information of an object class. It recorded a low rate of detecting wrong backgrounds and failed to detect small objects due to the location and size of a bound box and the rough division of the aspect ratio. Fig. 11 shows the structure map of YOLO [8, 21-22].

SSD is a neural network with improved performance and speed from YOLO. SSD uses the same VGG-19 network as CNN of Faster R-CNN, being a single shot detector that uses a single network like YOLO. SSD finds the bound boxes of objects and classifies their classes with a single network. In SSD, a feature map would have six output layers with different sizes including  $38 \times 38$ ,  $19 \times 19$ ,  $10 \times 10$ ,  $5 \times 5$ ,  $3 \times 3$ , and  $1 \times 1$ . When there are objects of different sizes, it would detect them across multiple scales with a convolution feature

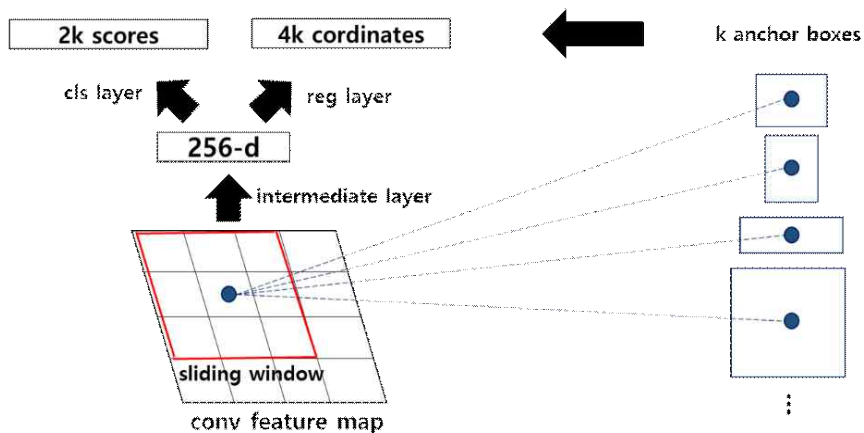


Fig. 10. The structure of class classification using Faster R-CNN.

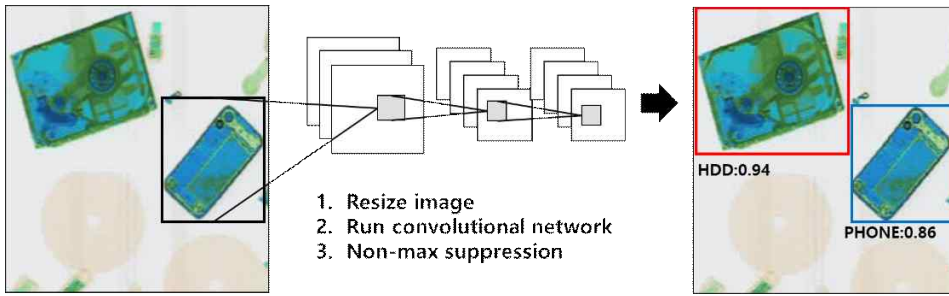


Fig. 11. The structure of class classification using YOLO.

map. Each feature map predicts the category scores and offsets of bound boxes in proper size. SSD improves its performance by detecting objects of different sizes on six feature maps of different sizes. It boasts higher detection speed than two-stage object detection techniques, but it has lower accuracy in the detection of small objects like YOLO. Fig. 12 shows the structure map of SSD [23].

These findings indicate that one-stage object detection methods such as YOLO and SSD improved object detection speed from two-stage object detection methods such as R-CNN, but they have a performance issue with 10~40% lower accuracy. One-stage object detection methods have inferior performance due to class imbalance issues [16]. They would detect backgrounds with no reduced weights as objects as most bound boxes to detect objects contained no objects in more cases. This issue raised a concern with model training to detect backgrounds rather than objects. The present study suggested that the class imbalance is-

issues in X-ray images could be solved by granting bigger or smaller loss values to certain class data in the calculation process.

Trying to solve this issue, the present study used the RetinaNet [9, 28] model and confirmed the possibilities of improving an object detection model with it. It will open a door for the application of an object detection technology to various X-ray images. The proposed object detection model can also be improved to learn diverse kinds of objects and be applied to necessary fields. A follow-up study will investigate an object detection model to detect and classify overlapping objects accurately by improving the proposed object detection model.

