



Delivery Invoice Information Classification System for Joint Courier Logistics Infrastructure

Youngmin Kim¹, Sunwoo Hwang², Jaemin Park¹ and Joouk Kim^{2,*}

¹Department of Systems Engineering, Ajou University, Suwon, 16499, Korea

²Advanced Logistics System Research Department, Korea Railroad Research Institute, Uiwang, 16105, Korea

*Corresponding Author: Joouk Kim. Email: jookim@krri.re.kr

Received: 27 January 2022; Accepted: 08 March 2022

Abstract: With the growth of the online market, demand for logistics and courier cargo is increasing rapidly. Accordingly, in the case of urban areas, road congestion and environmental problems due to cargo vehicles are mainly occurring. The joint courier logistics system, a plan to solve this problem, aims to establish an efficient logistics transportation system by utilizing one joint logistics delivery terminal by several logistics and delivery companies. However, several courier companies use different types of courier invoices. Such a system has a problem of information data transmission interruption. Therefore, the data processing process was systematically analyzed, a practically feasible methodology was devised, and delivery invoice information processing standards were established for this. In addition, the importance of this paper can be emphasized in terms of data processing in the logistics sector, which is expected to grow rapidly in the future. The results of this study can be used as basic data for the implementation of the logistics joint delivery terminal system in the future. And it can be used as a basis for securing the operational reliability of the joint courier logistics system.

Keywords: Joint courier logistics base infrastructure; logistics cooperation; urban public infrastructure; YOLOv4; object detection algorithm

1 Introduction

With the growth of the online market and the resulting growth of the logistics market, delivery volume is increasing, and various studies are being conducted to efficiently handle the increasing delivery volume. Among them, this is the point of time presented in the concept of point courier logistics system that can jointly process multiple courier cargoes. In addition, Joint courier logistics infrastructure is the world's first concept proposed to streamline the logistics process, and has not been used to date. Joint courier logistics system is a sub-terminal concept that classifies and releases multiple courier cargoes received from hub terminals by destination, but for smooth operation, information on courier invoices used by multiple courier companies is required. However, since the waybill used for delivery is used in different configurations for each courier, it is difficult to integrate information



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

because related information is managed individually for each courier, and such a system is expected to be inefficient in handling cargo due to information disconnection. Fig. 1 is the result of the study of Korea Maritime Institute to domestic logistics technology demand survey and analysis. This result shows the problem of disconnection of invoice information caused by each courier company using various forms of courier invoices when operating the provided system. Due to this information disconnection problem, the problem was found to be 16% in Transportation/Loading file, 12% in Cargo/Reservation, 8% in Information Acquisition/DB (Data Base) Construction, 10% in etc., and 7% in Standardization. The field of artificial intelligence have been improving consistently in the past few years since deep learning technologies developed. Due to these development of technologies, the field of object and pattern recognition advanced used by various deep learning algorithms. In the field of delivery, research on the application of detection of objects and patterns is being conducted to process unstructured cargoes of different shapes and sizes. In this paper, we intend to use object detection to solve the information disconnection problem. The results of this paper can be used as the basis data necessary to solve the problem caused by the increase in delivery volume by identifying the increase in cargo vehicles, traffic jams, and environmental problems caused by logistics transportation through individual delivery centers. In order to solve problems of complex smart systems and secure operational reliability, it is intended to present standards for data processing.

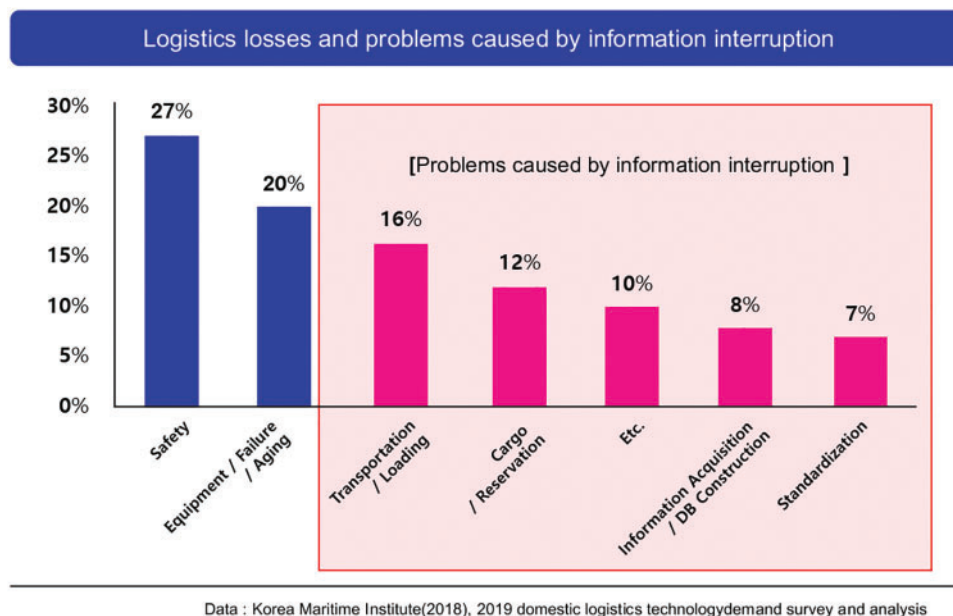


Figure 1: The loss problem of logistics systems due to lack of data

Fig. 2 shows a schematic conceptual diagram of the delivery invoice information detection system. Research to implement an actual physical system will be conducted. A multidimensional tunnel image recognition system will be implemented as a subsystem by utilizing video and image acquisition sensors such as cameras, and a bottom line scanner will be implemented under the path through which the courier cargo passes. The obtained courier cargo transport information data is transmitted to a field processor capable of processing barcode recognition, shape recognition, and delivery invoice information recognition for each courier. The transmitted image and video data are stored in large capacity storage and are processed and classified through an artificial intelligence based deep learning model. As a result of this, delivery invoice information will be transmitted through the interface with

the joint courier logistics system, which will be operated separately. In this study, a study was conducted on ways to improve data classification and object recognition algorithms as a way to increase the operational reliability of the system.

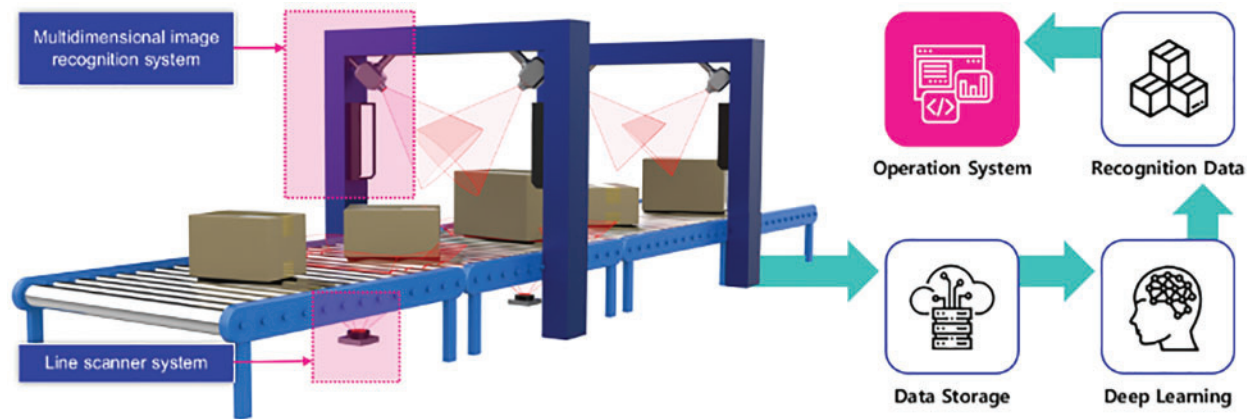


Figure 2: Schematic diagram of the system for recognizing delivery invoice information

