

Enhancing Real-time Object Detection with YOLO Algorithm

Gudala Lavanya¹ and Sagar Dhanraj Pande^{2,*}

^{1,2}School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

Abstract

This paper introduces YOLO, the best approach to object detection. Real-time detection plays a significant role in various domains like video surveillance, computer vision, autonomous driving and the operation of robots. YOLO algorithm has emerged as a well-liked and structured solution for real-time object detection due to its ability to detect items in one operation through the neural network. This research article seeks to lay out an extensive understanding of the defined Yolo algorithm, its architecture, and its impact on real-time object detection. This detection will be identified as a regression problem by frame object detection to spatially separated bounding boxes. Tasks like recognition, detection, localization, or finding widespread applicability in the best real-world scenarios, make object detection a crucial subdivision of computer vision. This algorithm detects objects in real-time using convolutional neural networks (CNN). Overall this research paper serves as a comprehensive guide to understanding the detection of objects in real-time using the You Only Look Once (YOLO) algorithm. By examining architecture, variations, and implementation details the reader can gain an understanding of YOLO's capability.

Keywords: computer vision, image processing, object detection, CNN, Accuracy

Received on 10 October 2023, accepted on 26 November 2023, published on 05 December 2023

Copyright © 2023 G. Lavanya *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetiot.4541

Corresponding author. Email: sagarpande30@gmail.com

1. Introduction

Generally, a person usually stares at an image and will get to know what is there in the image and how it interacts with each other. Real-time object detection refers to the ability to detect and localize objects in a continuous stream of data. The background of real-time object detection stems from the increasing need for efficient and accurate analysis of visual data in real-world scenarios. Detection of objects is a crucial work in many well-known fields and holds immense importance in various domains including navigation of robots, augmented reality, automation, and medical diagnosis [1]. Similarly, object detection plays the same role in detecting and differentiating images by neural networks. In such combination and unpredictable situations, all these detection methods will be built upon a deep learning perspective, those regions are constructed on neural networks (CNN), convolutional Networks as (SPP net), faster R-CNN, Regional based Fully Convolution Networks (R-FCN), fast R-CNN, YOLO algorithm and Feature Pyramid Networks (FPN) show better advantages than

traditional methods. Feature Pyramid Networks (FPN), Networks (R-FCN) and YOLO show more convenience and importance than already existing working methods [2]. You only look once algorithm is the best secured and quick object detection method with high accuracy and best training performance, also been improving since it was introduced, by involving all its yolo variations.

1.1 Unified Detection

The total concept of unified detection in YOLO is based upon dividing the given input source into a grid, and every grid cell estimates some certain bounding boxes in fixed numbers along with respective class probabilities. These bounding boxes are responsible for detecting objects that fall within the grid cell. For processing an entire image Yolo uses a neural network and makes predictions for all objects simultaneously. Human beings interface the indifferent elements of object detection mixing into one network. To predict each bounding box, the Yolo network takes every feature from that complete image [3]. Key components of

YOLO's unified detection approach include Grid Division, Anchor Boxes, Prediction Generation and Non-Maximum Suppression (NMS – eliminates redundant bounding boxes). First, the machine divides the given image which contains the object into S x S grid. The confidence score tells how confident the training model will be confident the box containing the object, with how much accuracy it thinks the box has and the prediction capacity of the box how accurate it is to predict [4]. Figure 1 represents grid division and the cell parameters. And Figure 2 represents image and object classifications of a single object image.

(a) How it works when there are multiple objects.

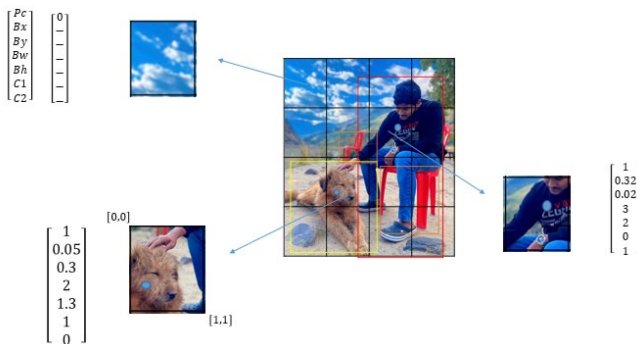


Figure 1. 4x4 by 7 volumes (four by four total grid of 16 cells and every cell of vector of size 7)

(b) Work progress when there is a single object.

