DOI 10.5162/ettc2022/5.1

Aircraft Tracking with Single Camera and RF - System Synchronization

Volkan Tutay[1], İrfan Karagöz[2]

[1] Gazi University, Graduate School of Natural and Applied Sciences, Department of Electrical-Electronic Engineering, 06560 Ankara, TURKEY,

[2] Gazi University Engineering Faculty Electrical and Electronics Engineering Department 06570 Maltepe/ANKARA

volkan.tutay@tai.com.tr

irfankaragöz@gazi.edu.tr

Abstract:

Flight tests are carried out to complete the flight test campaigns. During flight test, safety and time factors are at the forefront. The safety of aircraft during flight, reducing flight test schedule are provided by telemetry systems. In some cases, such as GPS loss, jamming, the radio frequency tracking methods used in these systems may be insufficient. In order to eliminate this situation, auxiliary systems have been developed for telemetry systems. In this study, a camera system will be used as an auxiliary system since camera systems provide visual tracking and are inexpensive. The aim of this study is to perform visual tracking with a single camera with deep learning methods and to synchronize with these systems to assist RF tracking systems.

Key words: Telemetry, Flight Test, Aircraft tracking, Computer vision, Deep learning.

1. Introduction

It is used for object detection and tracking in telemetry systems, cameras, mobile phones, autonomous vehicle technologies, security, automation systems, aviation and space fields.

The telemetry system as in show in Fig. 1. provides real-time monitoring and recording of sensor data and bus messages acquisition from aircraft and at the ground station so visual monitoring of experimental aircraft is critical in flight tests. Within the scope of this study, the detection and tracking of an aircraft at any speed and altitude will be discussed. The telemetry antenna will be synchronized with the location information of the detected aircraft and the antenna and camera system will work together.

Many problems are encountered while detecting an aircraft from camera. These are; instantaneous movements of the aircraft, background cloudiness, light and visibility variations, target diversity, noise in the image and real-time processing requirements [1]. Many studies have been carried out to eliminate these problems.

Object detection is basically divided into traditional methods and machine learning methods. In traditional methods, background subtraction, optical flow and frame difference methods are generally used [1], [2]. Although these methods have been prior in use, machine learning and deep learning algorithms provide higher accuracy and speed. Deep Learning Methods firstly extract a feature with convolutional features then classifier networks are used to recognize the features of the objects. Generally, these networks try to detect the object by scanning the whole frame or any region on the frame [3].

Many object detection methods have been studied in the literature, RCNN [4], Fast RCNN [5], Faster RCNN [6], YOLOv3 [7] methods have more accuracy or speed. In this study, these models performances will be compared in relation to accuracy and speed than traditional methods such as Scale-invariant feature transform (SIFT), Histogram of oriented gradients (HOG) features. In this application, it is not sufficient to just detect and classify the aircraft. It is also necessary to track the moving aircraft. Also the direction of movement of the aircraft must be estimated. Tracking algorithms are of two types, traditional and CNN based.