1.1 Related Literature Review

Current logistics is not just a process, but a combination of various technologies, showing the aspect of complex technologies. In order to implement an efficient logistics system, basic research through various approaches and technology development in which various fields are fused is essential. Accordingly, this paper proposes an invoice classification system according to an object detection algorithm. Chae et al. reviewed cases of public logistics-related business models and policy directions, and determined that a rational and systematic management and support system is necessary to implement a successful community logistics business model [1]. Shin conducted a study in order to applying the technology originated from the 4th industrial revolution to the field of logistics technology, such as IoT, AI and robot. In aspects of, it was judged that the combination of artificial intelligence and robots can well replace existing manpower and increase the efficiency of the logistics system [2]. Kwak et al. conducted a study that calculated the space and tracked parts, finished products and facilities in the manufacturing site through the simulation in order to smart SCM (Supply chain Management) solution by the distributed operating system based on NoSQL stored as big data [3]. Yu et al. identified the trends in intelligent robot technology in the logistics field and studied intelligent robot technology for logistics automation, and the results of this can be the basis for securing the reliability of robot development in the logistics field [4]. Choi et al. conducted overall robot research such as logistics robot operation technology and delivery robot technology, and identified key trends in robot technology related to the 4th industrial revolution [5]. Park et al. conducted an automatic classification logistics system study using image analysis. This study is a study using a model, but since control using a communication device was implemented, the meaning as a prototype was derived [6]. In order to implement an efficient logistics system, Won et al. conducted a study on a cargo automatic unloading system based on the YOLO (You Look Only Once) v2 model [7]. In order to analyze the factors affecting the intention to use logistics robots, Hwang et al. established a strategy to increase the utilization of logistics robots by applying a technology acceptance model [8]. Fig. 3 shows the procedure for the artificial intelligence algorithm-based courier invoice classification system to be presented in this study. Basically, an inspection of object data is performed. Next, through acquisition,

refine, and labeling steps, inspection is finally performed from an overall perspective. In this paper, the YOLOv4 algorithm based invoice classification to implement Joint courier logistics system was selected as the scope of research.

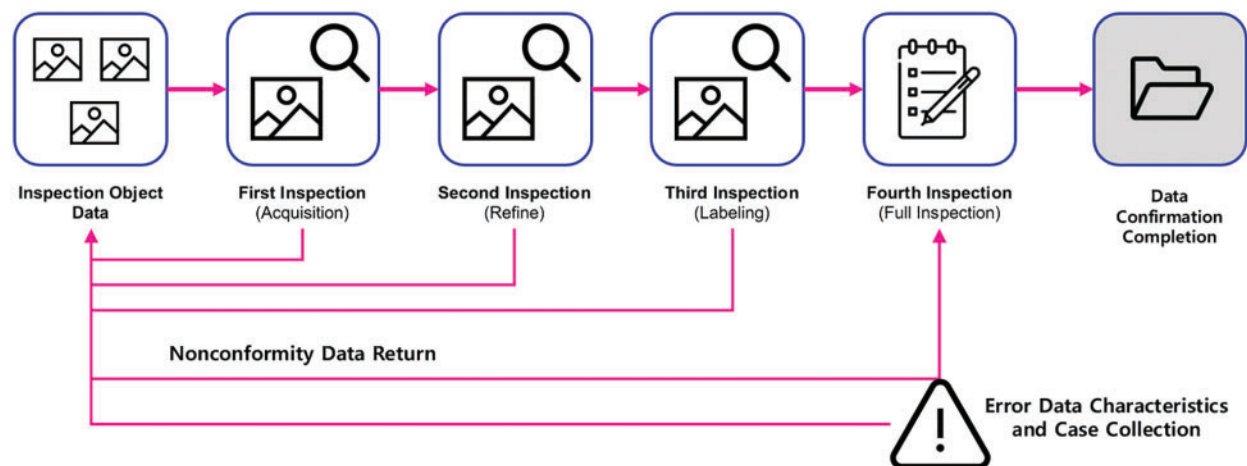


Figure 3: Schematic diagram of research methodology and procedures

1.2 Artificial Intelligence

Artificial intelligence technology imitating human intelligence is gradually developing. The Deep learning concept invented since the 1980s, but has received a lot of attention since it was published in 2006 by Hinton et al. [9]. Also, Rumelhart et al. proposed learning internal representations by error propagation, providing an opportunity to significantly develop the performance of artificial intelligence [10]. These days, the study in the field of the artificial intelligence which is related research on speech and image recognition proceeded since the paradigm of pattern recognition changed. CNN (Convolutional Neuron Networks) and RNNs (Recurrent Neural Networks) are deep learning models most generally used for diverse pattern recognition competition and commercial services. Gravesa et al. has been creating CNN by combining the backpropagation algorithm with Neocognitron. He makes field of image recognition has become popular by enlarging the number of pixels and converting it in various architectures of CNN algorithm. LSTM is the type of RNN. This methodology has been applied handwriting and speech recognition which in the field of time series analysis [11]. In spite of operating the algorithm successful, the period of neural network training was almost few days, and in general, there was an evaluation that it was not practical in other industry fields. Even if these circumstances, what is the reasons that deep learning used widely so far is as follows: Firstly, through various models, the disadvantages of the existing deep learning model have been overcome. Secondly, the development of hardware. In particular, the time required for complex deep learning tasks was greatly reduced through strong GPUs. Finally, the third is the use of big data that can be collected through the Internet and sensors. As such, after deep learning took the spotlight again, it achieved its best performance in various fields. In particular, in recent years, handwriting classification using MNIST data has been spotlighted in the field of computer vision. This provided an opportunity for non-experts to easily access deep learning. Like TIMIT, multiple test settings can be performed using a small amount of MNIST data. Accordingly, the importance of training for deep learning model implementation is emerging in the field of image recognition and object detection, these are main field

of computer vision. Deep learning was very effective for large-scale speech recognition at the time, making it suitable for large-scale tasks using large-scale deep CNN structures designed.

2 Methodology

2.1 Object Detection

Object detection which is major field of computer vision includes a series of processes for processing and analyzing videos to extract information contained in the data. Recently, it is used for various purposes, Park et al. proposed an object recognition method based on the depth information from the RGB-D camera for an indoor mobile robot. Such as perceiving object through distance and depth information and using these technology to indoor mobile robot [12]. In particular, there is a case where a study applying object recognition to a surveillance camera used in the security field has attracted attention [13]. The most common object detection method in vision algorithms is a method based on local image feature matching algorithm. Generally, local image feature matching process, an identifiable feature is selected to extract a feature vector from a local patch around the selected image feature point. One is the local scale unchanged image feature point detection method is SURF (Speeded Up Robust Features) is widely applied to derivative this feature vector [14,15]. As a study related to SURF method, Felzenszwalb et al. conducted an object detection study using a trained part-based model [16]. However, the problem with this method is that it is difficult to apply to the real world due to the slow calculation of complex feature vectors and the slow analysis process of matching valid object models to objects. For humans, it is a difficult process to design functions suitable for images object detection in all areas. Therefore, in terms of across the board performance and efficiency for the field of object detection, alternative means of deriving appropriate functions from the data itself are beginning to be required. Recently, fast development of artificial intelligence and deep learning technology has become a innovative point to solve almost problems related to object detection. Accordingly, the need for deep learning application is exploding in various fields, ranging from detection of defects in the factory automation system, precise robot work, to comprehensive convergence system based on object detection.