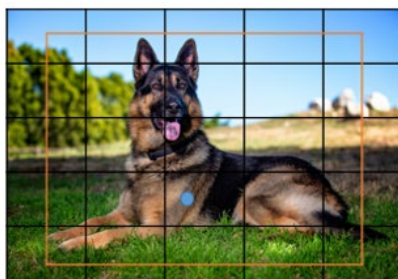
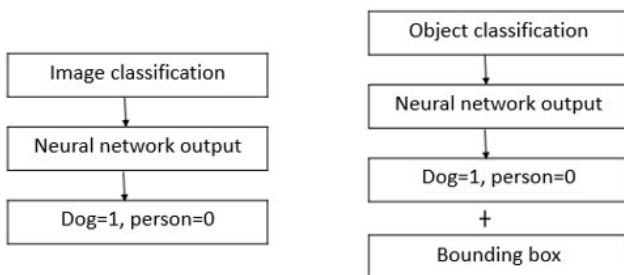


Figure 2. Image of single object and its classification

(c) Brief about 3 main versions and improvements of all versions:

YOLO is a single-stage object detection model [5]. A simple neural network predicts class and bounding box probabilities directly from the images in just one set of evaluations. If once the image is detected by the machine, the Yolo algorithm will start image processing after that it detects objects from the image using respective libraries. It faces errors and difficulties which appear as groups in detecting small objects. Object detection algorithms should not only be accurate in the prediction of object class but also with the location and must be incredibly fast while doing the process of the video processing in real-time demands. YOLO-V2 takes out all the connected layers which are only fully or linear structured and for the prediction of bounding boxes it introduced anchor boxes features like multi-scale training and included higher resolution capacity. In general, already the object detection models having tiny objects were facing the problems of poor performance and low precision. Instead of predicting the coordinates of the bounding box directly from that convolution network, it uses linear connected layers to predict bounding boxes [6]. From figure 3, it solves issues like low performance and precision, a model based upon a deep learning approach which was yolo-v2 having tiny objects, called O-YOLO-v2 (Optimized yolo v2) [7].

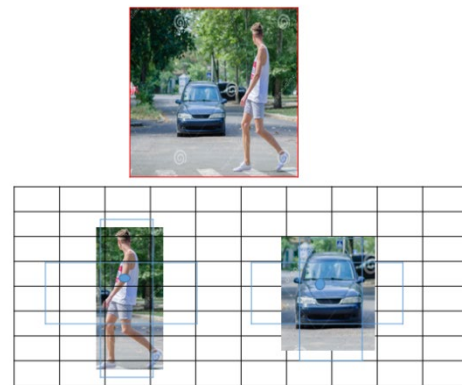


Figure 3. Yolo – v2 design.

To increase the improvement of object detection it targets various measurements mainly focused on smaller objects, and a different multi-sectoral detection algorithm is introduced involving the latest YOLO-V3 [8]. Yolo-V3 boosts excellent performance in a wide range of input or given resolutions YOLOv3 further improved upon its predecessors by introducing a few key enhancements. It utilized a variant of the dark net architecture called Darknet-53, which improved feature extraction capabilities. Yolo-V3 scored 37 mAP in the Test with a given input resolution of 608x608 on the COCO-2017 validation set [9]. The new upgradation of this algorithm with a comparatively smaller archetype size of artificial environments is tiny yolo v3. In real-time performance, the system falls short of detection accuracy on slow computational machines[10]. Figure 4 represents detection of fruits using yolov3 algorithm.



Figure 4. Yolo-v3 detection [11]