DOI 10.5162/ettc2022/5.1

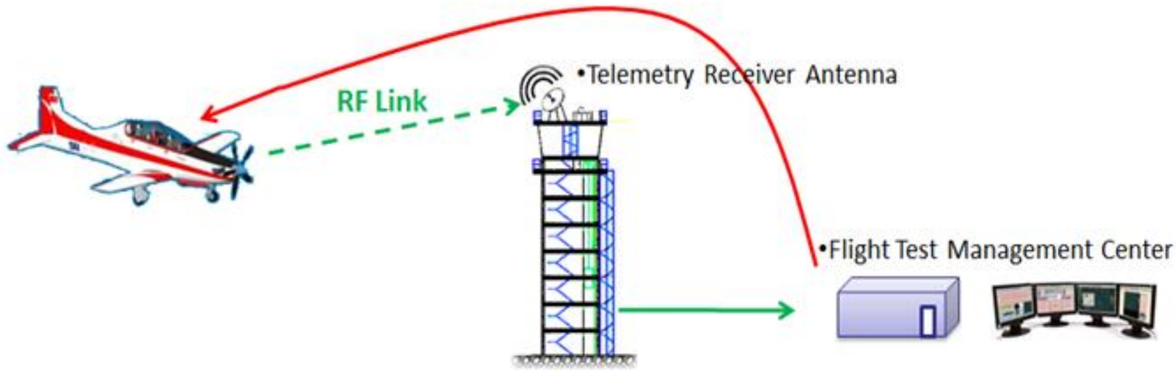

Fig. 1. Telemetry System

SORT, one of the most widely used tracking methods, will be used in this study.

The moving camera will be directed in accordance with the coordinates of the tracking aircraft. It will then move in synchronize with the telemetry antenna and camera.

2. Literature Review

There is no enough study in the literature within the scope of this study. For this reason, studies in which object detection and tracking are done with a moving camera are examined as separate subjects.

Dikbayır [8] used the Munich dataset, which includes cars, trucks, pickup trucks and buses in his study. The dataset includes vehicles photographed from different altitudes. Faster R-CNN and YOLOv3 algorithms were studied for each altitude. Although the YOLOv3 algorithm was successful in close vehicles, the Faster R-CNN algorithm was more successful in far vehicle photos. Although the YOLOv3 algorithm is faster for real time operations, it remains successful in detecting small objects. In this study, a new approach has been put forward by combining Faster R-CNN fed YOLOv3 algorithm and detection of vehicles was achieved with a higher success rate.

A performance comparison study was made by Wang et al. [9] using the Stanford University drone dataset with Faster R-CNN and SSD architectures on the RetinaNet algorithm. As a result of the study, they concluded that single-stage architectures successfully converged to two-stage architectures.

Another comparative study was done by Benjdira et al.[10]. In this study, they compared the one-stage YOLOv3 and two-stage Faster R-CNN model to realize vehicle surveillance and traffic monitoring. Dataset is created from images obtained from unmanned aerial vehicles. The performance evaluation and processing time of the models were examined. According to the results of the study, both algorithms achieved at least 99% and mention the object prediction accuracy in the dataset used. However, it was concluded that the YOLOv3 algorithm is more robust and has higher recall value than the Faster R-CNN model.

In the study conducted by Barış and Baştanlar[11], the classification of vehicles in traffic with PTZ camera was studied. The proposed method moves the camera according to the location of the detected object and then makes the classification of the vehicle. K-Nearest Neighbor method for object classification was tested in 4 different vehicle classes and the success of the method was 97.40%.

Maher et al.[12] proposed a target tracking system called deep-patch orientation network (DON) for tracking aircraft. This method predicts the direction of the target based on the training information. The DON method used the YOLOv3 and FrRCNN methods and the real-time tracking (SORT) algorithm. Experiments show that overall detection accuracy increases processing speed. Thus, the proposed method was more efficient for real-time operations.

3. Dataset and Methods

This section includes definition and preprocessing of dataset, object detection, object tracking and camera movements.

3.1 Dataset

In the literature, there are many data sets on object detection and tracking, but in this study, the data set was obtained from flight tests in Turkish Aerospace. This dataset was obtained by shooting 2 different fixed wing and 2 different rotary wing aircraft with different attributes from the ground camera. 1837 images were obtained from a total of 6 hours and 23 minutes of

DOI 10.5162/ettc2022/5.1

Fig. 2. Image labeling with LabelImg.

videos. These images were reproduced by mirroring, bleaching and rotation methods, which are data augmentation methods [13] and with the help of the augmentation methods 1837 images extended to 3280 images. The dataset split into training and testing as 70% of training, 30% of testing. To use in deep learning methods images need to be labelled so that create xml data for each images. Xml data include coordinate of image in frame and object class. Images were labeled with the LabelImg program as shown in Fig. 2.

3.2 Object Detection

Aircraft detection and classification were made in these images with R-CNN, Fast R-CNN, Faster R-CNN and YOLOv3, which are CNN based deep learning methods.

