

Article

Parcel Classification and Positioning of Intelligent Parcel Storage System Based on YOLOv5

Mirye Kim and Youngmin Kim *

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

* Correspondence: pretty0m@ajou.ac.kr

Abstract: Parcel storage provides last-mile delivery services as part of the logistics process. In order to build an intelligent system for parcel storage, we conducted a study on parcel box recognition using AI's deep learning technology. Box detection and location estimation studies were conducted using the YOLOv5 model for parcel recognition, and this model is presently being applied to many studies because it has excellent object recognition and is faster than previous models. The YOLOv5 model is classified into small, medium, large, and xlarge according to the size and performance of the model. In this study, these four models were compared and analyzed to perform an experiment showing the optimal parcel box recognition performance. As a result of the experiment, it was determined that the precision, recall, and F1 of the YOLOv5large model were 0.966, 0.899, and 0.932, respectively, showing a stronger performance than the other models. Additionally, the size of the YOLOv5large is half that of YOLOv5xlarge, and the YOLOv5large showed the optimal performance in object recognition of the parcel box. Therefore, it seems that the basis for building an intelligent parcel storage system, which shows optimal efficiency in real time using the YOLOv5large model, can be laid through the parcel object recognition experiment conducted in this study.

Keywords: parcel storage system; e-commerce logistics; object detection; YOLOv5; box detection



Citation: Kim, M.; Kim, Y. Parcel Classification and Positioning of Intelligent Parcel Storage System Based on YOLOv5. *Appl. Sci.* **2023**, *13*, 437. <https://doi.org/10.3390/app13010437>

Academic Editors: Emanuele Carpanzano, Bhanu Shrestha, Seongsoo Cho and Changho Seo

Received: 8 December 2022

Revised: 22 December 2022

Accepted: 26 December 2022

Published: 29 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In 2021, global e-commerce sales reached \$5945 trillion, up 19.5% from 2020. Among them, Korea ranked 5th in the worldwide e-commerce market share, showing the highest ratio of e-commerce to area ratio, and even today, the logistics industry is continuously increasing. In fact, in 2021, Korea's parcel delivery volume reached 3.62 billion, recording an all-time high, and the number of parcel deliveries per economically active population reached 128.4 per year. Alongside the growth of logistics, the latest logistics is growing in a direction that can promote optimization, efficiency, and cost reduction of the sector, reflecting the current society. This is also called an intelligent logistics system, and this logistics system is used in production, warehouse, service, and all other logistics fields, so that machines can efficiently help human convenience [1]. As the number of parcels increases, the application of intelligent logistics systems in the courier process becomes more critical since the labor force is limited. In the current logistics industry, due to the explosive increase in parcel delivery volume, complaints from consumers are increasing due to the loss and damage of parcels in the delivery process to the final consumer, which is the final stage of the parcel delivery process, and disputes between courier companies are occurring. In addition, as the problem of impersonating a courier driver has occurred over the past five years, consumers are exposed to crimes related to receiving parcels.

Therefore, various parcel storage facilities for storing the goods appear to be a countermeasure in order to solve the problem of exposing consumers to crime, due to the damage of parcels and impersonation of courier drivers [2]. Parcel storage is a safety measure between the delivery men and the consumers, and the delivery men can save themselves the cost of compensation for the loss or damage of the parcels by safely storing the goods,

and the consumer can avoid receiving damaged parcels and being exposed to the crime that may occur while receiving the goods [3]. The parcel storage facilities currently in operation and under development are not simple storage facilities, but are evolving into intelligent parcel storage facilities with artificial intelligence technology applied [4]. The reason why intelligent technologies are stored in parcel storage facilities is to facilitate the operation, reduce human resource consumption, and digitize and manage the operating process. Currently, unmanned parcel storage is installed and operated in countries worldwide. Examples include: CTT in Portugal (2020), Correos in Spain (2017), La Poste in France (2016), AustraliaPost in Australia (2021), and Hong Kong Post in Hongkong (2017), etc [5]. In Korea, delivery storage is operated in various places such as apartments, subway stations, and companies, and all existing parcel storage is attached to the wall side so that one parcel can be loaded into one storage box. Currently, this form of parcel storage utilization is labor-intensive and time-consuming for delivery drivers, and it is difficult to store many parcels in the space. Therefore, in this study, we intend to propose a system that can reduce human resource consumption and the labor of delivery drivers, and maximize the space available for parcel storage.

This study aims to implement a delivery storage system that can efficiently store parcels using AI technology. In order to efficiently manage delivery loading and storage in an unmanned delivery storage system, an object recognition experiment of the delivery box was conducted using the CNN deep learning model, a piece of AI technology. The YOLOv5 algorithm was used as a model in this research, which is the fifth version of the YOLO algorithm widely used in the field of object recognition and is drawing attention in terms of speed and using the latest techniques. The YOLOv5 algorithm is divided into four versions, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, depending on the width and depth of the model, and in this study, object recognition comparison analysis was conducted based on these four versions of the YOLOv5 algorithm.