2.1.1 CNN (Convolutional Neuron Networks)

Existing object detection studies used methods such as SIFT (Scale-Invariant Feature Transform), HOG (Histogram of Oriented Gradient) based on low-level functions. This approach has limitations in improving performance, and since 2012, various attempts have been made to detect objects using CNN algorithms. Classification of object pattern in the region is performed through CNN and SVM (Support Vector Machine) classifiers for each compartment [17]. The mask R-CNN (Region based Convolutional Neural Networks) can be seen as algorithms for synthesizing RPN (Region Proposition Network) and Fast R-CNN, also, new mask branches have synthesize to Fast R-CNN classification and localization branches. A ROI (Region Of Interest), are samples within a data set identified for a particular purpose. In computer vision and optical character recognition, the ROI defines the borders of an object under consideration. Also, FPN (Feature Pyramid Network) is a extractor generated in order to highest that detection accuracy and speed by using pyramid shape. FPN which is feature extractor generates various feature map layers, in aspect of it is instead like a Faster R-CNN. Firstly, the mask R-CNN scales the image. Next then, Mask R-CNN scales the input size according to the backbone network. ResNet-101 creates a feature map for each hidden layer, and FPN creates an add-on map from the feature map derived already. Individually RPN is applied to the finally created function map to generate an output result. Remove all anchor boxes except for the anchor box with the highest

score among the anchor boxes created by Non-max suppression and adjust the size of the anchor box with different sizes by ROI alignment. Finally, when the anchor box value is delivered to the mask branch, the final processing is completed. In aspects of Chen et al. conducted a road damage detection and classification study applying density mask R-CNN [18]. Meng et al. studied that novel steganography in order to solve unknown information and to hide unknown information by applying based on instance segmentation in image area. That methodology shown to extract to information by receiver. Also, it is efficiently and safety better than existing algorithm [19].

2.1.2 YOLO (You Only Look Once)

Currently, object detection algorithm technique based on deep learning solves the most problem by estimating. Also, location area of the object and configuring a process of classifying the object have been applied a derived various CNN algorithm. The R-CNN algorithm first estimates an object candidate area in an image, classifies and detects an object bounding box based on this. Due to many expected candidate areas in this process, performance of computer vision utilization application has limit. For examples, robot remote control and actual security systems. Therefore, various studies have been attempted to increase the quantitative performance measurements while maintaining the speed of object detection. Recently, the proposed YOLO algorithm for object detection for object detection which is popular generally is maded for conduct both bounding box detection and alignment in the last output process. The last output the YOLO network terminal is a functional tensor representing all the possibility perception of the object class and bounding box. This tensor separate a grid of specific sizes from the input image, indicating the probability that the bounding box generated by each grid will become the bounding box of the object class. YOLO is faster method than the Fast R-CNN but it has lower accuracy. YOLO, a object detection algorithm based on deep learning, predicts one object per cell by dividing the predicted image into grid cells. The location and size of the object are determined by a predetermined number of boundary boxes, and since each cell has only one predictable object, the impact of overlapping objects can be reduced. Each bounding box consists of location (x, y), size (w, h) and box confidence score. The probability of the box includes object and the accuracy of the bounding box are quantitative performance measurements indicated by the box confidence score. Also, YOLO anticipate numerous bounding boxes for each grid cell at the same time. But, to calculate the actual loss, YOLO must select the box containing the most detected objects. To this end, calculate the ground truth and IOU. Next, select the one IOU which is the most highest. Eq. (1) represents the loss function of YOLO.

Loss function:

$$\begin{aligned} & \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\ & + \lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\ & + \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left(C_i - \hat{C}_i \right)^2 \\ & + \lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{I}_{ij}^{noobj} \left(C_i - \hat{C}_i \right)^2 \end{aligned}$$

$$+ \sum_{i=0}^{s^2} \mathbb{I}_i^{obj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \quad (1)$$

where, x , y , w and h represent correct data. \hat{x} , \hat{y} , \hat{w} , and \hat{h} represent predicted data. C represents the confidence when an object is present. And \mathbb{I}_i^{obj} denotes if object appears in cell i and \mathbb{I}_{ij}^{obj} denotes that the j bounding box predictor in cell i is responsible for that prediction. Fig. 4 shows the conceptual diagram of YOLOv4. This YOLOv4 can be expressed based on YOLOv3, the backbone part can be applied as CSPDarknet53, and the Neck part can be expressed by applying SPP (Spatial Pyramid Pooling) and PAN (Path Aggregation Network).

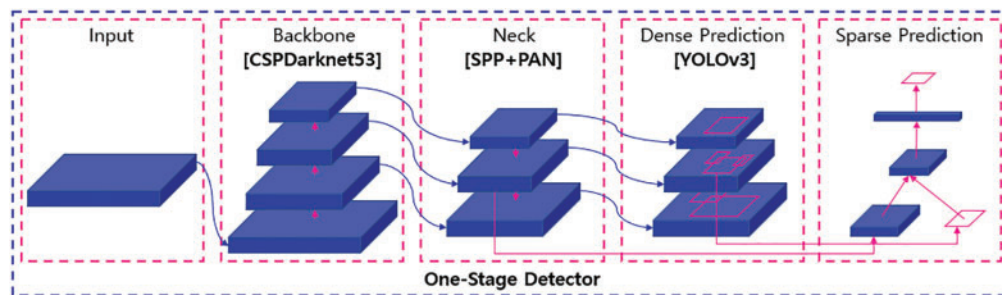


Figure 4: YOLOv4 based object detector

In a related study, Murthy et al. studied about pedestrian detection algorithm efficiently by applied YOLOv2 model. He applied the K-means clustering technique. In that way, the fuzzy fusion algorithm display more effective performance than legacy algorithms [20]. Amin et al. operate study of a Bi-LSTM-based human gait object detection by using sequences of video and derived 90% accuracy predictions by proposed algorithm application [21]. Amin et al. also conducted a three-dimensional semantics study of deep learning of leukemia networks and verified the accuracy, accuracy and sensitivity of fuzzy fusion algorithms using ALL-IDB1, ALL-IDB2, and LISC datasets [22]. Zhou et al. studied AT-YOLO deep mode-based helmet detection algorithms. As a result of the experiment, the mAP (Mean Average Precision) of the proposed algorithm was 96.5% and the detection rate was 27 fps, this measurements indicated that in aspects of detection accuracy and speed, it was superior than other conventional methods [23]. Baccouche et al. proposed two open mammograms, based on the YOLO model and the INbreast database to localize and classify simultaneously, suspected breast lesions in mammograph. As a result, the accuracy of mass and classified lesion detection in CBIS-DDSM, INbreast and personal datasets was 95.7%, 98.1%, 74.4%, 71.8%, and 73.2%, respectively [24]. Lee et al. proposed an improved two-stage license plate detection model using object division that can accurately detect rotating license plates due to the angle of inclination taken on the road [25]. Zhuang et al. conducted a YOLOv4 study to improve blurred gesture recognition. Consequently, it algorithm that based on YOLOv4 model derived highest effectiveness by measuring accuracy recognition effects in real time interaction of motion blur gestures batter than existing algorithm [26]. Matija et al. conducted a comparative study of the performance of two state-of-the-art convolutional neural network-based object detectors for ball detection under non-step real-world conditions, recognizing that many computer vision applications rely on accurate and fast object detection [27]. Soeleman et al. proposed a FCM (Fuzzy C Means) algorithm to generate adaptive thresholds for background subtraction in moving object detection. As a result, it was concluded that the MSE of FCM was lower than that of Otsu and the PSNR of FCM was higher than that of Otsu

[28]. Bochkovskiy et al. uses new features such as WRC, CSP, CmBN, SAT, Mish activation, Mosaic data augmentation, CmBN, DropBlock normalization, and CIOU loss, combining these features to achieve 43.5% APs (6.7%) of MS dataset real-time results [29]. Tan et al. many aspects of studied neural network architecture design selection in order that object detection and proposed several key optimizations to improve efficiency [30]. Fig. 5 shows a comparison of the performance of general object detection.