R-CNN method is the most basic model using the region recommendation approach as shown in Fig. 3. Fast R-CNN and Faster R-CNN are the developed and accelerated versions of this method. These models suggest regions with different sizes and in this model the window sizes are equalized by passing the relevant windows through conventional neural networks. At the end of the neural networks process, support vector machines (SVM) classifier is used to classify the object in that region. As a result of classification, it gives 4 coordinates indicating the location of the object in the image. On the other hand, Faster R-CNN classification is performed by linear regression. With the regression method used, the boundaries of the object are revealed [14].

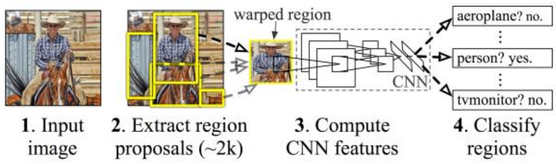

Fig. 3. R-CNN architecture

Unlike other deep learning methods, YOLOv3 does not operate with a regional-based approach, but in a convolutional network without fragmenting the image. It divides the image into grids of SxS, as shown in Fig.4., in accordance with its size and detects the object according to the similarity status [15]. In this way, the YOLOv3 algorithm is much faster than other deep learning methods. For this reason, it is widely used in the literature for object detection.

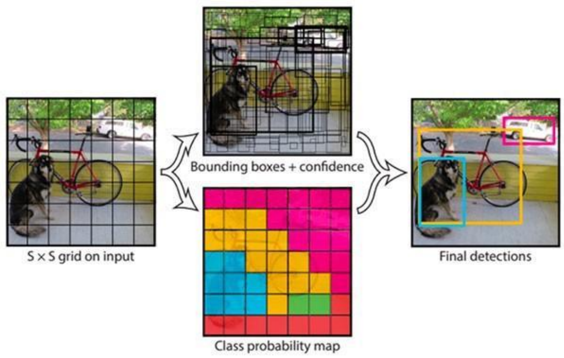

Fig. 4. YOLO SxS grid architecture

3.3 Object Tracking

SORT algorithm frequently used in real-time object tracking applicpations in the literature. SORT predicts the next location of the detected aircraft using the Kalman Filter. Intersection Over Union (IoU), one of the object association methods, is used in the SORT algorithm. This method makes object association according to the intersection of the previously detected

DOI 10.5162/ettc2022/5.1

object and the next location of which is estimated [16], as shown in Fig. 5.

SORT continues to track the object as long as the IoU score stays above a predefined threshold. If the object is lost for any reason, the tracking is interrupted and when the object is detected again, the tracking continues.

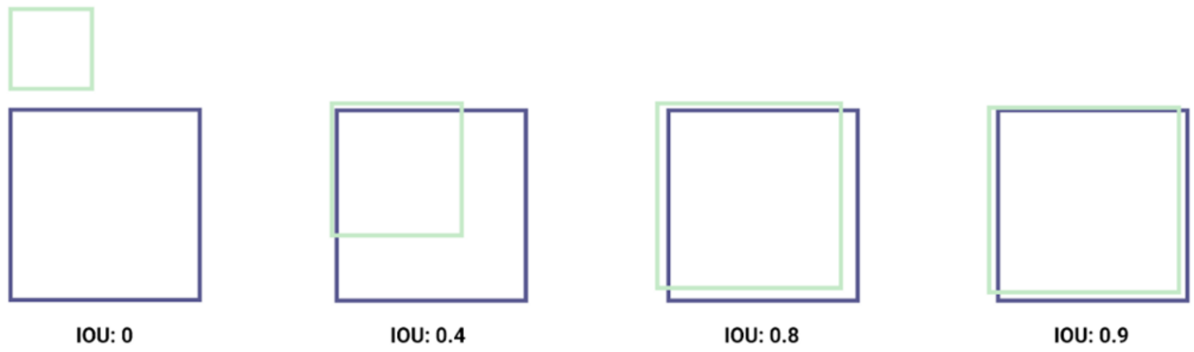

Fig.5. Sample IoU Scores

3.4 Camera Movements and Synchronization

The position values of the aircraft estimated by SORT are transmitted to the moving camera so that the camera can pan and tilt [17]. Camera movement is done according to the formula in (1).