This paper is structured as follows: Section 2 gives a literature review and a general system design study of the parcel storage system. Section 3 presents a study on deep learning-based object detection introduced into the parcel storage system, and algorithm selection. Section 4 discusses the experimental result of the object recognition accuracy and efficiency of parcel storage through actual object detection experiments. In addition, we proposed an algorithm that can be introduced into the parcel storage system by comparing results between models. Section 5 concludes the findings in relation to the contribution of the presented work.

2. Related Work

2.1. Literature Review

As personal and non-face-to-face life became commonplace after COVID-19, the courier logistics industry is also activating untact and e-commerce delivery. At the same time, AI (Artificial Intelligence) technology began to be applied to various fields including the manufacturing field, thus the logistics industry, which relied on manpower, is changing. The development of unmanned parcel storage that connects delivery drivers and consumers in the final stage of the delivery process, and the application of AI technology are increasing. However, while the speed at which AI technology is being applied to the existing unmanned delivery system is fast, there is a lack of research on the application of AI technology to the unmanned parcel storage system. Instead, research on parcel recognition, not parcel storage, has mainly been conducted to identify parcels through technologies such as GPS and RFID in the transportation and storage process of parcels [5,6]. This technology helps to track the delivery status and route of parcels, but many related studies have already been conducted since 2013, thus this research is now common, limiting further research because it is already widely used in the logistics industry.

Although few studies apply AI technology to parcel recognition, research is gradually increasing. In the study of V. Naga Bushanam and Ch.S atyananda Reddy, an image recognition experiment that can be applied to the logistics industry was conducted using

AI's deep learning technology. In the study, image classification was performed using EfficientNet, and the learning accuracy achieved was 83% [7]. Kim M. et al. conducted a study applying AI technology to classify parcels within the urban railway system. As a performance method, experiments were conducted by comparing several deep learning models such as MobileNet, ResNet50, and VGG16 [8]. The research conducted by Kim M. et al. was conducted before this research, and was conducted to classify both normal and defective parcel boxes, and the experiment focused on object classification rather than object recognition and location estimation. A study was also conducted to distinguish between breakage and non-damage in parcel recognition. An image recognition study was conducted to increase the recognition rate of parcels and speed by transforming the existing AI deep learning model. As a result of the study, the recognition accuracy of the delivery box was 86.75% and proved itself to be effective [9].

As shown in Table 1, looking at the studies conducted so far, GPS and RFID-related studies have been mainly conducted in the past, and now, based on these studies, most of them are applied to and operated in actual systems. Presently, studies on the recognition and classification of delivery boxes are mainly conducted, but the recognition rate remains at 80%, which performance is not particularly high.

Table 1. Comparison of related works.

Related Work	Research Classification	Method	Accuracy
[1]	Parcel tracking	GPS	-
[5]	Parcel tracking	RFID	-
[6]	Parcel tracking	Android 4.4.2	-
[7]	Parcel Classification	EfficientNet	83%
[8]	Parcel Classification	MobileNet, ResNet50, VGG16	84.6%
[9]	Parcel Classification	Improved MobileNet-v2	86.75%

Therefore, in this study, we propose an unmanned parcel system and conduct a study on an algorithm that can improve parcel recognition performance by applying suitable AI technology. Unlike the existing literature, this study intends to use object recognition technology to build a robot loading system for unmanned delivery systems in the future, rather than deep learning classification technology for box recognition.

2.2. Overall Design of the Parcel Storage System

The role of unmanned parcel storage is a system that serves as a connective element and safety measure between the delivery driver and the final consumer. The purpose of the delivery driver is to safely deliver the goods to the end consumer through parcel storage and to allow the final consumer to receive the stored goods without losing the goods or being exposed to other risk factors. Therefore, accurate parcel information recognition and safe storage are considered essential when designing a parcel storage system.

The main design of the system can be divided into a display unit (object recognition unit), a control unit (central control + backup unit), a robot unit, and a storage unit. To acquire continuous frame images from the storage entrance and display part (object recognition part) and recognize objects with a high rate of accuracy, it is essential to develop a camera and apply related AI technology. To detect an object, it must be designed as an appropriate AI deep learning model to balance detection speed and accuracy [10]. The control part stores the information received from the storage entrance and display part in the central control and backup, and retransmits it to the storage entrance and display part. Based on the object information transmitted from the control part, the mobile robot, receiving the command from the central control part, uses the gripper to move the parcel and load it into the storage space. The design of the proposed parcel storage system should consider the information exchange

and linkage between modules, data storage, real-time box recognition, and robot transfer technology. A diagram of the system is shown in Figure 1.

