International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Learning Algorithm Training and Performance Analysis for Corridor Monitoring

Woo-Jin Jung, Won-Hyuck Choi

Submitted: 11/10/2023 **Revised**: 12/12/2023 **Accepted**: 21/12/2023

Abstract: K-UAM will be commercialized through maturity after 2035. Since the UAM corridor will be used vertically separating the existing helicopter corridor, the corridor usage is expected to increase. Therefore, a system for monitoring corridors is also needed. In recent years, object detection algorithms have developed significantly. Object detection algorithms are largely divided into one-stage model and two-stage model. In real-time detection, the two-stage model is not suitable for being too slow. One-stage models also had problems with accuracy, but they have improved performance through version upgrades. Among them, YOLO-V5 improved small image object detection performance through Mosaic. Therefore, YOLO-V5 is the most suitable algorithm for systems that require real-time monitoring of wide corridors. Therefore, this paper trains YOLO-V5 and analyzes whether it is ultimately suitable for corridor monitoring. K-UAM will be commercialized through maturity after 2035.

***Keywords**: Urban Air Mobility, Advanced Air Mobility, Object Detection Algorithms, YOLO*

1. Introduction

Urban Air Mobility (UAM) is a sub-concept of the next generation of Advanced Air Mobility (AAM) and refers to an operating system that operates low-altitude environments in and outside the city [1] UAM can operate between urban points faster than conventional transportation [2].

It is anticipated to emerge as a novel substitute transportation system, addressing challenges related to cost associated with traffic congestion in urban areas [3].he commercialization of National Urban Air Transportation (K-UAM) is envisioned to progress through initial stages from 2025 to 2030, followed by a growth period spanning 2030 to 2035, and reaching maturity beyond 2035. It is necessary to designate a UAM-only corridor because there are many risks when UAM is allowed to fly freely. Currently, plans are underway to vertically separate and use the helicopter corridor [4]. As UAM is commercialized, the capacity of the corridor increases and a monitoring system is needed to monitor it.

In order to carry out the monitoring system, UAM must be trained and learned through an object detection algorithm. This entails the systematic acquisition of knowledge and skills by the UAM system, allowing it to identify and interpret relevant objects within its operational environment. The object detection algorithm serves as a crucial component, enabling the UAM to discern and categorize various entities or obstacles, thereby enhancing its ability to navigate, make informed decisions, and ensure the safety and efficiency of its operations. Through this continuous learning process, the UAM system can adapt to diverse scenarios, optimize performance, and contribute to the overall reliability and effectiveness of urban air transportation.

Training and learning with existing object detection algorithms is not suitable for real-time object detection that processes large amounts of data. In recent years, video and deep learning-based target inspection models have developed as image processing technology and artificial intelligence technology have developed rapidly [5, 6]. Therefore, in this paper, we propose a deep learning algorithm that enables real-time object detection to detect UAM, a moving object in the corridor, and perform fitness analysis through algorithm training and performance evaluation.

2. Existing object detection algorithms

Traditional object detection algorithms typically extract features through manual extraction [7]. The moving object is extracted from the video sequence and the extracted features are classified by a classifier to achieve the object identification purpose [8].

Object detection methods include a background update method, a frame difference method, and an optical flow method. Background update methods use the idea of weighted averages, and background update impacts often affect the completeness of target extraction and the accuracy of target detection [9]. The frame difference method aims to achieve the object extraction purpose by calculating a difference between adjacent frames [10]. This method is often greatly influenced by the time interval between the vehicle's speed and the continuous frame. The optical flow method is a method of estimating

Student, Department of Avionics, Hanseo University, Seosan 3162, Korea
20200130@office.hanseo.ac.kr

density at the pixel level. The difference between the previous frame and the current frame is used, and the movement is calculated and extracted through the relationship between the pixel value and the surrounding pixels [11]. There are many difficulties in implementing algorithms because the condition that there should be no change in lighting is necessary. Classification controllers are used to classify and accurately locate specific locations of each target in the candidate box. Since large amounts of data are separated in real-time detection, the algorithm used is limited. Generally, there are SIFT, HOG, Harr, etc. Classification controller devices include SVM and Adaboost.