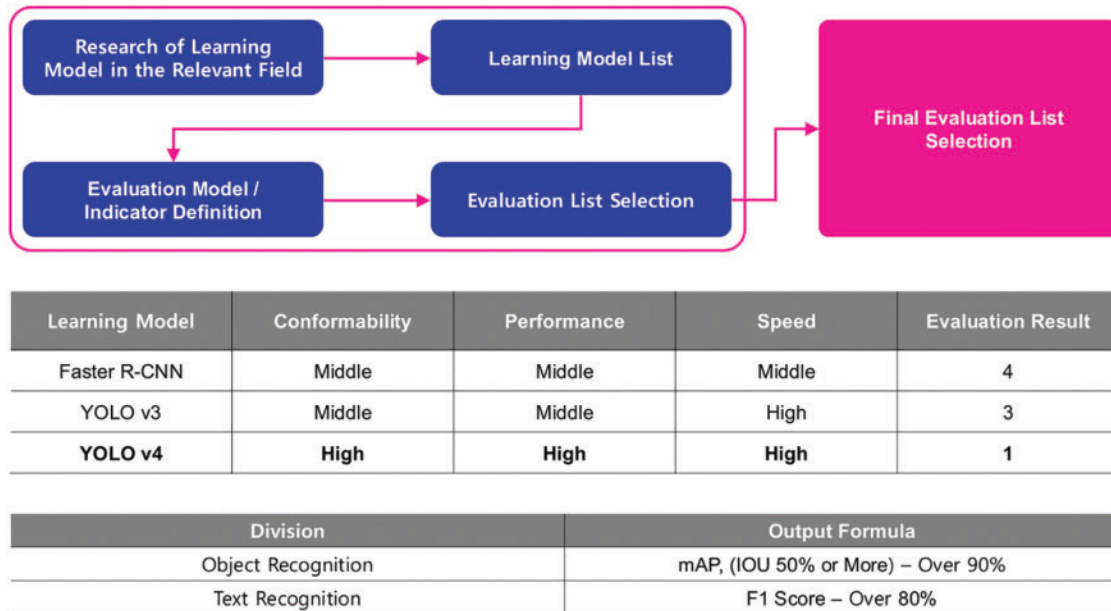


Figure 5: Comparison of the performance of object recognition algorithms

2.2 Joint Courier Logistics System

In order to solve the problems recognized in introduction, it is intended to present an artificial intelligence-based courier invoice algorithm procedure necessary for the operation of the point courier logistics system. The joint courier logistics system can be explained by the procedure shown in Fig. 6. First, the cargo departs from the manufacturing plant and is delivered to the final consumer through a logistics center, a joint fulfillment and herb logistics center, a sub-distribution center, and a retail stage. In this case, the manufacturing plant handles delivery invoice information such as destination, shipper, departure, etc., and cargo information such as expiration date and handling method. However, since each carrier and manufacturer handles different delivery invoices, information disconnection to the logistics center occurs, which is the next step. Four inspection are required to efficiently implement the joint courier logistics System. First, inspection object data is performed, and first inspection according to data acquisition is performed. Next, second inspection according to refine and third inspection according to labeling are performed. Finally, full inspection is performed to organize data completion as a result. By integrating the courier cargo invoice classification system operated separately for each courier company and operating the joint courier logistics System based on this, an efficient logistics system can be implemented in the city center. In this paper, to implement an efficient joint courier logistics system, we propose an artificial intelligence-based courier invoice collection, recognition, and classification system to systematically analyze courier cargo invoice classification procedures from the entire system to subsystem or components.

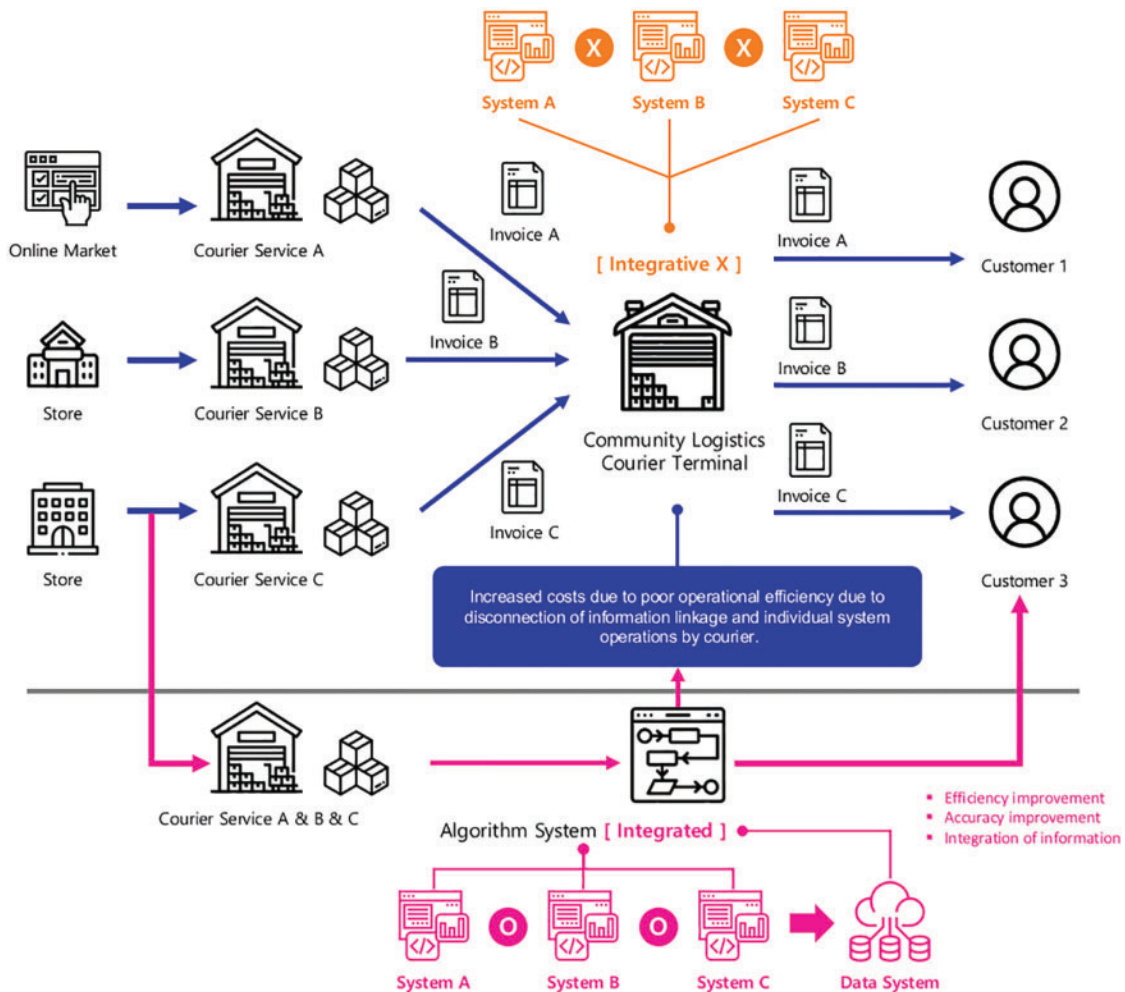


Figure 6: The problem of loss of logistics systems due to lack of data

3 Acquisition and Classification Standards for Invoice Data

3.1 Data Acquisition

This study aims to propose a delivery invoice information recognition system necessary to present standards for data collection to implement this system. Fig. 7 shows an example of a parcel delivery invoice used in the parcel delivery invoice system for each parcel delivery company. As highlighted, it can be seen that the courier cargo invoices used by each courier are different. This problem may cause information delivery problems in implementing the joint logistics delivery terminal system. This study proposes an artificial intelligence-based courier cargo invoice information recognition system for efficient implementation of joint logistics courier terminals, and prior to implementing the system, general courier cargo invoice information was classified. The target of collection is transport invoice information, and it is directly collected by presenting sensors necessary for image recognition. Thereafter, after sampling the acquired courier cargo invoice information, purification, labeling, and quality inspection operations are presented. In this study, in order to implement the joint logistics courier system, the target courier invoice information required for the courier invoice information

