5th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Aircraft detection using optical remote sensing images and YOLOv7 based deep learning method

Roya Talebi *1

1University of Tabriz, Faculty of Planning and Environmental Sciences, Department of Remote Sensing and GIS, Tabriz, Iran

Keywords
Remote sensing image
Aircraft detection
YOLOv7

Abstract

Target detection is an important application in remote sensing. Aircraft detection in remote sensing images (RSIs) has attracted widespread attention in recent years, which has been widely used in both military and civilian fields. In this study, we used YOLOv7 (You Only Look Once) model for Aircraft detection from satellite images. The results show that YOLOv7 can be well applied in the field of Aircraft detection. The YOLOv7 model with mAP, Precision, Recall, F1 score and average detection time of 95.62%, 93.27%, 93.41%, 93.34% and 0.025 s per image performed well in recognition.

1. Introduction

Aircraft detection in remote-sensing images is a fundamental task in civil and military applications. Deep learning techniques to achieve end-to-end object detection have attracted the attention of the Earth observation community (Lin & Chen, 2021).

Different disciplines and applications have benefited from DL methods. In the Remote Sensing (RS) domain, DL methods are also used for the detection of different geospatial objects, land cover/use segmentation, and pan-sharpening (Krizhevsky et al., 2017; Li et al., 2017; Redmon et al., 2016; Yaban et al., 2022).

YOLO is a commonly used single-stage target detection algorithm with the characteristics of fast and high accuracy (Li et al., 2022; Redmon et al., 2016). It exhibits satisfactory performance in detecting small and occluded targets in complex field environments and has better detection speed than other deep learning algorithms (Lu et al., 2019). YOLOv7 is the latest detector in YOLO series. This network is designed with trainable bag-of-freebies, which enable real-time detectors to greatly improve the accuracy without increasing the inference cost. It also involves extend and compound scaling so the target detector can effectively reduce the number of parameters and calculations, thereby greatly improving the detection speed (Wang et al., 2022; Wu et al., 2022). At pre-sent, YOLOv7, as a brand-new detector, has not been applied to Aircraft detection. Therefore, in the present work, YOLOv7 was used to detect Aircrafts.

The hardware used during computer development was to run the software. The object detection model was trained using, laptop computer with access to a Google Colab virtual machine, which offers free GPU cloud service that allows one to obtain 0.007 second inference time. That is, 140 FPS on a TESLA P100 GPU.

In this study, we aimed to automatically detect Aircrafts from very high-resolution satellite images using the DOTA dataset and a new test data set generated from satellite images of different airports and air bases obtained from the Google Earth platform.

2. Method

2.1. Data and Environment

In this study, we have used the following data:

(1) The DOTA dataset was used for training and testing purposes. It is an open-source dataset for object detection purposes from remote sensing images. The dataset includes satellite image patches obtained from the Google Earth© platform, and Jilin 1 (JL-1) and Gaofen 2 (GF-2) satellites. It contains 16 types of objects (plane, ship, storage tank, baseball diamond, tennis court, basketball court, ground track field, harbor, bridge, large vehicle, small vehicle, helicopter, roundabout, soccer ball field, swimming pool, and container crane) (Gong et al., 2022). The image sizes are in the range of 800 × 800 to 4000 × 4000. DOTA dataset was divided into 1600 training and 400 validation images.

* **Corresponding Author**

*(roya.talebi99@ms.tabrizu.ac.ir) ORCID ID 0000-0002-1553-1678

Cite this study

Talebi R., (2022). Aircraft detection using optical remote sensing images and YOLOv7 based deep learning method. 5th Intercontinental Geoinformation Days (IGD), 105-107, Netra, India

(2) In addition to this dataset, we also collected 100 images from Google Earth and used them as the independent test set.

In this study, we were the images split to the size of 608 × 608 for training YOLOv7 detectors. We implemented our experiments in the Google Colab Pro development platform.

2.2. YOLOv7 Network Architecture

YOLOv7, a latest detector with YOLO architecture, is an object detection network that has fast detection speed, high precision and easy to train and deploy characteristics. The speed and accuracy of the network is within the range of 5–160 FPS, surpassing currently known object detectors. The network is 120% faster than YOLOv5 in the same volume (FPS). The test results on the MS COCO dataset outperform the YOLOv5 detector (Ahmad et al., 2022; Wu et al., 2022).

2.3. Establishment of Model

The establishment of Aircraft object detection model was divided into training and testing stages. The YOLOv7 neural network was trained using the training set, and the evaluation indicators were verified on the validation set after model weights were obtained. Finally, the model with the best performance weight was selected as the preliminary model for object detection for Aircraft. In the testing phase, the detection model was run on the test set. The workflow is illustrated in Figure 6.