## 6. CONCLUSIONS

Many of the high-tech companies are recently operating X-ray security checkpoints at the entrance and exits of buildings to protect the intellectual property rights and technologies of com-

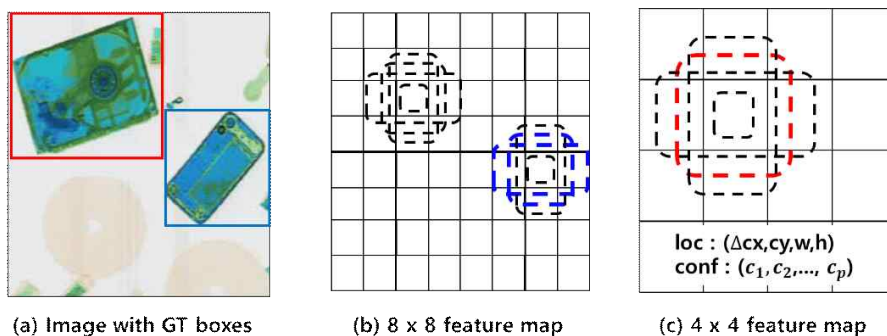


Fig. 12. The structure of class classification using SSD.

panies recently. However, in many cases, security guard are directly detecting storage media from X-ray images. As a result, there is a need for an automatic storage medium detection technology at the X-ray security checkpoint. In this study, we propose an automatic detection method of storage media using deep learning in X-ray security checkpoint images to prevent the leakage of technical information of high tech companies. We employed the RetinaNet, which is as fast as a one-stage detector, yet has an accuracy that surpasses all existing top performing detectors, as the storage media detector in X-ray security checkpoint images. We define the storage media detection as object detection problem of 10 classes of object, i.e., HDD, SSD, mobile phone, USB memory stick, pentype USB memory, SDCARD, laptop computer, tablet PC, auxiliary battery, and portable fan. According to the experimental results, the proposed approach achieved a mean average precision (mAP) of 95.92 % on private dataset. For future research, we plan to conduct research on the image preprocessing process in consideration of various X-ray image acquisition environments, and we plan to conduct research on object detection algorithms and image segmentation specialized in detecting storage media in X-ray images.

## REFERENCE

- [ 1 ] A. Chavallaz, A. Schwaninger, S. Michel, and J. Sauer, "Expertise, Automation and Trust in X-Ray Screening of Cabin Baggage," *Frontiers in Psychology*, Vol. 10, pp. 256, 2019.
- [ 2 ] S. Baeriswyl, A. Krause, and A. Schwaninger, "Emotional Exhaustion and Job Satisfaction in Airport Security Officers Work Family Conflict as Mediator in the Job Demands Resources Model," *Frontiers in Psychology*, Vol. 7 pp. 663, 2016.
- [ 3 ] T. Halbherr, A. Schwaninger, G.R. Budgell, and A. Wales, "Airport Security Screener Competency: A Cross-Sectional and Longitudinal Analysis," *The International Journal of Aviation Psychology*, Vol. 23, No. 2, pp. 113-129, 2013.
- [ 4 ] L. Swann, V. Popovic, A.L. Blackler, and B.J. Kraal, "Airport Security Screeners Expertise and Implications for Interface Design," *Proceeding of International Design Research Society Conference*, pp. 1465-1481, 2014.
- [ 5 ] S. Michel, S. M. Koller, J.C. de Ruiter, R. Moerland, M. Hogervorst, and A. Schwaninger, "Computer-Based Training Increases Efficiency in X-Ray Image Interpretation by Aviation Security Screeners," *Proceeding of 41st Annual IEEE International Carnahan Conference on Security Technology*, pp. 201-206, 2007.
- [ 6 ] A. Petrozziello, and I. Jordanov, "Automated Deep Learning for Threat Detection in Luggage from X-Ray Images," *Proceeding of International Symposium on Experimental Algorithms*, pp. 505-512, 2019.
- [ 7 ] K. Liang, J. B. Sigman, G. P. Spell, D. Strellis, W. Chang, F. Liu, et al., "Toward Automatic Threat Recognition for Airport X-ray Baggage Screening with Deep Convolutional Object Detection," *arXiv Preprint*, arXiv: 1912.06329, 2019.
- [ 8 ] S. Akcay and T.P. Breckon, "Towards Automatic Threat Detection: A Survey of Advances of Deep Learning within X-Ray Security Imaging," *arXiv Preprint*, arXiv: 2001.01293, 2020.
- [ 9 ] S. Akcay, M.E. Kundegorski, M. Devereux, and T.P. Breckon, "Transfer Learning Using Convolutional Neural Networks for Object Classification within X-Ray Baggage Security Imagery," *Proceeding of 2016 IEEE International Conference on Image Processing*, pp. 1057-1061, 2016.
- [ 10 ] M. Bastan, M.R. Yousefi, and T.M. Breuel, "Visual Words on Baggage X-Ray Images," *Proceeding of International Conference on Computer Analysis of Images and Patterns*,