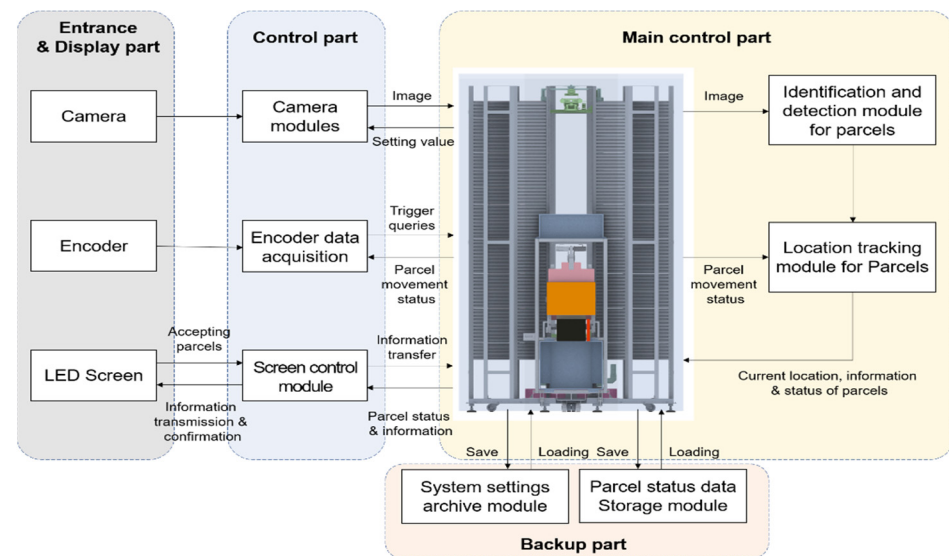


Figure 1. System block diagram.

2.2.1. Storage Entrance and Display Part

This part is responsible for acquiring images from the camera and setting the parameters of the control part. In this part, information recognition is performed using a camera to select a loading position. The encoder module communicates and processes information between the system and the encoder. Additionally, in this part, deep learning object recognition technology is used to analyze and detect parcel images, obtain location information, and transmit information to other parts.

2.2.2. Control Part

This part controls the image and information received in the previous part and builds a query program to obtain the moving state information data of the delivery service from the encoder. The control part exchanges information with the main control part in real time and can mostly transmit real-time image detection results and information to the central control part. Furthermore, this part allows querying of the interface layout to query historical data such as parcel detection.

2.2.3. Main Control Part

The main control part manages movement and loading information while tracking the loading position of parcels in real-time. In addition, it sets reasonable parameters, provides information to the control part, and plays a role in storing optimal parameters. This part aims to analyze the information of the parcels, track the moving process of the object, and save in the parcel storage.

2.2.4. Backup Part

This part is for backing up the setting information of the system during system operation and saving the state and storage data of the parcel box. This part plays a role in efficiently managing the system through data storage and preventing data loss.

3. Methodology

This section intends to detect and analyze parcels using a deep learning-based model in the parcel storage system. The first method introduces the basic performance of deep learning-based object detection models, and an experiment and analysis were conducted by deriving a model

that can be applied to the parcel storage system. Secondly, an intelligent e-commerce parcel storage system was proposed in the context of a deep learning detection system.

3.1. Deep Learning-Based Object Detection

Deep learning-based object detection algorithms are divided into R-CNN and YOLO model series. The R-CNN series is excellent in object detection with a high rate of accuracy but is disadvantaged due to its' detection speed being slower than the YOLO series [11]. Therefore, there is a limit to its application in detecting objects, for example parcels that move at high speed in real time in actual logistics sites. In this context, the YOLO algorithm series was developed to quickly learn the generalized characteristics of an object and solve the speed problem using the concept of regression. The YOLO Algorithm stands for 'You Only Look Once', and YOLO's various versions of models accurately and quickly classify the location and classification of detection targets using one-stage neural networks [12]. In 2015, Redmon et al. introduced the first version of YOLO [13], and in the past few years, several successor versions, such as yolov2, v3, v5, and v6 with improved performance in terms of computational speed and object detection have been announced [14–16].

YOLO has a simple structure that can directly output the location and category of the boundary box through the neural network as shown in the Figure 2 [17]. The YOLO architecture consists of 24 convolutional layers and two fully connected layers. YOLO estimates several bounding boxes per grid cell, but the bounding boxes which have the closest IoU (Intersection over Union) with the Ground Truth (GT) are selected. It is called non-maximum suppression [18].