2.4. Evaluation Indicators of Model

In this paper, Precision, Recall, Mean Average Precision (mAP) and F1 score were used to accurately and objectively evaluate the performance of the model. Precision is the most common evaluation index, and it is the number of right targets divided by the number of detected targets. In general, the higher the Precision is, the better the detection effect will be. Precision is a very intuitive evaluation index, but sometimes high Precision does not represent all. Therefore, mAP, Recall and F1 score were introduced for comprehensive evaluation. Precision (Eq.1), Recall (Eq.2), Average Precision (Eq. 3) mAP (Eq.4), and F1 score (Eq. 5) were calculated as follows:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

$$AP = \int_0^1 P(r)dr \quad (3)$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_I \quad (4)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (5)$$

where TP (True Positive) represents the number of Aircraft objects that are correctly detected; FP (False Positive) represents the number of other objects detected as Aircraft; and FN (False Negative) represents the number of Aircraft that are undetected/missed (Wu et al., 2022).

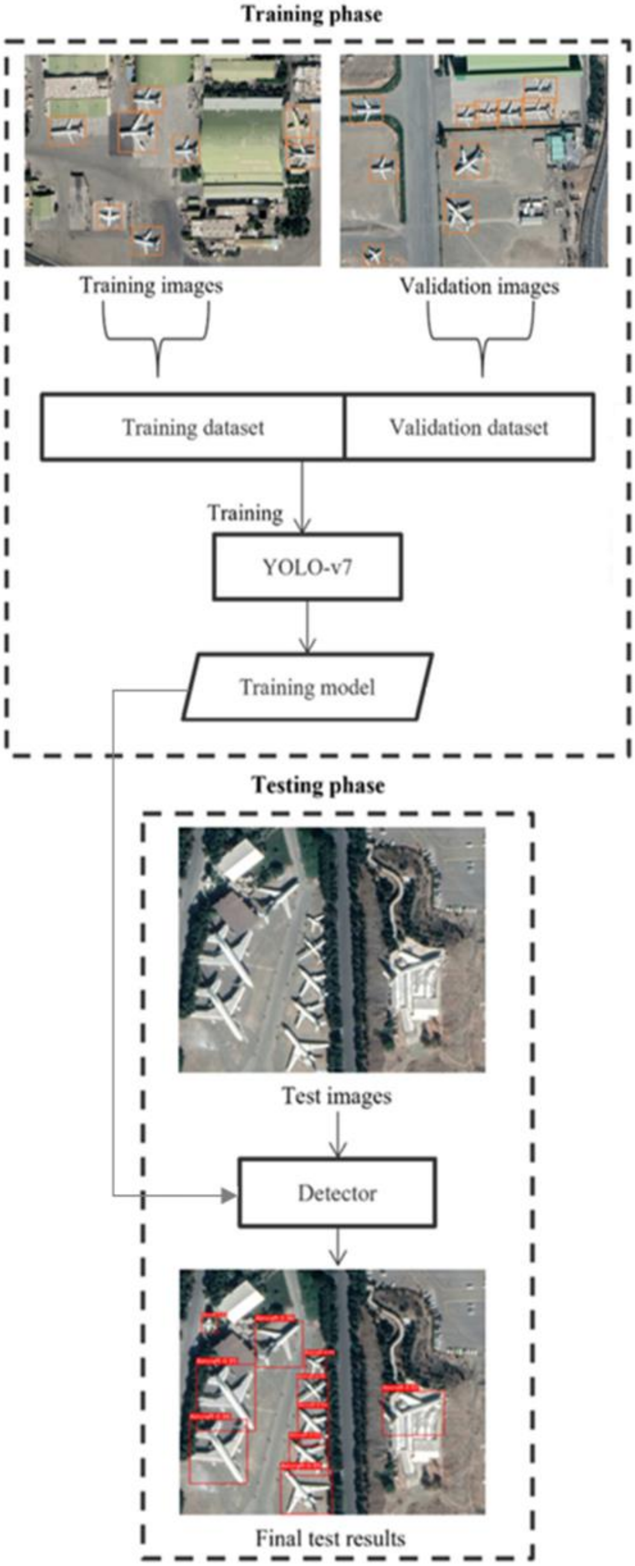

Figure 1. Workflow of the proposed study

3. Results and Discussion

The YOLOv7 model for object detection of Aircraft was established based on the original dataset and the

YOLOv7 network. This study determined that the model after 258 epochs was the suitable detection model for Aircraft.

The performance indicators of the YOLOv7 model shown in the Table 3. It can be seen that The YOLOv7 model with mAP, Precision, Recall, F1 score and average detection time of 95.62%, 93.27%, 93.41%, 93.34% and 0.025 s per image performed well in recognition.