recognition system was set to more than 300 video data, more than 100,000 images, and text labeling data. Videos used in the registration and inspection of delivery invoice are not images obtained throughout the parcel invoice, but images of delivery invoice to be inspected are separately cut and used. The courier invoices used by each courier will be attached horizontally or vertically to the cargo. All videos of the courier invoice used in the learning mode or inspection mode rotate so that the horizontal length of the transport information is longer than the vertical length, otherwise the part with a small lead of the courier invoice rotates upward to change the form of the courier invoice the same.

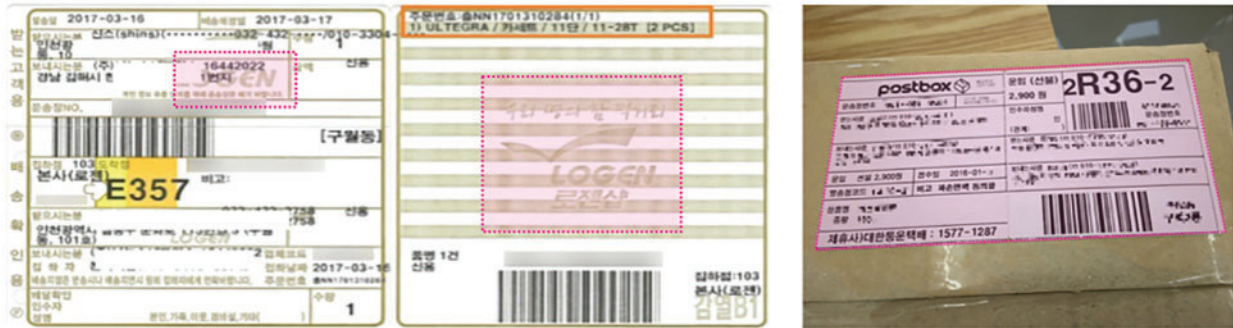


Figure 7: Delivery information for each company

3.2 Classification Standards

In order to implement the delivery invoice information recognition system, it is necessary to define the acquisition and classification procedure for data integration. In order to construct a systematic data acquisition and classification procedure, this paper planned a data sensitive information and quality processing plan. Data sensitive information and quality processing methods consist of four categories: refinement, redundancy review, de-identification of sensitive information, and image data division. The purification process performs primary purification in consideration of photographed video and image review and image quality and object identification. The redundancy review process discards data when overlapping factors occur in consideration of photographing time, photographing distance, photographing place, and photographing height for the same object. For the libraryization of parcel invoices, the types of data or characteristics between each data are analyzed. There are five types of courier invoices used in this paper: Company C, Company H, and Company L, and the types of delivery invoice are subdivided according to the size of courier invoices in each type. Features such as shape, size, and font are features that can distinguish the types of each parcel invoice when using a deep learning algorithm, so ROI areas are determined by analyzing the characteristics of paper colors and characters for each parcel invoice type. The biggest feature is that characters exist in each parcel invoice, and the characters are long vertically. Class of parcel invoices is classified by dividing them based on the type, width, and length of parcel invoices, and the class is named type × width × length. There are a total of five parcel invoices, as shown in Table 1, and CJD is classified into 3 types, HJD is 1 type, LGD is 2 types, LTD is 2 types, and POD is 1 type. Even if the shape or color is similar, if the size of the parcel invoice is different, it is a different class. When defining a class, the unit that measures the size of the parcel invoice is cm, and the class is classified using the type, horizontal length, and vertical length. In this way, parcel invoices for each courier company were classified only by type, and divided by type and size and classified secondarily. Classification by parcel invoice can be classified into different parts even if they are the same size or different types when classifying by size, so this

paper conducts research with the aim of intermediate classification. When classifying parcel invoices by size, parcel invoices of similar size can be classified into one class, and even parts of the same class often have different shapes, sizes, and colors. The goal of this study is to classify parts well even if there are such deviations, but classification errors due to contamination due to transportation environment can be a problem, so they were excluded. The class type is shown in [Table 1](#), and the class name is named by the horizontal and vertical length of the parcel invoice, and the class name is named by attaching the type of parcel invoice after the size of the parcel invoice.

Table 1: Selection of the class name of the parcel invoice according to the type and size

Classification	Size [cm]	Class name	Count
CJD	10 × 13	1013-CJD	610
	10 × 18	1018-CJD	613
	10 × 19	1019-CJD	621
HJD	10 × 20	1020-HJD	598
LGD	10 × 17	1017-LGD	589
	10 × 20	1020-LGD	613
LTD	10 × 13	1013-LTD	600
	10 × 20	1020-LTD	618
POD	11 × 17	1117-POD	619
Total experimental data			5,481

The process of de-identifying sensitive information performs blur or mosaic processing when sensitive information that violates personal information and security is exposed. Finally, an image data segmentation process is performed. The criteria for purification consist of four types: clarity, prevention of redundancy, purification work manual, and purification tool. The clarity of the refinement criteria evaluates whether a refinement standard suitable for the purpose of data use has been established, and the prevention of redundancy evaluates whether redundancy has been confirmed after comparing data information. The refinement work manual evaluates whether a manual for refinement work has been created, and the refinement tool evaluates whether it is familiar with how to secure and use software tools to be used for refinement work. There are seven categories of inspection items, 1. Check whether necessary or meaningless data is removed, 2. Check whether noise-generating data is removed, 3. Check repeated data and folder/file overlap, 4. Check appropriate frame division criteria. 5. Check sensor data analysis and abnormal data deidentification. Refinement criteria for delivery invoice information and sensitive information including inspection items can contribute as an evaluation and performance manual when the actual system is operated. [Fig. 8](#) shows the contents of this. The data labeling process can be judged as a key element in implementing a successful object detection algorithm. Therefore, as an element necessary for the procedure for recognizing the delivery invoice information for implementing the joint logistics delivery terminal system, the definition of the procedure should be systematically performed.

[Fig. 9](#) shows the procedure for labeling transport information for the delivery invoice information recognition system. Next, a label suitable for delivery invoice information is added, and label information corresponding thereto is designated. After recognizing the invoice and forming the bounding box, a process of reviewing whether the position of the bounding box is appropriate is performed.

Next, the position of the bounding box is readjusted and modified, and the processing data is finally reviewed. Finally, the labeled transport information is stored. Details of the labeling operation can be composed of a total of four things: common, annotation, inspection, and administrator. Data labeling was performed by applying the YOLO category in the “labelImg” program and inputting information corresponding to image data. The initial work of labeling the contents in the invoice was done manually. In the case of simple tasks such as copying labeling contents that have been carried out since then, it was automatically carried out through an automated labeling program.