Figure 5. Detecting different objects

2. FAST YOLO: A fast you only look once machine for real-time embedded object detection in recordings

The most challenging thing in computer vision is object detection because it involves both image classification which classifies the image and image localization which localizes the image. For achieving maximum object detection output as compared to some other approaches, the deep neural networks (DNNs) were revealed, as all know YOLO version2 is an existing state in DNN-dependent object detection techniques in both terms of accuracy and processing speed. Even though yolov2 can achieve real-time high performance on a powerful graphics processing unit, it remains very objectifying for holding this method for actual time detection of objects in video on devices like embedded systems with only related computer memory or power. Here there is a proposal of a new framework which is Fast YOLO, called framework as fast You Only Look Once which advances yolo version2 to perform detection of objects in running video in a real-time manner on devices that are embedded [12]. This type of a single convolution network simultaneously forecasts numerous bounding boxes and the probability of classes for the boxes [13]. With zero processing of batch on a GPU TitanX this base network will run at 45 fps and the faster version will run at more than 150 frames per second and by this we can process the detection of any streamed videotape in real-time with minimal 25ms of latency [14]. YOLO is globally the best when comes to the image and its process of prediction. Contrastive to the sliding window method and regional-based proposal methods YOLO can observe the entire image in the process of test time and training. So that it completely encodes the provisional info about its appearance and classes. YOLO has the learning capability to recognize the generalized portrayal of objects. As soon as YOLO is trained on general images and tested in real-time work, YOLO outruns best object detection methods like DPM and regions with CNN by an extensive edge [15]. From figure 5, it visualizes the detection of different objects. yolo-SA is an improved version of the one-stage detection model YOLO v4 [16]. Table 1 represents comparative analysis of all yolo versions in the division of Years, improvements, and total no. of layers in the network.

Table 1. Comparison of YOLO version improvements

Yolo Versions	Years	Improvements	Total no. of layers in the network
Yolov2	2017	Included higher resolution and anchor boxes	Contains 5 max-pooling layers which is by dark net and 19 convolutional layers.
Yolov3	2018	Performance on smaller objects	Increased number of layers to 106.
Yolov4	2020	Optimal speed and accuracy	Have 53 convolutional layers with certain sizes.
Yolov5	2020	Introduced mosaic augmentation	It has 80 classes of 3 parts
Yolov6	2021	Architectural improvements	Not fixed
Yolov7	2021	Infers faster and greater accuracy	Not fixed
Yolov8	2023	improved adaptive training, customizable architecture advanced data augmentation	Not fixed

3. Network Architecture

Next to network design of the detection of objects was increasingly trending and has grown broadly, in the Deep Learning generation. So, the network architecture consists of mainly three different layers: The convolution layer, the Max pool, and the Fully Connected layers. Yolo network

has 24 convolution layers which is connected by 2 fully connected layers. As a substitute of the established units worn by Google net, it generally uses 1×1 reductant layers coming after 3×3 convolution layers. With 9 convolution layers instead of 24 layers and some filters in those layers will be used by neural networks of Fast YOLO. By not considering the dimensions of the network, coaching and trailing parameters are so equal between YOLO and Fast YOLO [17]. Fig6 presents the Network architecture of the Yolo algorithm as this model advances for sum squared misconception or error in the desired output. The main advantage of the sum squared error is its ease to optimization even though this doesn't have perfect alignment with the moto of increasing mean accuracy. This loads classified error the same as the localized error which wouldn't be absolute. We can also observe in every image, that not all grid cells contain objects that may be empty. So that those empty grids' confidence scores will be pushed towards zero. Therefore, it affects model instability [18]. Coming for analysis of error in this algorithm in comparison to fast regional CNN proves that localization errors in YOLO have a large signification count. For the regularization model, we use batch or group normalization. For removing dropouts from representation without overfitting for this we use Batch Normalization. Convolutional layers with pre-well-defined bounding boxes with custom widths and heights which are called Anchor boxes. For the prediction of the bounding box coordinates Yolo uses a feature extractor that is fully connected layers which are high in the convolutional layers. But we get some decrease in its accuracy due to the use of anchor boxes. Generally, YOLO can predict 98 boxes for any single image but by using predefined bounding boxes (anchor boxes) it can predict more than thousands of boxes. If Yolo doesn't use anchor boxes, then the basic model assures to get 69.5 mAP with 81% of recall [19]. In addition to anchor boxes, the model assures to get 69.2 mAP with 88% of recall [20]. It implies there is a decrease in MAP and increases in recall percentage so there is a chance to improve the performance of the model. After it increases the loss of predictions from bounding box coordinates then considers boxes that don't contain objects and decreases the dropping from Confidence predictions. In this way we mainly focus on the improvement of localization and recall percentage and to maintain classification accuracy it adds the Batch Normalization method on every convolution layer present in YOLO then we can increase 2% improvement in mAP. Below figure 6, visualizes the network architecture of the Yolo algorithm.