$X = x + dx(1+\alpha),\ Y = y + dy(1+\alpha)$ (1)

The position values obtained after X and Y estimation, dx and dy instantaneous velocity values and the necessary variable for updating the α position are defined. Motion information is transmitted to the camera system via serial communication protocol (RS-232). With the generated PWM messages, the motors in the camera pedestal are moved accordingly.

The rate of change obtained from the movement of the camera motors is sent to the antenna control unit (ACU) of the telemetry antenna by RS-232. As a result of this, with the generated PWM messages, the antenna motors are moved synchronously with the camera motors.

4. Experimental Results

Four different methods used in object detection were compared. This comparison is provided by precision, recall, F1-Score, quality and IoU metrics.

Mean average precision (mAP) values obtained from precision, recall and IoU metrics. The results obtained with 3280 images are shown in Table 1.

Tab. 1: Metrics comparison

Method	mAP%
RCNN	86
Fast RCNN	90.4
Faster RCNN	94.2
YOLOv3	87.8

According to the test results, it was observed that the scores of the YOLO method were more successful. In addition to images, it has also been tested on videos from which the dataset was created. The fps values of the video of rotary wing aircraft performing taxi and take-off activities are given in Table. 2. In addition to these, the fps comparison of a section is as in Fig. 6.

Tab. 2: Average fps comparison

Method	Average fps
RCNN	30.7
Fast RCNN	4.8
Faster RCNN	43.1
YOLO	72.4

DOI 10.5162/ettc2022/5.1

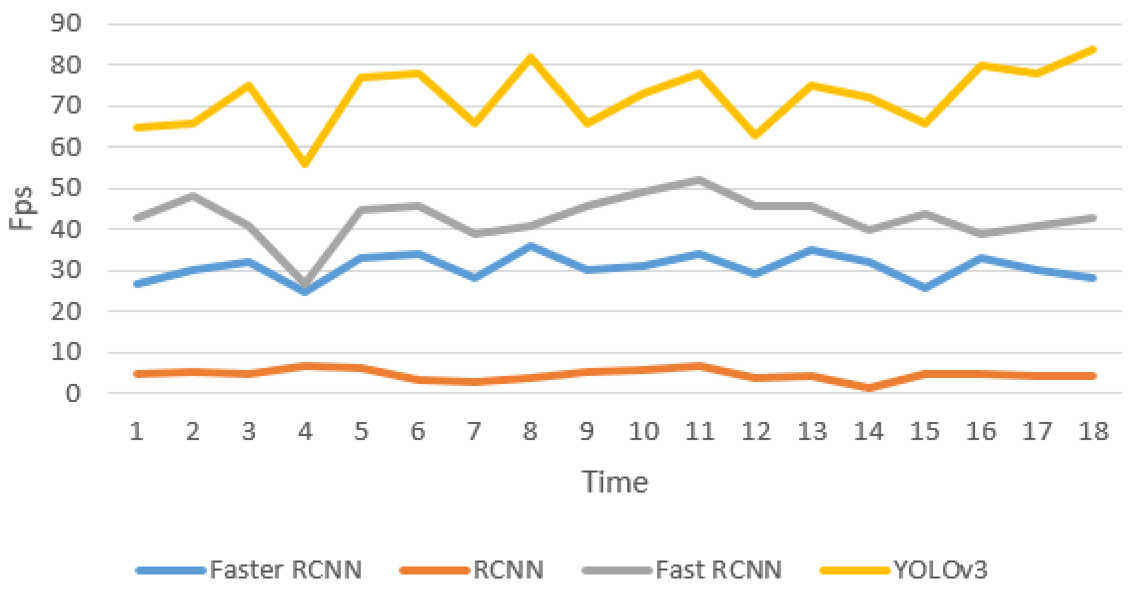

Fig.6. Fps comprision in video track

The sample results of the rotary wing aircraft detected by YOLOv3 are shown in Fig. 7 and Fig. 8.