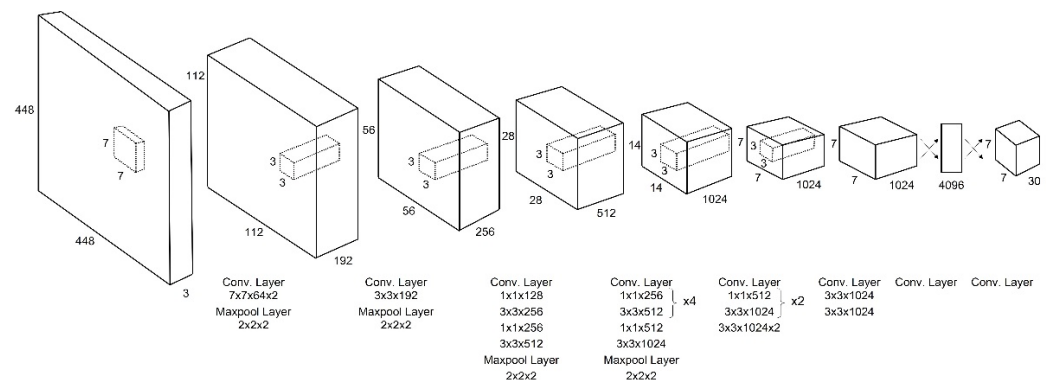


Figure 2. Architecture of YOLO Algorithm.

The original YOLO model has two disadvantages. The first is inaccurate location estimation, and the second is the low recall rate. Therefore, YOLOv2 presented an improved upgrade version in these two aspects. YOLOv2 does not deepen or widen the network but instead simplifies the algorithm to improve performance and speed [14]. Additionally, compared to YOLOv2, YOLOv3 is characterized by using multi-scale features for object detection and coordinating the original network structure. The feature extraction network of YOLOv3 used the residual model. The architecture of YOLOv3 contained 53 convolution layers, so it was called Darknet-53, compared with Darknet19 used by YOLOv2 [15]. The YOLOv4 algorithm has a great change when compared with previous models; it focuses on comparing data and significantly improves speed and performance [19]. YOLO v4 and v5 were proposed by developers other than the original YOLO developer.

YOLOv5 has features that can be applied to iOS using PyTorch. YOLOv5 may be relatively low in terms of accuracy compared to other models, but it shows an overwhelming performance in terms of speed. However, unlike existing models, YOLOv5 does not have published papers or technical reports. YOLOv5 is the latest enhancement that incorporates several new technologies that have already been proven in other networks [20]. Depending on the width and depth of the model, YOLOv5 can be divided into four versions: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x.

In the case of YOLOv6, while having a deeper network structure, it started detecting objects of more diverse sizes by increasing them from the previous three scales to four. It is characterized by having a much faster speed than the last v5 model [3]. The performance, speed, and analysis settings of YOLO algorithms are summarized in Table 2. Among various YOLO algorithms, YOLOv5, is the most efficient in terms of speed and performance and is being applied in several recent research papers, is used for this analysis to conduct research [11,21].

Table 2. Performance comparison between deep learning-based object detection.

Model	Size (Pixels)	Test Dataset	mAP (0.5)	FPS	GPU
YOLOv1	448 × 448	VOC2007	63.4	45	Titan X
YOLOv2	544 × 544	MS COCO	44.0	40	Titan X
YOLOv3	608 × 608	MS COCO	57.9	20	Titan X
YOLOv4	608 × 608	MS COCO	65.7	62	Tesla V100
YOLOv5s	640 × 640	MS COCO	56	140	Tesla V100
YOLOv5m	640 × 640	MS COCO	63.9		Tesla V100
YOLOv5l	640 × 640	MS COCO	67.2		Tesla V100
YOLOv5x	640 × 640	MS COCO	68.9		Tesla V100
YOLOv6s	640 × 640	MS COCO	35.9	550–620	Tesla V100
YOLOv6l	640 × 640	MS COCO	52.5		Tesla V100

3.2. Parcel Object Detection Algorithm

According to the depth of the network and the width of the feature map, YOLOv5 can be divided into four models: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x: S, M, L, and X, representing Small, Medium, Large, and Xlarge, respectively. It means that the model's overall structure remains constant as shown in Figure 3, but the size and complexity of each model are scaled [22]. YOLOv5s has the fastest processing speed and YOLOv5x has the highest detection accuracy among these models [23]. To find the applicable optimal algorithms for an unmanned parcel storage system, we trained and tested the four scales of YOLOv5. Table 3 shows the basic calculation amount and speed of YOLOv5.

Table 3. Basic Parameters Setting Values of YOLOv5.

