A Comparative Study of Various Versions of YOLO Algorithm to Detect Drones

Gayathridevi K[1], Dr. S. Kanmani[2]

Department of Information Technology (Puducherry Technological University), Puducherry, India

E-mail: [1]gayathridevi.k@pec.edu, [2]kanmani@pec.edu

Abstract

Object detection algorithms with various versions of YOLO are compared with parameters like methodology, dataset used, image size, precision, recall, technology used etc. to get a conclusion as which algorithm would be the best and effective for the detection of objects. Nowadays, due to the low price and ease of use, drones can pose a malicious threat. In the field of public security and personal privacy, it is important to deploy drone detection system in restricted areas. This comparative analysis model gives a wide picture of how various object detection algorithms work, and helps in understanding the best algorithm to be used for the detection of drones with highest accuracy and precision.

Keywords: YOLO, object detection, drone detection, precision, recall.

1. Introduction

Object detection has become a research hotspot, due to the broad development prospect and huge commercial value. The newest object recognition algorithm is called YOLO, or You Only Look Once. The business has seen an increase in the number of YOLO models since they were first introduced in 2015. There are several variants of the YOLO algorithm, including YOLOv1, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, and YOLOv7.

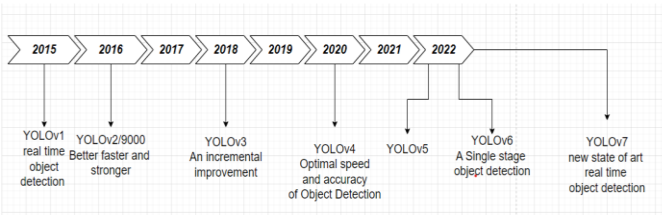

Figure 1.1. YOLO versions

1.1 Architecture of YOLO

YOLO specifically creates the bounding box coordinates and probabilities for each class through direct regression. This greatly increases detection speeds when compared to faster R-CNN. In the training set, each image is split into SxS grids using the YOLO network. When an object's ground truth centre is in a particular grid, that grid is in charge of detecting the object. Each grid also anticipates bounding boxes, conditional probabilities, and confidence ratings for those bounding boxes. [21].

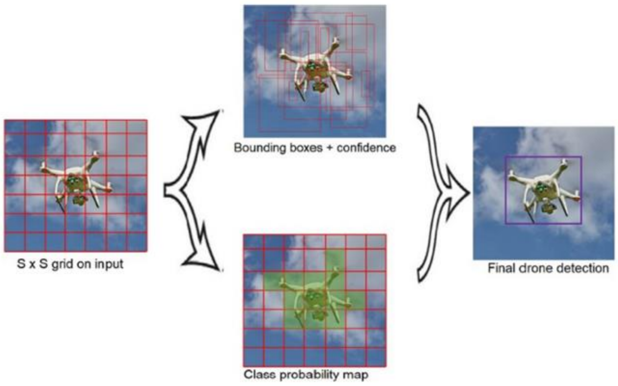

Figure 1.2. YOLO functionality [21]

2. Literature Review

YOLOv1 was released in 2015. You only need to look at an image once to determine what objects are visible and where they are located. This model's limitation is that it can only forecast one of the two objects that enter the same cell. Small objects that emerge in groups pose a challenge for this model [1]. 2016 saw the introduction of YOLOv2. Comparing YOLO to region proposal-based methods, YOLOv2 was improved [3] and faster, and it had a relatively poor recall. A labelled drone collection was needed for training the YOLO drone detector [2].

2018 saw the introduction of YOLOv3. The drone object recognition based on YOLOv3 algorithm for anti-drones was introduced for the first time. Instead of using the last three scales of feature maps, it used the last four scales. YOLOv3 was modified to provide an algorithm for real-time drone detection. YOLOv3 multiscale detection and network structure upgrades were included in the modified version. Training collection for drone detection was implemented [11]. The issue with the YOLOv3 model's excessive number of factors was its limitation. YOLOv3 multi-scale recognition needs to be improved [4].

In 2020, the YOLOv4 algorithm was developed. Compared to YOLOv3, SSD, and YOLOv4, which all belong to one-stage algorithms, the YOLOv4 algorithm has better real-time detection accuracy and speed. There are still deficiencies in accuracy compared with the two-stage method [4].

In [5], the algorithm was trained using datasets from drones and birds. The trained YOLOv4 model was assessed using precision, F1-score, mean Average Precision (mAP), recall and frames per second. Labelling datasets in a specific format was required for various object recognition algorithms, which requires time. Further, when selecting methods, speed was taken into consideration. The research [8] aimed to help develop early detection systems that can provide high security for areas where drones are misused, whether they are low-quality or high-quality images at long or near range. Its main goal was to assist in the development of early detection systems.

Along with studying drone recognition, onboard systems for real-time identification was also researched [9]. YOLO-v3 performed better than SSD and Faster R-CNN in the same testing setting, making it the best algorithm among the three [6].