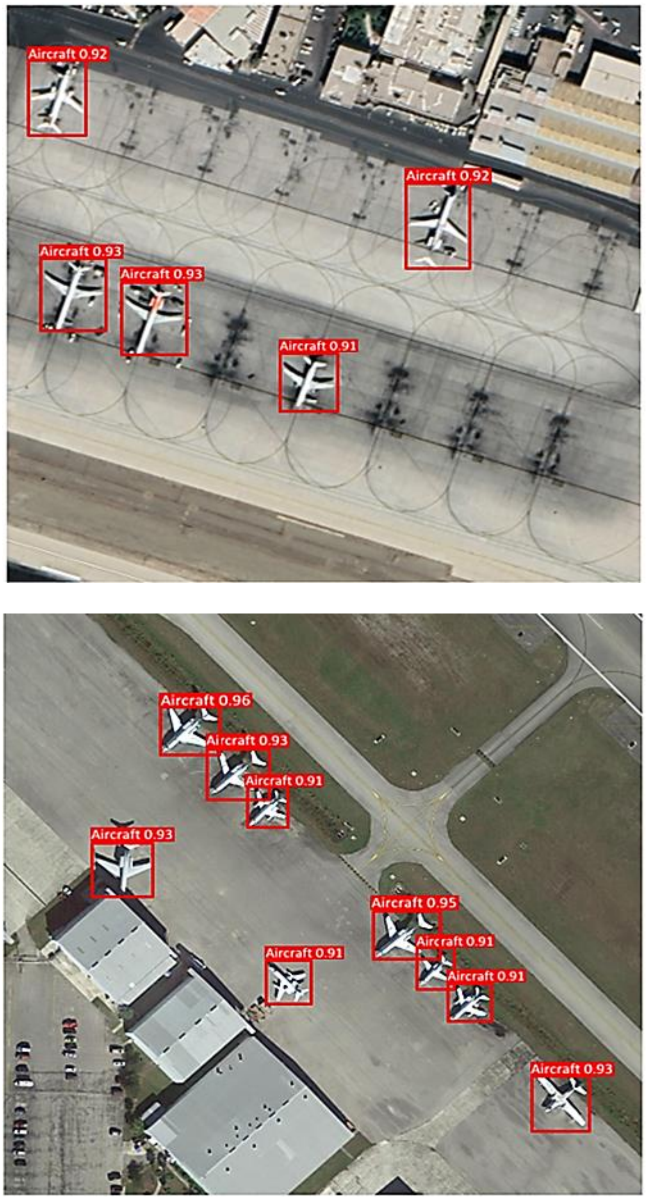

Figure 1. Detection previews from YOLOv7 Architecture

Table 1. Evaluation indexes results

Target Detection Network	mAP (%)	Precision (%)	Recall (%)	F1 Score (%)	Average Detection Speed (s/Image)
YOLOv7	95.62	93.27	93.41	93.34	0.025

Table 2. Detection result of YOLOv7 model

Model	Detection Results	Numbers
YOLOv7	Number of detected objects	1338
	Number of right objects (TP)	1248
	Number of wrong objects (FP)	90
	Number of missed objects (FN)	88

4. Conclusion

Preferring smaller scale models of YOLOv7 and using more powerful graphics cards can enable model training with higher number of batch sizes, thus may result in higher success rates indirectly. Furthermore, if the inference time is not a strong requirement, many more models can be built to create an ensemble with higher accuracy.

References

Ahmad, I., Yang, Y., Yue, Y., Ye, C., Hassan, M., Cheng, X., Wu, Y., & Zhang, Y. (2022). Deep Learning Based Detector YOLOv5 for Identifying Insect Pests. *Applied Sciences*, *12*(19), 10167.

Gong, H., Mu, T., Li, Q., Dai, H., Li, C., He, Z., Wang, W., Han, F., Tuniyazi, A., & Li, H. (2022). Swin-Transformer-Enabled YOLOv5 with Attention Mechanism for Small Object Detection on Satellite Images. *Remote Sensing*, *14*(12), 2861.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.

Li, G., Suo, R., Zhao, G., Gao, C., Fu, L., Shi, F., Dhupia, J., Li, R., & Cui, Y. (2022). Real-time detection of kiwifruit flower and bud simultaneously in orchard using YOLOv4 for robotic pollination. *Computers and Electronics in Agriculture*, *193*, 106641.

Li, X., Wang, S., Jiang, B., & Chan, X. (2017). Airplane detection using convolutional neural networks in a coarse-to-fine manner. 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC),

Lin, Y.-C., & Chen, W.-D. (2021). Automatic aircraft detection in very-high-resolution satellite imagery using a YOLOv3-based process. *Journal of Applied Remote Sensing*, *15*(1), 018502.

Lu, S., Wang, B., Wang, H., Chen, L., Linjian, M., & Zhang, X. (2019). A real-time object detection algorithm for video. *Computers & Electrical Engineering*, *77*, 398-408.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition,

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*.

Wu, D., Jiang, S., Zhao, E., Liu, Y., Zhu, H., Wang, W., & Wang, R. (2022). Detection of Camellia oleifera Fruit in Complex Scenes by Using YOLOv7 and Data Augmentation. *Applied Sciences*, *12*(22), 11318.

Yaban, B., Alganci, U., & Sertel, E. (2022). Aircraft detection in very high-resolution satellite images using YOLO-based deep learning methods. *Intercontinental Geoinformation Days*, *4*, 270-273.