Model	mAP (0.5)	mAP (0.5:0.95)	Params (M)	FLOPs@640 (B)
YOLOv5s	56.0	37.2	56	Tesla V100
YOLOv5m	63.9	45.2	63.9	Tesla V100
YOLOv5l	67.2	48.8	67.2	Tesla V100
YOLOv5x	68.9	50.7	68.9	Tesla V100

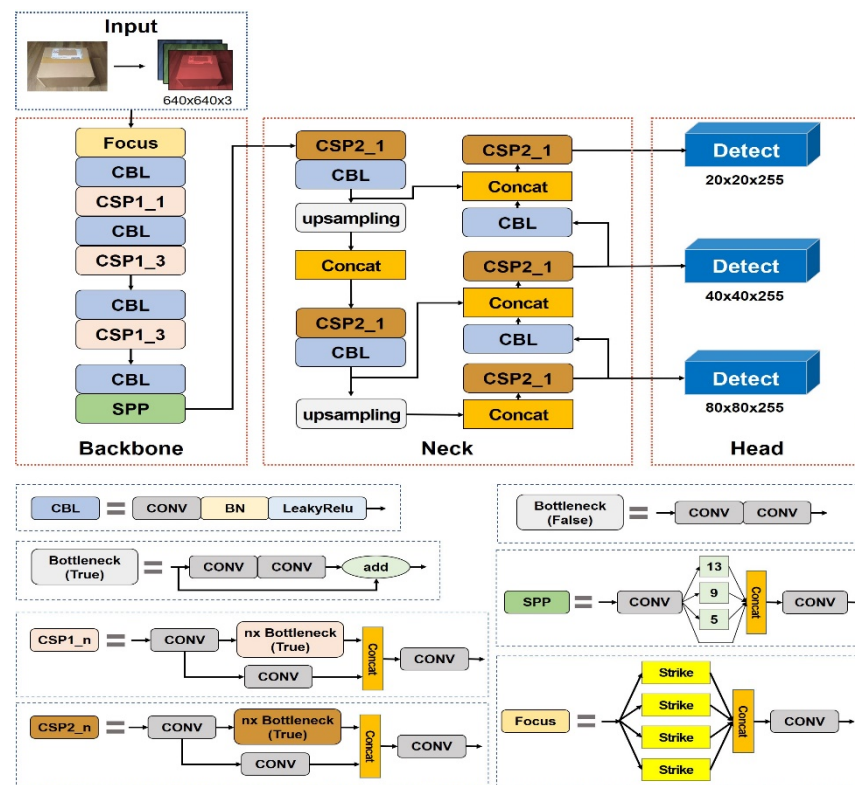


Figure 3. The architecture of the YOLOv5.

4. Experiments

In this chapter, experiments were conducted using the collected data sets to evaluate the performance of detailed versions of YOLOv5, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x.

4.1. Settings

The training process of the YOLOv5 model runs on the Windows 10 operating system and the PyTorch framework. The software analysis environment is CUDA11.2, CUDNN8.1, and Python3.9.2. The CPU used for training the dataset is an 11th Gen Intel Core i7 11,700 k @ 3.60 GHz 16 G, and the GPU is a GeForce RTX 3080 Ti 12 G. The CPU used for training the dataset is an 11th Gen Intel Core i7 11,700 k @ 3.60 GHz 16 G, and the GPU is a GeForce RTX 3080 Ti 12 G.

To avoid model oscillations due to high initial learning rates during model training, warmup training is introduced for all experimental learning rates in this paper to maintain the deep stability of the model [24]. During the warmup training phase, the learning rate rises from 0 to the set size of 0.01. After the warmup, the learning rate is updated [23]. The specific hyperparameters are set as follows: batch size is 16, learning rate(lr) is 0.01, momentum parameter is 0.937, weight attenuation factor to 0.0005, and there are 200 training epochs for the box dataset. In this paper, the input size is $640 \times 640 \times 3$ and data augmentation is used, such as mosaic, changing brightness, and rotation to expand the dataset before training is performed.

4.2. Performance Measurement

To analyze the results of the performance of YOLOv5, precision, recall, mAP@0.5, mAP@0.5:0.95, F1 score, and model size were used as evaluation indicators in this study. Precision is the ratio of correctly predicted positives to the total number of predicted positives, as shown in Equation (1). Precision refers to how accurately the model when compared with positive sample features. In other words, it is an indicator of how accurate the prediction value of the model is, and it is the ratio that correctly fits the positive among

the predictions. The higher the accuracy, the more accurate the prediction of positive samples [8].

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Recall is the percentage of correctly predicted positive numbers to the total number of true positives as shown in Equation (2). The higher the recall, the more accurately the target sample is predicted and the less likely it is to miss a bad sample [25].

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

where TP (True Positive) denotes the number of predicted positive cases that are correctly classified, and TN (True Negative) is the number of predicted negative cases classified as negative. FP (False Positive) is the number of predicted positive cases that are incorrectly classified as negative cases, and FN (False Negative) is the number of predicted negative cases that are incorrectly classified as positive cases [26]. The criteria for evaluating performance are defined using the above four criteria.

F1 score is the harmonic mean of precision and recall, shown in Equation (3) [27]. The F1 score accurately evaluates the performance of the model when the data label is an unbalanced structure and expresses the performance in a number. It is also called the test accuracy of the model. The value of F1 is from 0 to 1, and the more perfect precision and recall values are, the closer it is to 1, and the lower precision and recall values, the closer it is to 0 [28].

$$F1 \text{ score} = 2 \times \frac{(\text{precision} \times \text{recall})}{\text{precision} + \text{recall}} \quad (3)$$

In addition, mAP is the performance of the multiple object detection algorithm in one scalar value. It is calculated by taking the mean of the AP (Average Precision) of the classes [29].