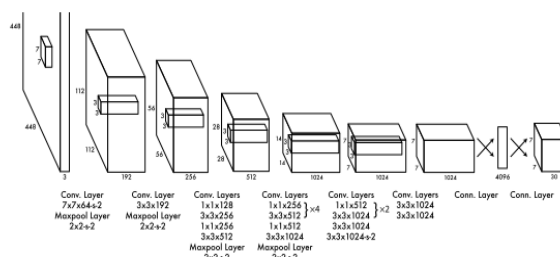


Figure 6. Network architecture of yolo algorithm [21]

4. Literature Survey

For the detection of objects, YOLO requires a neural network which is only one side front propagation. Between two frame images for the frame differencing, it uses pixel-wise differentiation which is background representation for the detection of objects and background subtraction for detecting moving regions by a Gaussian mixture model [22]. The proposal of a backward history model for identifying the object is raised by Stauffer and Grimson [23]. Background subtraction for the detection of moving objects is proposed by Liu et al that is in any image is done by recording the pixel-by-pixel difference between reference and present background pictures [24]. Sungandi et al. have introduced the detection of objects in low-resolution images by using frame differences [25]. The proposal of a new shadow detection in video clippings and untouched background model is done by Jacques et al [26]. Instead of the classification method, yolo got a name like detection of objects as a regression problem by the authors of YOLO [27] which implies that YOLO performed with more accuracy and much faster. Even it can also predict artwork perfectly. Figure 7 represents the flow chart of the process. Ghosh et al. (2023) embarked on a comprehensive study to assess water quality through predictive machine learning. Their research underscored the potential of machine learning models in effectively assessing and classifying water quality. The dataset used for this purpose included parameters like pH, dissolved oxygen, BOD, and TDS. Among the various models they employed, the Random Forest model emerged as the most accurate, achieving a commendable accuracy rate of 78.96%. In contrast, the SVM model lagged behind, registering the lowest accuracy of 68.29%[33]. Alenezi et al. (2021) developed a novel Convolutional Neural Network (CNN) integrated with a block-greedy algorithm to enhance underwater image dehazing. The method addresses color channel attenuation and optimizes local and global pixel values. By employing a unique Markov random field, the approach refines image edges. Performance evaluations, using metrics like UCIQE and UIQM, demonstrated the superiority of this method over existing techniques, resulting in sharper, clearer, and more colorful underwater images.[34].

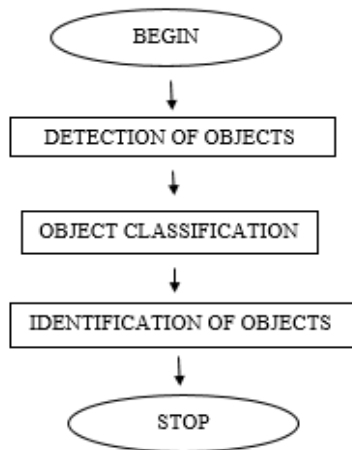


Figure 7. Flow chart of object detection model

5. Yolo object detection algorithm is crucial because of the given reasons:

Speed: As YOLO predict objects in real-time, it improves detection speed. Compared to other algorithms yolo can perform much faster running at 45 frames per second. Another main difference is YOLO has the capability to see complete images at only once which is not present in previous methods [28]. We will run the image on CNN for only one time at run time. All the testing and training parameters are as same as between fast Yolo and YOLO [29].

High accuracy: YOLO has a high prediction capacity that gives the best results with fewer background mistakes. There are some different heuristics to increase yolo accuracies like cosine learning rate scheduler, data augmentation, batch normalization (synchronized) and image mix-up [30]. Having a larger pixel quality improves accuracy but takes off with inference and slow training time. For more accuracy large pixel quality may help the model to detect small objects.

Learning capabilities: yolo has a high learning capability, which allows one to find out the patterns of the objects and apply them in the process of detection of objects. Yolo acquired the object detection by division of an image into N grids, of equal dimensions $S \times S$. Based on the COCO dataset (common objects in context), this algorithm can detect classes of 80 COCO objects: bus, person, car, Bicycle, motorbike, aeroplane, truck, train, boat [31].

High-resolution classifier: Generally, before training, a real YOLO neural network uses 224×224 pixels and then changes to 448×448 P while recognition. In the process of changing from one model to another model which is classification to detection model, the model adjusts to the

classification of the image. YOLO version 2 segments the prior training method into just two steps: train the network with 224×224 Pixels and convert the pixels to 448×448 [32].