In order to defend restricted areas, YOLOv5 was introduced. It detects objects applying deep learning architectures. In research [7], attempts were made to include a variety of bird species, sounds, planes, and drones. Additionally, blur drone images were included.

In 2022, the YOLOv6 algorithm was created. The state-of-the-art quantization techniques were studied and utilized [16]. YOLOv7 algorithm was proposed in 2022, with enhanced modules of several trainable bag of freebies method, and re-parameterized modules with extend and compound scaling methods [22].

3. Comparative Study

YOLOv1 uses the image size of 448x448 and a precision value of 86% was obtained when compared to fast RCNN method and tested on PASCAL VOC detection dataset with GoogLeNet network technique with 45 frames per second.

YOLOv2 uses the image size of 416x416 and obtained a precision value of 88.25% and recall of 85.44% tested on Public-domain drone dataset and Customised dataset of 60 YouTube videos. It uses Darknet19 and Darknet-10 network technology with 67 frames per second.

YOLOv3 uses the image size of 416x416, giving the precision value 84.14% and recall of 89.27%, tested on ImageNet dataset and COCO dataset with ResNet and Darknet53 network technology with 45 frames per second.

YOLOv4 uses CSPDarknet 53 (cross stage Partial) network technology with the image size of 416x416 tested on COCO dataset, giving precision value of 89.32% and recall of 92.48% with 38 frames per second.

YOLOv5 uses Darknet 53 network technology with the image size of 224x224 tested on Kaggle dataset giving precision value as 94.7% with 140 frames per second.

YOLOv6 uses python/pytorch for network technology with image size of 640x640 tested on COCO dataset with mAP of 43.3% with 1242 frames per second.

YOLOv7 uses python/pytorch for network technology with image size of 640x640 tested on COCO dataset and 56.8% AP with 161 frames per second.

Table 3.1. Comparative analysis of YOLO algorithms

YOLO Version	Methodology	Dataset	Image size	Precision/ recall
YOLOv1	The input image is divided into a 7 x 7 grid by the system. A grid cell is in charge of detecting an object if its center lies within that grid cell.	PASCAL VOC detection dataset. ImageNet 1000-class competition dataset.	448x448	86% and 54% mAP
YOLOv2	It uses neural network to determine the class probabilities and bounding box in a single round of evaluation using data from the entire image.	Drone dataset in the public realm and customized dataset of 60 YouTube videos	416x416	Precision =88.25% Recall=85.44%
YOLOv3	It clusters the bounding area using k-means. It utilizes the 4 dimensions of feature maps to forecast object bounding boxes.	ImageNet dataset and COCO dataset	416x416 320x320	Precision =84.14% Recall =89.27%
YOLOv4	It contains head, backbone, and neck. Neck is the backbone and output layer. It uses Spatial Pyramid Pooling and Path	COCO dataset	416x416 512x512	Precision =89.32% Recall =92.48%

	Aggregated Network and Feature Pyramid Network			
YOLOv5	It contains CSPNet-Backbone 4 CSPBottleneck, Spatial Pyramid Pooling. Backbone- Feature Extraction Neck- Feature fusion Head/output – object detection.	Kaggle dataset	224x224	Precision =94.7%
YOLOv6	Anchor-free paradigm, SimOTA Tag Assignment Policy, and SIoU bounding box regression loss	COCO	640x640	43.3% AP
YOLOv7	For real time object detector, it uses extended and compound scaling methods.	Kaggle COCO	640x640	56.8% AP

4. Analysis and Result

From the above comparative survey, it can be observed that each YOLO algorithm is more effective than the prior. The performance and precision/recall values of YOLOv7 are higher than the other YOLO versions, and also frames per second value is much more when compared with others. The main advantages of YOLO v7 is its speed. It can process images at a rate of 155 frames per second, which is much faster than other state-of-the-art object detection algorithms. YOLO v7 with 9 anchor boxes, enables to perceive broader shapes and dimensions than the existing algorithms.

With its faster and more robust network design, YOLOv7 is much appropriate to identify the tiny objects by overcoming the limitations of other YOLO versions due to its increased accuracy, detection capability, stability, training and labelling. Additionally, YOLOv7 has a better resolution than the earlier versions. It analyses images at a resolution of 640 by 640 pixels, as opposed to YOLOv3's use of 416 by 416 pixels. YOLOv7 can identify smaller objects with greater accuracy thanks to its higher resolution.

5. Conclusion

Various works conducted on YOLO have been compared by taking into account the parameters like methodology, dataset used, image size, precision, recall, technology, and frames per second, by which it can be concluded that YOLOv7 is much suitable to be used in the detection of drones. The feasibility of detecting drones are much more enhanced using YOLOv7 algorithm, since the limitations in detection of drones are overcome by the exclusive features of YOLOv7 algorithm.

References

[1] Joseph Redmon, Santosh Divvala, Yoss Girshick, Ali Farhadi, “You Only Look Once: Unified, Real-Time Object Detection” arXiv:1506.02640v1 [cs.CV] 8 Jun 2015.