- pp. 360–368, 2011.
- [11] D.G. Lowe, “Object Recognition from Local Scale Invariant Features,” *Proceeding of the International Conference on Computer Vision*, pp. 1150–1157, 1999.
- [12] D. Turcsany, A. Mouton, and T.P. Breckon, “Improving Feature-Based Object Recognition for X-Ray Baggage Security Screening Using Primed Visual Words,” *Proceeding of the IEEE International Conference on Industrial Technology*, pp. 1140–1145, 2013.
- [13] M. Bastan, W. Byeon, and T. Breuel, “Object Recognition in Multi-View Dual Energy X-ray Images,” *Proceeding of the British Machine Vision Conference*, p. 11, 2013.
- [14] D. Mery, E. Svec, and M. Arias, “Object Recognition in Baggage Inspection Using Adaptive Sparse Representations of X-Ray Images,” *Proceeding of Image and Video Technology*, pp. 709–720, 2016.
- [15] E.M. Kundegorski, S. Akcay, M. Devereux, A. Mouton, and T.P. Breckon, “On Using Feature Descriptors as Visual Words for Object Detection within X-Ray Baggage Security Screening,” *Proceedings of the International Conference on Imaging for Crime Detection and Prevention*, p. 12, 2016.
- [16] N. Jaccard, T.W. Rogers, E.J. Morton, and L.D. Griffin, “Using Deep Learning On X-Ray Images To Detect Threats,” *Proceeding of Cranfield Defence and Security Doctoral Symposium*, pp. 1–12, 2016.
- [17] J. Yuan and C. Guo, “A Deep Learning Method for Detection of Dangerous Equipment,” *Proceeding of International Conference on Information Science and Technology*, pp. 159–164, 2018.
- [18] M. Xu, H. Zhang, and J. Yang, “Prohibited Item Detection in Airport X-Ray Security Images via Attention Mechanism Based CNN,” *Proceeding of Chinese Conference on Pattern Recognition and Computer Vision, Lecture Notes in Computer Science*, pp. 429–439, 2018.
- [19] H.S. Lee, K.S. Kim, and S.H. Jung, “Deep Learning-Based Storage Media Detection for X-Ray Security Checkpoints,” *Proceedings of the 16th International Conference on Multi-media Information Technology and Applications*, pp. 20–21, 2020.
- [20] T.Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal Loss for Dense Object Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, No. 2, pp. 318–327, 2018.
- [21] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” *Proceeding of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7263–7271, 2017.
- [22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” *Proceeding of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, 2016.
- [23] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, et al., “SSD: Single Shot MultiBox Detector,” *Proceeding of European Conference on Computer Vision*, pp. 21–37, 2016.
- [24] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation,” *Proceeding of 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, 2014.
- [25] R. Girshick, “Fast R-CNN,” *Proceeding of IEEE International Conference on Computer Vision*, pp. 1440–1448, 2015.
- [26] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 6, pp. 1137–1149, 2016.
- [27] Z.Q. Zhao, P. Zheng, S.T. Xu, and X. Wu,

“Object Detection with Deep Learning: A Review,” *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 30, No. 11, pp. 3212–3232, 2019.

- [28] J.Y. Kim, S.H. Jung, and C.B. Sim, “A Study on Object Detection using Restructured Retina Net,” *Journal of Korea Multimedia Society*, Vol. 23, No. 12, pp. 1531–1539, 2020.



Han-Sung, Lee

He received his BS, MS, and PhD degrees in computer science from Korea University, Rep. of Korea, in 1996, 2002, and 2008, respectively. From July 1996 to July 1999, he worked for DAEWOO engineering company, Rep. of Korea. From November 2009 to November 2014, he worked for Electronics and Telecommunications Research Institute (ETRI), Rep. of Korea as a senior member of the research staff. He worked for the Samsung Electronics Co., Ltd., Rep. of Korea, from December 2014 to August 2019. He was with Youngsan University, Rep. of Korea as an assistant professor from September, 2019 to February 2021. He joined Andong National University, Rep. of Korea in March, 2021. His current research interests include machine/deep learning, computer vision, computer/network security, optimization, data mining, and big-data analytics.



Kang-San Kim

He received his BS degree in computer engineering from Youngsan University (Y'sU). He joined Co., Ltd TNMTECH, Rep. of Korea in January, 2021. His current research interests include machine learning, deep learning and computer vision.



Won-Chan, Kim

He received his associate degree in business administration, Seattle Central College in June, 2017. He took a leave of absence from Dept. of data science, the University of Washington, Seattle in July 2019. He joined Co., Ltd TNMTECH, Rep. of Korea as AI software engineer in September, 2020. His current research interests include machine learning, deep learning, big data analytic and visualization, technology statistics, artificial neural network, and image processing.



Tea-Kun, Woo

He received his BS, MS degrees in computer science from Korea University, Rep. of Korea, in 1998 and 2000, respectively. From September 1999 to May 2005, he worked for TeleVideo, USA. From June 2005 to June 2007, he worked for DT Research, USA. He worked for the Nano Technology Lab., USA, from July 2007 to February 2012. He was with Dual Aperture, USA from March, 2012 to May, 2017. He joined Co., Ltd TNMTECH, Rep. of Korea, as a vice president in June, 2018. His current research interests include AIoT, machine learning, deep learning, soft computing, and computer vision.



Se-Hoon, Jung

He received his BS, MS, and PhD degrees in multimedia engineering from Suncheon National University, Suncheon city, Rep. of Korea, in 2010, 2012, and 2017, respectively. He worked for the youngsan university, Rep. of Korea, from September 2018 to February 2020 and he worked for the andong national university, Rep. of Korea from March 2020 to August 2022. He joined in department of computer engineering from Suncheon National University in september, 2022. His current research interests include bigdata processing, data-mining, reinforcement learning, blockchain.