Physically based training: For the detection of images with different resolutions then this model enables the same before the network. We know the training speed is fast when the given input is small and when the given input size is high then the speed of training is low. Physically based training can also improve its accuracy so that there will be a good balance between speed and accuracy.

6. Yearly Trends

This section has organization of all the publication data for the purpose of displaying yearly growth of YOLO versions. Table 2 explains the count of educational research papers of all versions of yolo are yolo v1, yolo v2, yolo v3, yolo v4, yolo v5, yolo v6, yolo v7 and yolo v8. This breakdown shows that the publication number of those papers has increased slowly in the 2020 and 2021. Apart from, YOLO V3, YOLO and V2 versions have interested most of the researchers due to its properties, here the time factor comes under separate element. YOLO V5, V6, v7 and V8 versions count is low because both are recent to the trend now so they will improve in future years. Fig 8 represents the graphical view of the mentioned table2. Table 3 represents various comparative analysis of object detection algorithms which includes invention year, novelty of algorithm and recent searches of mentioned algorithms.

Table 2. Yearly trends of publication data

	Yolo v3	Yolov4	Yolov5	Yolov6	Yolov7	Yolov8	total
2017	10	0	0	0	0	0	10
2018	50	19	0	0	0	0	34
2019	48	210	0	0	0	0	258
2020	36	496	81	13	9	0	635
2021	418	734	440	175	23	8	1798
total	529	1459	521	188	32	8	2737

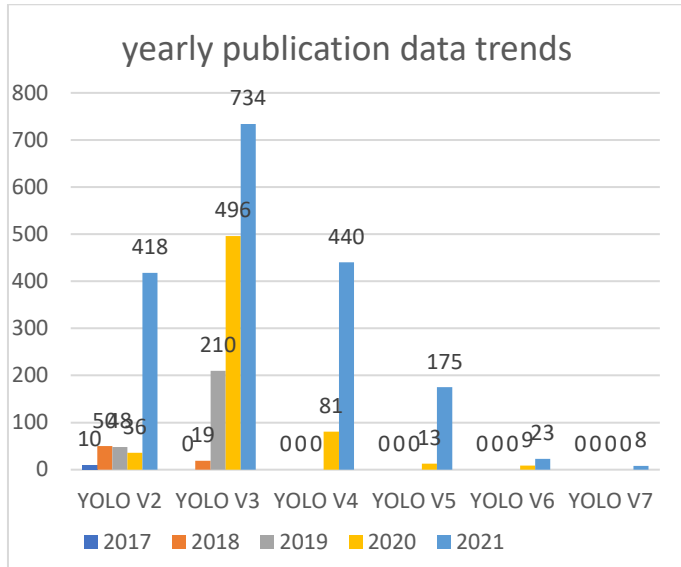


Figure 8. Graphical representation of yearly publication data

Table 3. Comparative analysis of different object detection techniques

Object detection algorithm	Invented in year	The novelty of the algorithm	Topmost Google searches in the below algorithms
1. Region-based Convolution Neural Networks (R-CNN)	2013	Joins proposal of a rectangular region with CNN features.	R-CNN object detection, R-CNN implementation
2. Fast R-CNN	2015	For the creation of a set of regions, it uses a regional method.	Fast R-CNN GitHub, PyTorch
3. Faster R-CNN	2015	It quickly predicts the locations of different objects.	Faster R-CNN Architecture, algorithm
4. Mask R-CNN	2017	It improves the speed and efficiency of object detection	Mask R-CNN image segmentation
5. Region-based Fully Convolutional Network (R-FCN)	2016	It was used to detect and classify stop and yield signs from the images.	R-FCN keras, tensor flow
6. Mesh R-CNN	2019	It predicts object instances in that image and infers their 3D shape.	Mesh R-CNN GitHub and real-time uses
7. Histogram of Oriented Gradients (HOG)	2005	In all the localized portions of image, it will count the occurrences of gradient orientation.	HOG face detection, HOG Matlab
8. Single Shot Detector (SSD)	2016	By using multi box it Detects multiple objects in the given image.	SSD single shot. Multi box detector bibtex
9. You Only Look Once (YOLO)	2015	Increases accuracy in predictions.	YOLO latest versions. And yolo full form