4.3. Data Acquisition

Object recognition experiments should accurately identify the detection and location of the parcel boxes. Therefore, 3410 datasets were constructed by collecting delivery box images in various domestic environments. In the experiment, the ratio of learning, validation, and testing was 6:2:2 and 2728 pictures were applied to the training set, 680 pictures to the verification set, and 682 pictures to the verification set. As for the form of the original data, the experiment was conducted with various parcel sizes with different image pixels, as shown in Figure 4.



Figure 4. Parcel image dataset sample.

4.4. Experimental Results and Comparisons

In this section, experiments were conducted with various types of YOLOv5 models, such as s (small), m (medium), l (large), and x (xlarge). Parameter values have high values in the order of YOLOv5 s, m, l, and x, and the model of the x version can be regarded as the heaviest.

YOLOv5s, m, l, x can be seen in Figure 5, which displays different performance metrics for both the training and validation sets. The improvement of YOLOv5s, m, l, and x can be seen in Figure 5, which displays different performance metrics for both training and validation sets. Training and validation results, shown in Figure 6, showed rapid improvement in terms of precision and recall, plateauing after approximately 180 epochs. The performance evaluation index also showed a sharp decrease at approximately 180.

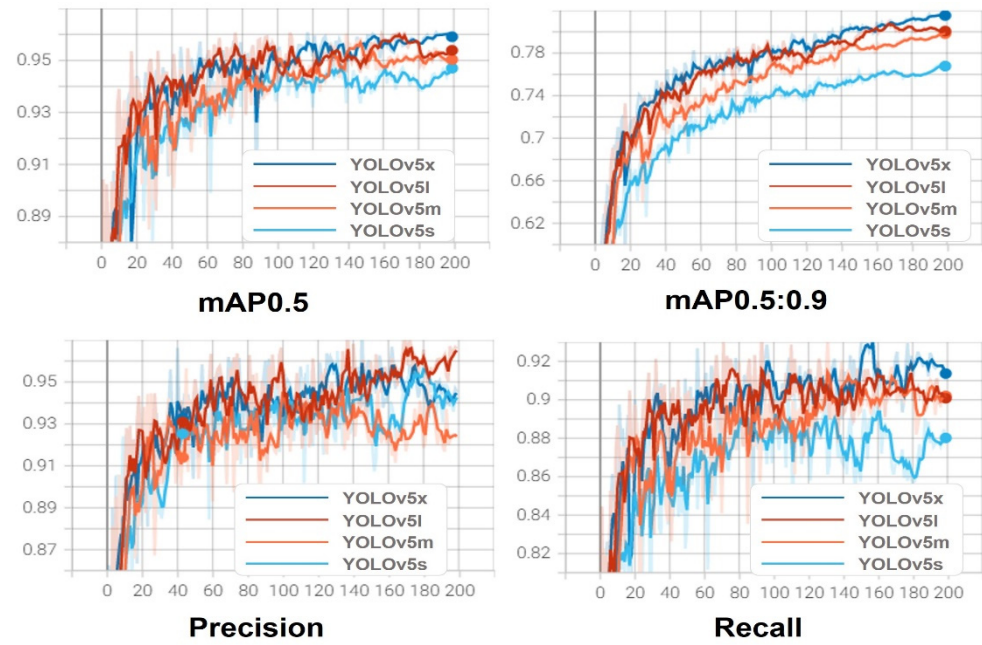


Figure 5. Experiment results of YOLOv5s, m, l, x.

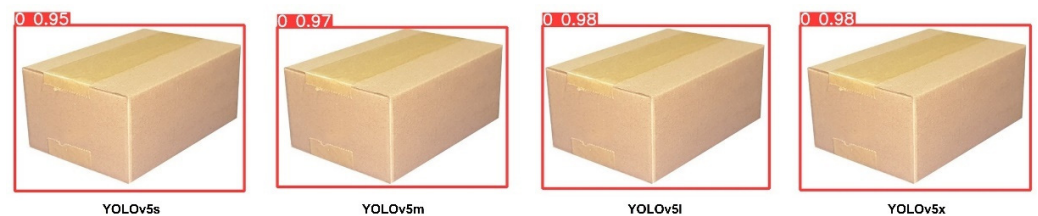


Figure 6. Comparison of test results of YOLOv5 models.

As a result of comprehensively analyzing s, m, l, and x of the YOLOv5 model, the precision of YOLOv5l was the highest at 0.966, followed by x, s, and m values. In the case of recall, the value of YOLOv5x was the highest at 0.910, followed by m, l, and s. In the case of mAP (0.5:0.95), YOLOv5x was the highest at 0.815, followed by l, m, and s, respectively. Finally, in the case of the F1 value that measures accuracy, YOLOv5l had the highest performance at 0.932, followed by x, m, and s, respectively. The performance values of YOLOv5 for each detailed model are shown in Table 4.