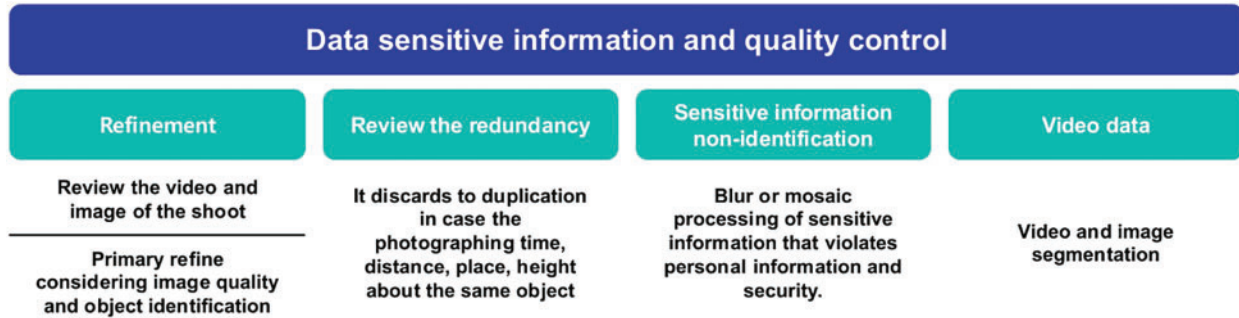


Figure 8: Data sensitive information and quality control

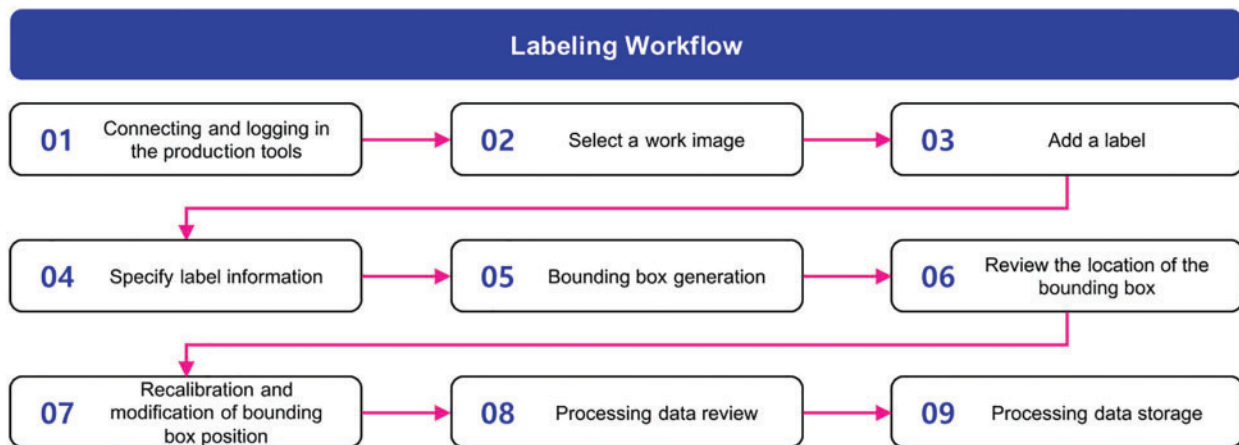


Figure 9: Labeling workflow

Fig. 10 shows the work of labeling an invoice image file. Information on inspection is determined by providing the assigned image list annotation results for each worker, providing key meta information for each image, providing detailed results for each annotation object, and passing and failure for the annotation operation results. Details of the work from the manager’s perspective include designation of worker information inquiry/registration/modification, image registration, and detailed status inquiry, allocation/inspection work, and monitoring progress. Even if the implemented algorithm shows high performance results with an efficient configuration, a successful object detection system cannot be implemented without solving the problem with the data itself. The problem with this is the most emerging problem at the time when artificial intelligence is currently being combined and studied in various fields. Implementing and operating a successful system requires efficient architecture configuration by eliminating systematic and unnecessary procedures from the data collection stage to

the classification stage. Therefore, for successful implementation of an object detection system in the logistics field, definition of criteria for obtaining and classifying courier cargo transport data may act as an important factor.

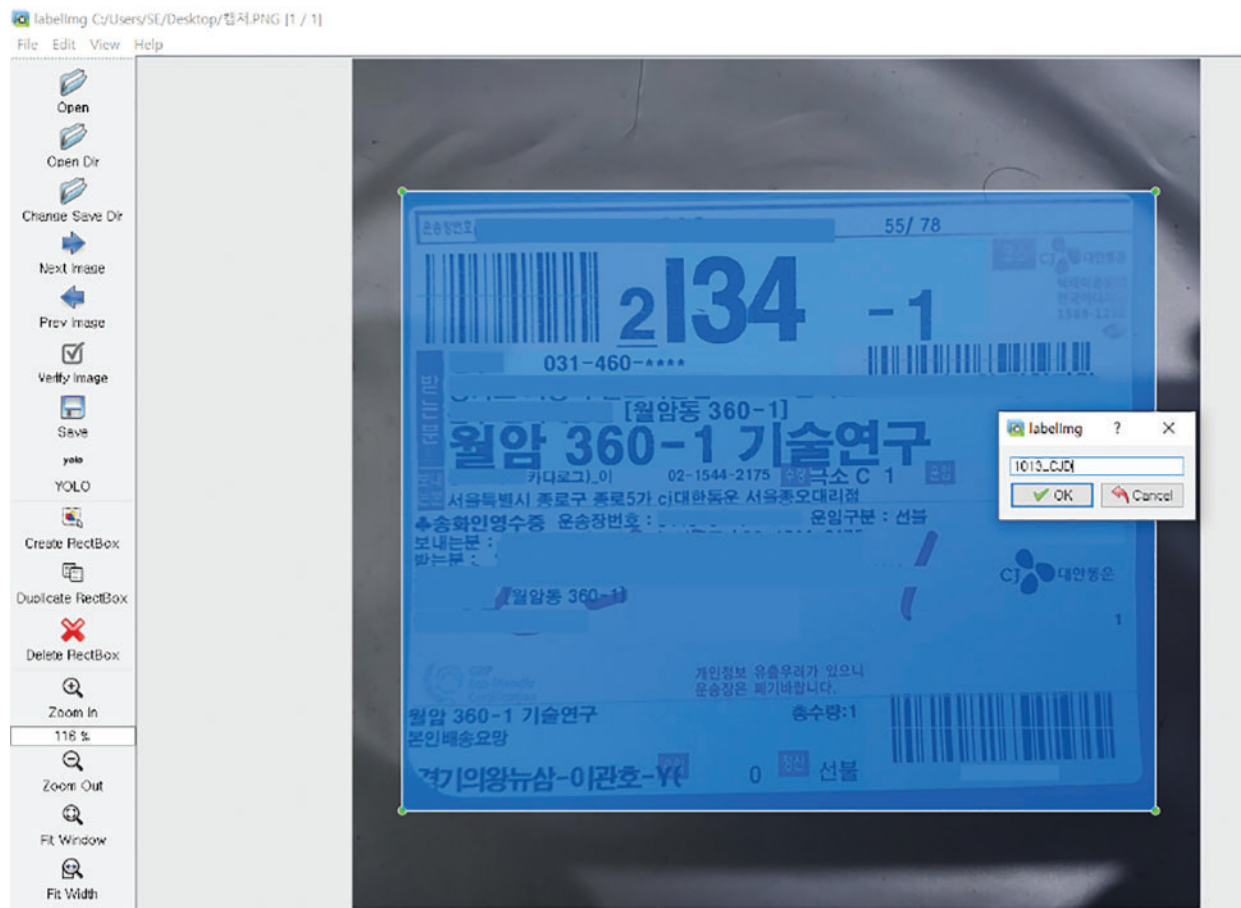


Figure 10: The process of labeling invoice image data

4 Algorithm Verification

In this study, general standards were established through the integration of delivery invoice, data acquisition, and definition of labeling procedures for each courier for the implementation of the joint courier logistics system. However, additional procedures are needed to determine whether these criteria are accurately met. Therefore, this study proposes a 1-cycle-centered inspection for efficient system operation. From the perspective of the entire system, meta-items and processing information inspection and feedback through checklist establishment are key factors, and this process includes four items: acquisition, purification, processing, verification, and structural inspection. Although more precise data inspection can be performed by detailing the inspection process, time and cost economic feasibility must be considered for efficient system operation. Each of the transport information data inspection checklist items consisting of transport information acquisition and purification, processing, verification, and structural inspection includes corresponding contents. First, the acquisition and