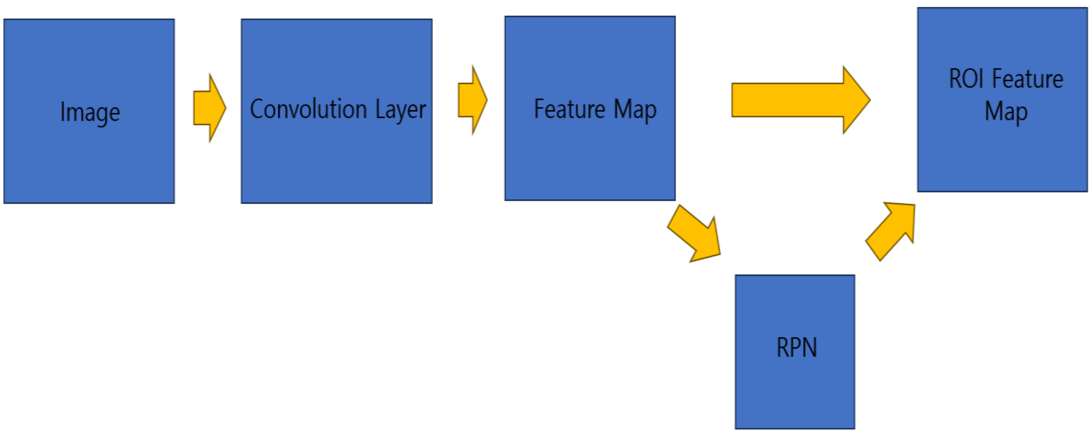
3. Real-time Object Detecting Algorithm

1. Real-time Object Detecting Algorithm

Algorithms used for time object detection can be divided into two categories. One is an R-CNN series calculation method based on two steps of the candidate region. The Region Based Convolutional Neural Network (R-CNN) is a neural network structure proposed by ross girsshick in 2014. R-CNN is the most basic model of region-based neural network structures, and it is an algorithm in which classification and localization occur sequentially. The basic components of RCNN are RoI extractors, feature extraction modules, classification modules, and location-specific modules. R-CNN [12], The computational load is very large and slow because more than 2,000 RoIs are proposed for each image and these areas are all structures that pass through the entire pipeline. To address these issues, it has developed into Ross girshick's Fast R-CNN in 2015 and Xiao’s Fast R-CNN in 2016. The structure of the Fast R-CNN algorithm is shown in Fig. 1. The Fast R-CNN [12]. algorithm introduced Region Proposition Network (RPN) instead of selective inspection based on the existing R-CNN, and RPN extracted RPN into a depth-grossing neural network instead of selective search and predicted the reliability score of the target boundary box and category at each location simultaneously to speed network calculation.

Since the series of R-CNN is a two-stage object detection algorithm, selecting area candidates and classifying objects does not occur at the same time. Therefore, in terms of speed, the performance is not good compared to the algorithm described later. Although Fast R-CNN has improved this problem, the improved performance value is also lower than the subsequent algorithm.

Fig 1. Faster R-CNN



2. One-Stage Detecting Algorithm

Unlike the two-step method, the one-step method does not generate candidate boxes in advance. The prediction and detection processes occur simultaneously in the candidate box. Typical algorithms include YOLO [13] and SSD. YOLO is the first object detection neural network proposed by Joseph Redmon in 2016. The image is divided into a grid and the boundary box and object classification are performed directly on the split area. As a result, a large number of boundary box candidates are generated, which is reduced to the final prediction result using MMS. YOLO [13] reached a speed of 45fps in 2016 and its accuracy is higher than that of RCNN, making it best suited for real-time object detection. YOLO [13] improved the loss function by adding FPN and other structures while upgrading to YOLO 9000 and YOLO V3[3] versions. In 2020, AlexAB attempted a new backbone network combination of YOLO V4 [14] to existing YOLO, increasing its average accuracy by 10% compared to YOLO V3 [14]. In the same year, Jocher Glenn released YOLOV5. The YOLO V5 consists of input, backbone, neck, and prediction, and has a similar structure to the existing V3 and V4, but the details have been slightly modified. Mosaic data augmentation was put in the input module, focus, CSP [14] structure in the backbone module, FPN+PAN configuration in the neck module, and GIOU_Loss loss function in the prediction to compensate for small object detection vulnerabilities, which are problems in the existing version. The backbone consists of five CSP modules of CSPnet, contains a total of 72 volume cores, and can extract 3x3 volume layers. The CSP module divides the features into two parts and integrates them in stages to secure accuracy while reducing the amount of calculation to 608x608x3 and outputs a 19x19 feature diagram through five CSP modules. The neck structure is mainly used to create a feature pyramid with a path integration network (PANET), and the feature pyramid allows the model to recognize targets on different scales and recognize the same object on multiple scales.