Table 4. Experiment result of YOLOv5s, m, l, x.

Model	Precision	Recall	mAP (0.5)	mAP (0.5:0.95)	F1	Parameter
YOLOv5s	0.952	0.865	0.946	0.773	0.906	7,012,822
YOLOv5m	0.923	0.903	0.951	0.798	0.913	20,852,934
YOLOv5l	0.966	0.899	0.955	0.801	0.932	46,108,278
YOLOv5x	0.945	0.910	0.958	0.815	0.928	84,173,414

As a result of learning and testing the detailed model of YOLOv5, it was concluded that YOLOv5l and YOLOv5x have the highest performance and highest Ground Truth, as shown in Figure 6. Among them, considering the parameter value and calculation speed, the YOLOv5l model is best suited for application to parcel boxes, shown in their performance and Ground Truth as analyzed in Figure 7.

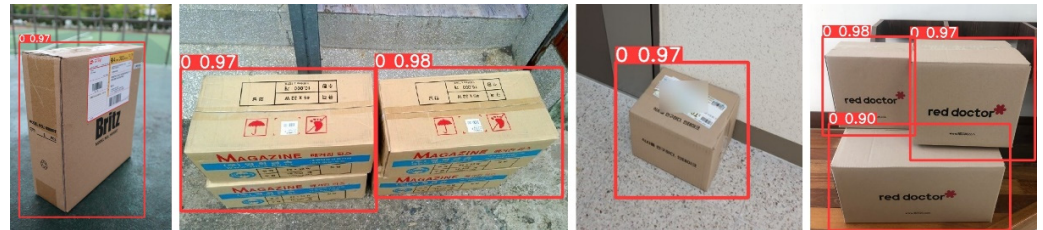


Figure 7. Test result of YOLOv5l.

5. Conclusions and Future Work

In this paper, based on YOLOv5, a parcel box recognition study was conducted to build an intelligent parcel storage system. The YOLOv5 model, which is divided into four detailed versions according to the performance and size of the model, was used for comparative analysis of parcel box recognition, and 3410 images were collected for the experiment. As a result of the experiment, it was analyzed that the precision, recall, and F1 of the YOLOv5l model were 0.966, 0.899, and 0.932, respectively, showing a stronger performance than the other models. The YOLOv5l model showed a performance 2.6% higher in precision, 0.6% higher in recall, and 1.6% higher in F1 score compared with the other version. In the case of the parameter value indicating the size of the model, the YOLOv5l model showed a larger value than YOLOv5s, m but showed a value half that of the YOLOv5x model. As a result, YOLOv5l showed the best performance in terms of the size of the model, and it was analyzed as an object recognition model suitable for application to an actual delivery box. In this study, a parcel recognition model with optimal performance was proposed through an experiment using the YOLOv5 model, and a basis for applying object recognition technology to an actual intelligent parcel storage system was laid.

In this study, a parcel recognition experiment was conducted to build an intelligent parcel storage system using the YOLOv5 model, and the following research limitations and future research tasks were derived. Firstly, it is necessary to conduct experiments in more diverse environments and more diverse datasets by securing additional datasets. Secondly, with the rapid development of the YOLO model, the YOLOv7 model appeared in July 2022. Therefore, it is necessary to conduct parcel recognition research using the latest model. In addition, through model transformation such as pruning, it is possible to propose an improved model which is optimized for parcel recognition as a focus of future study. Finally, by applying the object recognition model to the parcel storage system under development, future research will be able to confirm its performance in the real environment.

Author Contributions: Conceptualization, Y.K.; methodology, M.K.; software, M.K.; resources, M.K. and Y.K.; data curation, writing—original draft preparation, M.K.; writing—review and editing, Y.K.; supervision, Y.K.; project administration, Y.K.; funding acquisition, Y.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant from R&D program of the Korea Evaluation Institute of Industrial Technology (20014664).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to personal information in the invoice of the parcels.