purification process includes file name generation rules, file forms, file validity, meta-items, and meta-validity, and the processing process includes data omission, meta-completeness, compliance with transcription principles, and meta-validity. Next, the verification stage includes regular verification stage 1, 2, and 3 inspections, and additionally includes third-party inspections by external advisory agencies. The structural inspection process includes structural inspection, formal inspection, and statistical inspection. Fig. 11 shows the checklist items presented in this study and the delivery invoice information quality inspection process. The implementation of artificial intelligence-based delivery invoice information recognition algorithms for joint courier logistics system is also important in the process of successfully implementing only the algorithm learning model itself, but data used as a material for algorithms is also a key factor. This requires data collection, processing, and classification processes. However, without any criteria for data handling, the process of processing data can be difficult to implement a successful algorithm. Therefore, if an object detection algorithm system in the logistics field is implemented in consideration of the criteria for obtaining, processing, and classification of courier cargo transport information data presented in this paper, it is judged that more systematic and efficient work progress is possible. In general, the process of implementing an object detection algorithm consists of data acquisition, data purification, data labeling, data quality inspection, and application of a learning model, and the courier transport information data classification algorithm presented in this paper is also configured like this. In the process of each step, the criteria for data processing methods are key elements in the implementation of successful algorithms, and in this paper, systematic criteria are presented.

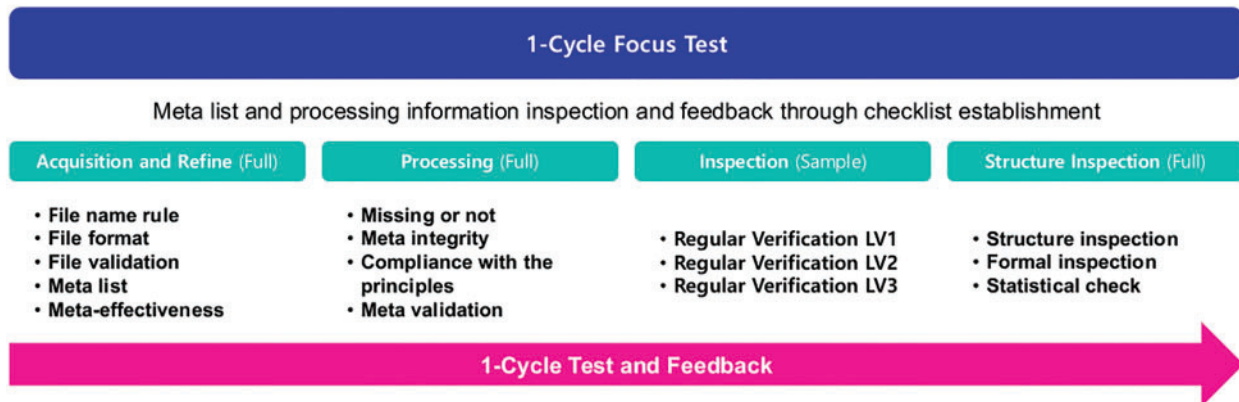


Figure 11: Checklist generation and data inspection procedures

It is judged that efficient algorithm implementation is possible by applying these data processing criteria. In the future, we would like to present two ways to apply the learning model to be carried out based on the delivery invoice information processing standards presented in this paper. The first is the development of an expiration date detection model through image data analysis. Due to the separate expiration date notation method used by many logistics and courier companies and manufacturers, there is a problem of recognizing expiration dates that cannot be solved with general barcode or OCR (Optical Character Recreation) equipment. To solve this problem, we intend to use the results of this paper to develop an expiration date recognition model through image data analysis. The second is the development of an atypical packaging logistics delivery invoice information object detection model of this paper. Recognizing the transport information of unstructured courier cargoes such as small packaging and individual packaging, not formal courier box packaging, is currently very difficult, and to solve this problem, it is necessary to obtain data on unstructured courier cargo delivery invoice

information. In the future, the results of this study may contribute to establishing a logical criterion in the processing process for obtaining and classifying courier cargo transport information data. The YOLOv4 algorithm was implemented in a environment with CPU Intel i7-11700k, GPU NVIDIA RTX-3060, and RAM 32GB (16GB + 16GB). It was divided into 8:2 learning data and experimental data, and because there are several invoice information included in each class and the background is diverse, 900 sheets of learning data are included. The learning data consisted of 14,519 chapters and the experimental data consisted of 5,481 chapters. Fig. 12 shows the results of invoice image recognition using the YOLOv4 algorithm. Table 2 is the result of comparing the YOLOv4 algorithm used in this study with the existing algorithms, Fast R-CNN and YOLOv3. For Fast R-CNN, FPS Average was 26, detection rate was 96.48%, FPS Average was 45, and detection rate was 88.44%. In contrast, the FPS Average of YOLOv4 was 51 and the detection rate was 95.36%. Even if classes with less data are included in learning, if the deviation of learning data per class is severe, classes with more data will be overfitting, and classes with less data will not be properly learned, so classes with less data will be excluded from classification classes. Since the basic anchor box does not contain large barcodes, the low detection accuracy for small texts is a disadvantage of CNN-based object detection models. Table 3 shows the detailed composition of each class of learning data. Finally, as shown in Table 3, the classification performance by delivery invoice information is 95.36%, showing very good results. As a result of applying the YOLOv4 algorithm, the performance of 51 FPS was shown. Also, When analyzing the results of these contrast experiences, the algorithm optimized to date was found to be YOLOv4.



Figure 12: YOLOv4-based invoice image detection

Table 2: Comparison of the existing Fast R-CNN, YOLOv3 and YOLOv4 algorithm

Algorithm	FPS average	Detection rate
Faster R-CNN	26	96.48
YOLOv3	45	88.44
YOLOv4	51	95.36

Table 3: Experiment result of network modification model (YOLO v4) for delivery invoice information

Class	Count	Pass	Fail	Accuracy
1013-CJD	600	574	26	95.72%
1018-CJD	618	593	25	95.88%
1019-CJD	589	561	28	95.25%
1020-HJD	613	579	34	94.45%
1017-LGD	598	571	27	95.56%
1020-LGD	610	580	30	95.08%
1013-LTD	613	586	27	95.57%
1020-LTD	621	593	28	95.42%
1117-POD	619	590	29	95.26%
Total	5,481	5,226	255	95.36%

5 Conclusion

Due to the growth of the online logistics market due to the development of Internet technology, the demand for logistics and courier cargo among general consumers is rapidly increasing. Accordingly, in the case of urban areas, road traffic congestion and environmental problems caused by cargo vehicles are mainly occurring. As a way to solve this problem, a joint courier logistics system was proposed. The joint logistics delivery terminal aims to implement an efficient logistics transportation system by utilizing one joint courier logistics system by multiple logistics and delivery companies. However, these systems have problems due to information delivery disconnection because the delivery invoice information processing systems used by many logistics and courier companies are different. Therefore, in this paper, an artificial intelligence based delivery invoice information recognition algorithm for implementing the joint courier logistics system is presented. In addition, in order to successfully implement the algorithm, a series of criteria are required as key elements in the process of acquiring and classifying transport information data. In this paper, only five types of large courier companies were selected as research subjects, but further studies will deal with more types of courier companies. In addition, for further research, additional cloud and DB construction studies that can efficiently collect and process large amounts of delivery invoice information will be conducted. Finally, this paper systematically analyzed the data processing process required for algorithm implementation, devised a practically feasible methodology, and established delivery invoice information processing standards for this. The delivery invoice information processing standards presented as research results can be used as base data for implementing the joint courier logistics system in the future. And it can be used as a basis for securing the operational reliability of the joint courier logistics system. From a macroscopic point of view of the logistics industry, the results of this study can efficiently improve the logistics process by eliminating social costs such as delivery delays and misdelivery.