3. Select Algorithm

Recent YOLO models have steadily improved accuracy through FPN, FPN+PAN, and CSP modules to existing YOLO models. Among them, the YOLOV5 model reinforced the small object detection performance by adding Mosaic data augmentation to the input module. This is an algorithm suitable for corridor surveillance systems that monitor vast ranges through real-time streaming.

Therefore, this paper trains the YOLOV5 algorithm through helicopter and drone images and then analyzes the suitability of corridor monitoring through tests.

4. Algorithm Training and Results Analysis

1. Classifying and Learning Data Set

Roboflow's helicopter and drone images were divided into 3,829 train images, 692 validation images, and 1,476 test images to conduct learning. The batch and epoch were set to 16 and 30 respectively for simulation, and the total time required was 2 hours.

Fig 1. Faster R-CNN

2. Analysis of Results

Table 1. is a confusion matrix used to evaluate algorithm performance.

Table 1. Confusion matrix

		Actual Values	
		Positive	Negative
Predicted Values	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (NP)

The average average precision (mAP), which is most commonly used as an algorithm's performance evaluation indicator, refers to the average value of the lower area of the curve that displays the recall and precision on the x and y axes, respectively. Recall refers to the proportion of data whose predicted value and actual value match positively among those whose actual value is positive, and precision refers to the proportion of data whose predicted value and actual value match positively among those whose prediction is positively positive.

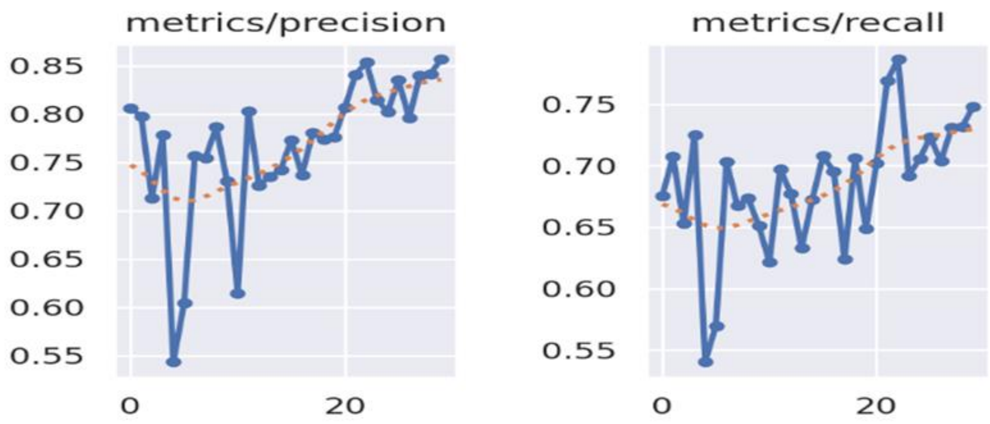


Fig 3. Train Results

Recall showed a value of 0.75 as training progressed, and precision showed a high value of 0.85 or more. Figure 4 shows the result value of the PR curve implemented through recall and precision.