Acknowledgments: The authors would like to thank the participants who volunteered to participate in this experiment.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Louis, E. Courier Management System. Master's Thesis, Riara University, Nairobi, Kenya, 2021.
- Wang, W.; Zhong, Y.C.; Ma, J.M. The status quo and future of smart parcel cabinets in the development of smart technology in my country's logistics express industry. *Mod. Bus* **2018**, *34*, 20–21.
- Zhang, S.Z.; Lee, C.K.M. Flexible vehicle scheduling for urban last mile logistics: The emerging technology of shared reception box. In Proceedings of the 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2016), Bali, Indonesia, 4–7 December 2016; pp. 1913–1917.
- Kolarovszki, P.; Vaculík, J. Intelligent storage system based on automatic identification. *Transp. Telecommun. J.* **2014**, *15*, 185–195. [\[CrossRef\]](#)
- Muzhir, S.; Al-Ani, M. Packages tracking using RFID technology. *Int. J. Bus. ICT* **2015**, *1*, 12–20.
- Mohd, A.; Rashid, R.A.; Hamid, A.H.F.A.; Sarijari, M.A.; Rahim, M.R.A.; Sayuti, H.; Rashid, M.R.A. Development of android-based apps for courier service management. *Indones. J. Electr. Eng. Comput. Sci.* **2019**, *15*, 1027–1036. [\[CrossRef\]](#)
- Bushanam, V.N.; Reddy, C.H. An Efficient CNN a deep learning approach applied on the image matching context. *Int. J. Eng. Technol.* **2018**, *7*, 507–512. [\[CrossRef\]](#)
- Kim, M.; Kwon, Y.; Kim, J.; Kim, Y. Image Classification of Parcel Boxes under the Underground Logistics System Using CNN MobileNet. *Appl. Sci.* **2022**, *12*, 3337. [\[CrossRef\]](#)
- Kim, M. Improvement of Recognition Performance through Refinement of Parcel Damage Classification Algorithm based on CNN. Ph.D. Thesis, University of Science & Technology, Daejeon, Korea, 2022.
- Chen, X. E-Commerce Logistics Inspection System Based on Artificial Intelligence Technology in the Context of Big Data. *Secur. Commun. Netw.* **2022**, *2022*, 3418466. [\[CrossRef\]](#)
- Yao, J.; Qi, J.; Zhang, J.; Shao, H.; Yang, J.; Li, X. A real-time detection algorithm for Kiwifruit defects based on YOLOv5. *Electronics* **2021**, *10*, 1711. [\[CrossRef\]](#)
- Ruan, J. *Design and Implementation of Target Detection Algorithm Based on YOLO*; Beijing University of Posts and Telecommunications: Beijing, China, 2019.
- Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 79–788.
- Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
- Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
- Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
- Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of Yolo algorithm developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [\[CrossRef\]](#)
- Jamtsho, Y.; Riyamongkol, P.; Waranusast, R. Real-time Bhutanese license plate localization using YOLO. *ICT Express* **2020**, *6*, 121–124. [\[CrossRef\]](#)
- Fan, S.; Liang, X.; Huang, W.; Zhang, V.J.; Pang, Q.; He, X.; Li, L.; Zhang, C. Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOV4 network. *Comput. Electron. Agric.* **2022**, *193*, 106715. [\[CrossRef\]](#)
- Chen, S.; Yang, D.; Liu, J.; Tian, Q.; Zhou, F. Automatic weld type classification, tacked spot recognition and weld ROI determination for robotic welding based on modified YOLOv5. *Robot. Comput.-Integr. Manuf.* **2023**, *81*, 102490. [\[CrossRef\]](#)
- Song, Q.; Li, S.; Bai, Q.; Yang, J.; Zhang, X.; Li, Z.; Duan, Z. Object Detection Method for Grasping Robot Based on Improved YOLOv5. *Micromachines* **2021**, *12*, 1273. [\[CrossRef\]](#)
- Benjumea, A.; Teeti, I.; Cuzzolin, F.; Bradley, A. YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles. *arXiv* **2021**, arXiv:2112.11798.
- Dong, X.; Yan, S.; Duan, C. A lightweight vehicles detection network model based on YOLOv5. *Eng. Appl. Artif. Intell.* **2022**, *113*, 104914. [\[CrossRef\]](#)
- Xiong, R.; Yang, Y.; He, D.; Zheng, K.; Zheng, S.; Xing, C.; Liu, T. On layer normalization in the transformer architecture. In Proceedings of the International Conference on Machine Learning (PMLR 2020), Virtual, 18 July 2021; Volume 119, pp. 10524–10533.
- Zhou, F.; Zhao, H.; Nie, Z. Safety helmet detection based on YOLOv5. In Proceedings of the 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), Shenyang, China, 22–24 January 2021; pp. 6–11.
- Yang, Y.; Zhang, L.; Du, M.; Bo, J.; Liu, H.; Ren, L.; Deen, M.J. A comparative analysis of eleven neural networks architectures for small datasets of lung images of COVID-19 patients toward improved clinical decisions. *Comput. Biol. Med.* **2021**, *139*, 104887. [\[CrossRef\]](#)
- Davis, J.; Goadrich, M. The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd international conference on Machine learning, Pittsburgh, PA, USA, 25–29 June 2006; pp. 233–240.

28. Nepal, U.; Eslamiat, H. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. *Sensors* **2011**, *22*, 464. [[CrossRef](#)]
29. Maxwell, A.E.; Warner, T.A.; Guillén, L.A. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 1: Literature review. *Remote Sens.* **2021**, *13*, 2450. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.