[2] Joseph Redmon, Ali Farhadi “YOLO9000: Better, Faster, Stronger” arXiv:1612.08242v1 [cs.CV] 25 Dec 2016

[3] Hu, Yuanyuan, Xinjian Wu, Guangdi Zheng, and Xiaofei Liu. "Object detection of DRONES for anti-DRONES based on improved YOLO v3." In *2019 Chinese Control Conference (CCC)*, pp. 8386-8390. IEEE, 2019.

[4] Shi, Qingbang, and Jun Li. " Objects detection of DRONES for anti-DRONES based on YOLOv4." In *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT*, pp. 1048-1052. IEEE, 2020.

[5] Singha, Subroto, and Burchan Aydin. "Automated Drone Detection Using YOLOv4." *Drones* 5, no. 3 (2021): 95.

[6] Maske, Shubham Rajabhau. "Micro-DRONES Detection using Mask R-CNN." PhD diss., Dublin, National College of Ireland, 2021.

[7] Al-Qubaydhi, Nader, Abdulrahman Alenezi, Turki Alanazi, Abdulrahman Senyor, Naif Alanezi, Bandar Alotaibi, Munif Alotaibi, Abdul Razaque, Abdelaziz A. Abdelhamid, and Aziz Alotaibi. "Detection of Unauthorized Unmanned Aerial Vehicles Using YOLOv5 and Transfer Learning." *Electronics* 11, no. 17 (2022): 2669.

[8] Naseri, Ahmed, and Nada Hussein M. Ali. "Detection of drones with YOLOv4 deep learning algorithm." *International Journal of Nonlinear Analysis and Applications* 13, no. 2 (2022): 2709-2722.

[9] Dadrass Javan, Farzaneh, Farhad Samadzadegan, Mehrnaz Gholamshahi, and Farnaz Ashatari Mahini. "A Modified YOLOv4 Deep Learning Network for Vision-Based UAV Recognition." *Drones* 6, no. 7 (2022): 160.

[10] Samadzadegan, Farhad, Farzaneh Dadrass Javan, Farnaz Ashtari Mahini, and Mehrnaz Gholamshahi. "Detection and Recognition of Drones Based on a Deep Convolutional Neural Network Using Visible Imagery." *Aerospace* 9, no. 1 (2022): 31.

[11] Alsanad, Hamid R., Amin Z. Sadik, Osman N. Ucan, Muhammad Ilyas, and Oguz Bayat. "YOLO-V3 based real-time drone detection algorithm." *Multimedia Tools and Applications* (2022): 1-14.

[12] Valappil, Najiya K., and Qurban A. Memon. "CNN-SVM based vehicle detection for UAV platform." *International Journal of Hybrid Intelligent Systems* Preprint (2021): 1-12.

[13] Bahri, Haythem, David Krčmařík, and Jan Kočí. "Accurate object detection system on hololens using yolo algorithm." In *2019 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO)*, pp. 219-224. IEEE, 2019.

[14] Sazdić-Jotić, Boban, Ivan Pokrajac, Jovan Bajčetić, Boban Bondžulić, and Danilo Obradović. "Single and multiple drones detection and identification using RF based deep learning algorithm." *Expert Systems with Applications* 187 (2022): 115928.

[15] Pham, Vung, Du Nguyen, and Christopher Donan. "Road Damages Detection and Classification with YOLOv7." *arXiv preprint arXiv:2211.00091* (2022).

[16] Li, Chuyi, Lulu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, Liang Li, Zaidan Ke et al. "YOLOv6: a single-stage object detection framework for industrial applications." *arXiv preprint arXiv:2209.02976* (2022).

[17] Ammar, Adel, Anis Koubaa, Mohanned Ahmed, Abdulrahman Saad, and Bilel Benjdira. "Vehicle detection from aerial images using deep learning: A comparative study." *Electronics* 10, no. 7 (2021): 820.

[18] Xin Wu, Member, Wei Li, Senior Member, Danfeng Hong,Ran Tao, and Qian Du, “Deep Learning for UAV-based Object Detection and Tracking: A Survey” arXiv:2110.12638v1 [cs.CV] 25 Oct 2021

[19] John, Anand, and Divyakant Meva. "A comparative study of various object detection algorithms and performance analysis." *Int J Comput Sci Eng* 8 (2020): 158-163.

[20] Kurdthongmee, Wattanapong. "A comparative study of the effectiveness of using popular DNN object detection algorithms for pith detection in cross-sectional images of parawood." *Heliyon* 6, no. 2 (2020): e03480.

[21] Alsanad, Hamid R., Amin Z. Sadik, Osman N. Ucan, Muhammad Ilyas, and Oguz Bayat. "YOLO-V3 based real-time drone detection algorithm." *Multimedia Tools and Applications* (2022): 1-14. Springer

[22] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." *arXiv preprint arXiv:2207.02696* (2022).