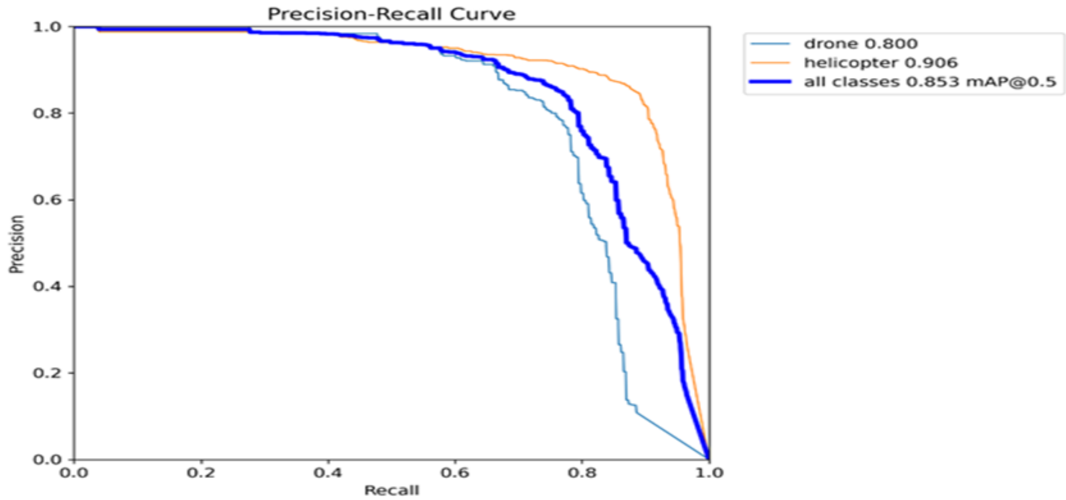


Fig 4. PR-Curve

Factors contributing to the development of marine leisure sports were initially presented through The mAP of helicopters and drones was shown using mAP@0.5. mAP@0.5 is a mAP result value in which the threshold value of the overlap rate (IoU) is 0.5. The mAP@0.5 of the helicopter class and the drone class were 0.906 and 0.800, respectively, with the object size within the helicopter's image being larger than that of the drone object and the predicted value of the helicopter was higher. The average value of mAP@0.5 in the two classes is 0.853, which is considered suitable for object detection in the corridor.

5. Conclusion

This study is based on the support of the technology development project of the Ministry of SMEs and Startups in 2023 [202201410002]

In this paper, the YOLOV5 model was selected as a deep learning algorithm for corridor monitoring. This was learned to analyze whether the actual YOLOV5 algorithm is suitable for corridor monitoring. The learning result of the YOLO V5 algorithm showed a high result value of mAP@0.5 with an average of 0.853. Since the YOLO algorithm, which has advantages in FPS, has improved accuracy, the model learned through class subdivision, data set refinement, and quantitative enhancement is considered to have sufficient performance to be used in real-time corridor monitoring systems

Acknowledgments

Reference

[1] Cohen, A. P., Shaheen, S. A., & Farrar, E. M. (2021). Urban air mobility: History, ecosystem, market potential, and challenges. IEEE Transactions on Intelligent Transportation Systems, 22(9), 6074-6087.

[2] Straubinger, A., Verhoef, E. T., & De Groot, H. L.

(2021). Will urban air mobility fly? The efficiency and distributional impacts of UAM in different urban spatial structures. Transportation Research Part C: Emerging Technologies, 127, 103124.

[3] Cho, S. H., & Kim, M. (2022). Assessment of the environmental impact and policy responses for urban air mobility: A case study of Seoul metropolitan area. Journal of Cleaner Production, 360, 132139.

[4] Tang, H., Zhang, Y., Mohmoodian, V., & Charkhgard, H. (2021). Automated flight planning of high-density urban air mobility. Transportation Research Part C: Emerging Technologies, 131, 103324.

[5] Zheng, X., Zheng, S., Kong, Y., & Chen, J. (2021). Recent advances in surface defect inspection of industrial products using deep learning techniques. The International Journal of Advanced Manufacturing Technology, 113, 35-58.

[6] Ye, X. W., Jin, T., & Yun, C. B. (2019). A review on deep learning-based structural health monitoring of civil infrastructures. Smart Struct. Syst, 24(5), 567-585.

[7] Blundell, J. S., & Opitz, D. W. (2006). Object recognition and feature extraction from imagery: The Feature Analyst® approach. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(4), C42.

[8] Dedeoğlu, Y. (2004). Moving object detection, tracking and classification for smart video surveillance (Doctoral dissertation, Bilkent Universitesi (Turkey)).

[9] Weng, M., Huang, G., & Da, X. (2010, October). A new interframe difference algorithm for moving target detection. In 2010 3rd international congress on image and signal processing (Vol. 1, pp. 285-289). IEEE.

[10] Wei, H., & Peng, Q. (2018). A block-wise frame difference method for real-time video motion detection. International Journal of Advanced Robotic Systems, 15(4), 1729881418783633.

[11] Agarwal, A., Gupta, S., & Singh, D. K. (2016, December). Review of optical flow technique for moving object detection. In 2016 2nd international conference on contemporary computing and informatics (IC3I) (pp. 409-413). IEEE.

[12] Shaoqing Ren, Kaiming He, Cho, Ross Girshick, Jiam Sun, "Faster r-cnn: Toward real-time object detection with region proposal networks," Advances in neural information processing systems 28, 2015

[13] Joseph Redmon, Santosh Divvala, Ross Girshik, Ali Farhadi, "You only look once: Unified, real-time object detection,"Proceedings of the IEEE conference on computer vision and pattern recognition, 2016

[14] J.A.Bochkovskiy, C. Y. Wang, and H. M. Liao, "YOLOV4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2024.10934, 2020