Fig.7. Experimental result-1

Fig.8. Experimental result-2

DOI 10.5162/ettc2022/5.1

5. Conclusion and Future Works

In this study, metrics and frames per second (fps) values of RCNN, Fast RCNN, Faster RCNN and YOLOv3 models were compared for the for the detection of aircraft with motion cameras. Although the best mAP was obtained with the Faster RCNN model, the model with the highest fps value was YOLOv3. Since this application is processed in real-time, the YOLOv3 model is used.

SORT model was used for object tracking. Depending on the movement direction and speed of the tracked aircraft, the moving camera made pan-tilt movement and the camera and telemetry antenna synchronization was achieved.

In the future works, we will propose a new method that will be faster and have more accuracy.

6. References

[1] Liang, R., Yan, L., Gao, P., Qian, X., Zhang, Z., & Sun, H. (2010, October). Aviation video moving-target detection with inter-frame difference. In 2010 3rd International Congress on Image and Signal Processing (Vol. 3, pp. 1494-1497). IEEE.

[2] Pakfiliz, A. G. (2017). Automatic detection of aerial vehicle in cloudy environment by using wavelet enhancement technique. Radioengineering, 26(4), 1169-1176.

[3] Zhu, M., & Wang, H. (2017, October). Fast detection of moving object based on improved frame-difference method. In 2017 6th International Conference on Computer Science and Network Technology (ICCSNT) (pp. 299-303). IEEE.

[4] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 580–587, 2014, doi: 10.1109/CVPR.2014.81.

[5] R. Girshick, “Fast R-CNN,” Proc. IEEE Int. Conf. Comput. Vis., vol. 2015 International Conference on Computer Vision, ICCV 2015, pp. 1440–1448, 2015, doi: 10.1109/ICCV.2015.169.

[6] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real- time object detection,” IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.

[8] BÜLBÜL, H. (2020). DERİN ÖĞRENME YÖNTEMLERİ KULLANARAK GERÇEK ZAMANLI ARAÇ TESPİTİ.

[9] X. Wang, P. Cheng, X. Liu ve B. Uzochukwu, «Fast and Accurate, Convolutional Neural Network Based Approach for Object Detection from UAV,» arXiv.org, 2018.

[10] B. Benjdira, T. Khursheed, A. Koubaa, A. Ammar ve K. Ouni, «Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3,» arXiv, 2018.

[11] Barış, İ., & Baştanlar, Y. (2016, May). Classification of vehicles with omnidirectional and PTZ cameras. In 2016 24th Signal Processing and Communication Application Conference (SIU) (pp. 441-444). IEEE.

[12] Maher, A., Taha, H., & Zhang, B. (2018). Realtime multi-aircraft tracking in aerial scene with deep orientation network. Journal of Real-Time Image Processing, 15(3), 495-507.

[13] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1), 60. doi: 10.1186/s40537-019-0197-0

[14] Resul, D. A. Ş., Polat, B., & Tuna, G. (2019). Derin öğrenme ile resim ve videolarda nesnelerin tanınması ve takibi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 31(2), 571-581.

[15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only lookonce: Unified, real- time object detection,” IEEE Comput. Soc. Conf.Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 779–788, 2016,doi: 10.1109/CVPR.2016.91.

[16] Bathija, A., & Sharma, G. (2019). Visual object detection and tracking using Yolo and sort. International Journal of Engineering Research Technology, 8(11).

[17] G. Scotti, L. Marcenaro, C. Coelho, F. Selvaggi, and C. Regaz-zoni, “Dual camera intelligent sensor for high definition 360 degreessurveillance,” IEEE Proceedings-Vision, Image and Signal Processing,vol. 152, no. 2, pp. 250–257, 2005.