Funding Statement: This work was supported by a grant from R&D program of the Korea Evaluation Institute of Industrial Technology (20015047).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] M. N. Chae and H. K. Gwon, "Business models and policy directions of joint logistics projects funded by government," *Korea Society of Transportation*, vol. 13, no. 3, pp. 28–35, 2016.
- [2] H. J. Shin, "A study on trends in the use of logistics technology based on the 4th industrial revolution," *The e-Business Studies*, vol. 21, no. 2, pp. 17–27, 2020.
- [3] K. J. Kwak, S. G. Hwang, D. J. Shin, K. W. Park, J. J. Kim *et al.*, "Study of logistics object tracking service for smart SCM," *Journal of the Korean Institute of Industrial Engineers*, vol. 46, no. 1, pp. 71–81, 2020.
- [4] W. P. Yu, Y. C. Lee and D. H. Kim "Technical trends of robot task intelligence in intelligent logistics/agriculture," *Electronics and Telecommunications Trend*, vol. 36, no. 2, pp. 22–31, 2021.
- [5] S. I. Choi, D. H. Kim, J. Y. Lee, S. H. Park, B. S. Seo *et al.*, "Logistics and delivery robots in the 4th industrial revolution," *Electronics and Telecommunications Trends*, vol. 34, no. 4, pp. 98–107, 2019.
- [6] C. H. Park, S. D. Bae, S. G. Choi, S. H. Choi, J. W. Choi *et al.*, "Automatic picking/classification system using video analysis," in *Proc. of the Korean Society of Computer Information Conf.*, Jeju, Korea, vol. 28, no. 2, pp. 661–662, 2020.
- [7] J. U. Won, M. H. Park, S. W. Park, J. H. Cho and Y. T. Kim, "Deep learning based cargo recognition algorithm for automatic cargo unloading system," *Journal of Korean Institute of Intelligent Systems*, vol. 29, no. 6, pp. 430–436, 2019.
- [8] H. C. Hwang and S. H. Song, "A study on the factors affecting the acceptance of logistics robot in the fulfillment center using the technology acceptance model," *Journal of the Korea Academia-Industrial Cooperation Society*, vol. 20, no. 12, pp. 287–297, 2019.
- [9] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [10] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning internal representations by error propagation," in D. E. Rumelhart, J. L. McClelland and The PDP Research Group (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge, MA, USA: MIT Press, vol. 1, pp. 318–362, 1986.
- [11] A. Gravesa and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, 2005.
- [12] J. K. Park and J. B. Park, "An object recognition method based on depth information for an indoor mobile robot," *Journal of Institute of Control, Robotics and Systems*, vol. 21, no. 10, pp. 958–964, 2015.
- [13] S. I. Park and M. Y. Kim, "Multiple-background model-based object detection for fixed-embedded surveillance system," *Journal of Institute of Control, Robotics and Systems*, vol. 21, no. 11, pp. 989–995, 2015.
- [14] H. Bay, T. Tuytelaars and L. V. Gool, "Surf: Speeded up robust features," in *Computer Vision–European Conf. on Computer Vision*, Graz, Austria, pp. 404–417, 2006.
- [15] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. of the 7th IEEE Int. Conf.*, Kerkyra, Greece, vol. 2, pp. 1150–1157, 1999.
- [16] P. F. Felzenszwalb, R. B. Girshick, D. McAllester and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [17] R. Girshick, J. Donahue, T. Darrell and J. Malik Rich, "Feature hierarchies for accurate object detection and semantic segmentation," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 580–587, 2014.
- [18] Q. Chen, X. Gan, W. Huang, J. Feng and H. Shim, "Road damage detection and classification using mask R-CNN with DenseNet backbone," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 2201–2215, 2020.
- [19] R. Meng, Q. Cui, Z. Zhou, C. Yuan and X. Sun "A novel steganography algorithm based on instance segmentation," *Computers, Materials & Continua*, vol. 63, no. 1, pp. 183–196, 2020.
- [20] C. B. Murthy, M. F. Hashmi, G. Muhammad and S. A. AlQahtani, "YOLOv2PD: An efficient pedestrian detection algorithm using improved YOLOv2 model," *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3015–3031, 2021.

- [21] J. Amin, M. A. Anjum, M. Sharif, S. Kadry, Y. Nam *et al.*, “Convolutional bi-LSTM based human gait recognition using video sequences,” *Computers, Materials & Continua*, vol. 68, no. 2, pp. 2693–2709, 2021.
- [22] J. Amin, M. Sharif, M. A. Anjum, A. Siddiqa, S. Kadry *et al.*, “3D semantic deep learning networks for leukemia detection,” *Computers, Materials & Continua*, vol. 69, no. 1, pp. 785–799, 2021.
- [23] Q. Zhou, J. Qin, X. Xiang, Y. Tan and N. N. Xiong, “Algorithm of helmet wearing detection based on AT-YOLO deep mode,” *Computers, Materials & Continua*, vol. 69, no. 1, pp. 161–174, 2021.
- [24] A. Baccouche, B. Garcia-Zapirain, C. C. Olea and A. S. Elmaghraby, “Breast lesions detection and classification via YOLO-based fusion models,” *Computers, Materials & Continua*, vol. 69, no. 1, pp. 1407–1425, 2021.
- [25] S. W. Lee, B. S. Choi and Y. S. Kim, “An enhanced two-stage vehicle license plate detection scheme using object segmentation for declined license plate detections,” *Journal of the Korea Society of Computer and Information*, vol. 26, pp. 49–55, 2021.
- [26] H. Zhuang, Y. Xia, N. Wang and L. Dong, “High inclusiveness and accuracy motion blur real-time gesture recognition based on YOLOv4 model combined attention mechanism and deblurganv2,” *Applied Sciences*, vol. 11, no. 21, pp. 1–19, 2021.
- [27] B. Matija, P. Miran and M. I. Kos, “Ball detection using YOLO and mask R-CNN,” in *Conf.: 2018 Int. Conf. on Computational Science and Computational Intelligence*, Las Vegas, NV, USA, pp. 319–323, 2018.
- [28] M. A. Soeleman, M. Hariadi and M. H. Purnomo, “Adaptive threshold for background subtraction in moving object detection using fuzzy C means clustering,” in *Proc. of TENCON IEEE Region 10th. Conf.*, Cebu, Philippines, 2012.
- [29] A. Bochkovskiy, C. Y. Wang and H. Y. M. Liao, “YOLOv4: Optimal speed and accuracy of object detection,” *arXiv, Computer Vision and Pattern Recognition*, arXiv:2004.10934, 2020.
- [30] M. Tan, R. Pang and Q. V. Le, “EfficientDet: Scalable and efficient object detection,” *arXiv, Computer Vision and Pattern Recognition*, arXiv:1911.09070, 2020.