7. Conclusion

In this paper, it is the general view based on YOLO object detection and object classification. Detection of objects is a significant technique in Computer Vision for instance location of objects in images and videos. As we compared to previous classification techniques. It offers several advantages, including real-time processing, simplicity, and effective handling of small objects which performs classification and object detection in one pass through the network in YOLO's unified detection approach. Although there are some issues yet to be solved with some versions of Yolo regarding both larger and smaller objects, it struggles with output to get perfect alignment of objects in the image. Further, these issues are surely going to be rectified and worked with the best outcome. YOLO algorithm is mainly based on a regression model, it predicts all bounding boxes and classes for the entire image in one time of the algorithm, instead of selecting the region of interesting part in an Image, and Object Detection is composed of general tasks such as localization, object classification and segmentation. In the performance, the Yolo algorithm gives its best for detecting objects. We mainly reviewed unified detection, types of Yolo versions, applications based on Yolo, network architecture and comparative analysis. Overall, YOLO-based real-time object detection has revolutionized computer vision applications by providing efficient and accurate solutions.

8. Future Scope

Object detection in real-time is the main ability that is wanted by most robots and computer vision systems. It's making great progress and giving output in many directions because of the early research in this area. It has to be considered that object detection with the Yolo algorithm is not used much in many areas where it could be of great help and this could be improved in future. In fact, YOLO object detection in images has received a lot of observation in the pattern recognition sectors and computer vision in recent years. The future of these mechanisms is in the process of proving and could give freedom from routine jobs which will be done more precisely by systems and machines. Keys areas of future exploration are improved accuracy, handling complex scenes, multi-object tracking and domain-specific object detection.

References

- [1] **Conference:** Redmon, Joseph, et al. "You only look once: Unified, real-time object detection. "Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [2] **Review article:** W. Zhiqiang, L. Jun, A review of object detection based on convolutional neural network, in: 2017 36th Chinese Control Conference (CCC), 2017, pp. 11104– 11109. doi: 10.23919/ChiCC.2017.8029130
- [3] **Journal article:** Arya MC, Rawat A. A review on YOLO (You Look Only One)-an algorithm for real time object detection. J Eng Sci. 2020;11:554-7
- [4] **Review paper:** Arya, Mukesh Chandra, and Anchal Rawat. "A review on YOLO (You Look Only One)-an algorithm for real time object detection." J Eng Sci 11 (2020): 554-7.
- [5] **Conference:** Redmon et al. in You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [6] **Review paper:** Tsang, Sik-Ho. "Review: YOLOv2 & YOLO9000—You Only Look Once (Object Detection (accessed on 24 February 2019) (2019) u Only Look Once: Unified, Real-Time Object Detection
- [7] **Conference:** M. Takahashi, Y. Ji, K. Umeda and A. Moro, "Expandable YOLO: 3D Object Detection from RGB-D Images," 2020 21st International Conference on Research and Education in Mechatronics (REM), 2020, pp. 1-5, doi: 10.1109/REM49740.2020.9313886.
- [8] **Journal article:** Ju, M.; Luo, H.; Wang, Z.; Hui, B.; Chang, Z. The Application of Improved YOLO V3 in Multi-Scale Target Detection. Appl. Sci. 2019, 9, 3775.
- [9] **Journal article:** wei fang 1,2, (member, ieee), lin wang 1 , and peiming ren 1, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments.
- [10] **Journal article:** Yi, Zhang, Shen Yongliang, and Zhang Jun. "An improved tiny-yolov3 pedestrian detection algorithm." Optik 183 (2019): 17-23.FANG 1,2, (Member, IEEE), lin wang 1 , and peiming ren 1, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments.
- [11] **Journal article:** Tian, Yunong, et al. "Apple detection during different growth stages in orchards using the improved YOLO-V3 model." Computers and electronics in agriculture 157 (2019): 417-426
- [12] **Article:** Shafiee, Mohammad Javad, et al. "Fast YOLO: A fast you only look once system for real-time embedded object detection in video." arXiv preprint arXiv:1709.05943 (2017).
- [13] **Journal article:** George, Jose, Shibon Skaria, and V. V. Varun. "Using YOLO based deep learning network for real time detection and localization of lung nodules from low dose CT scans." Medical Imaging 2018: Computer-Aided Diagnosis. Vol. 10575. SPIE, 2018.mad Javad, et al. "Fast YOLO: A fast you only look once system for real-time embedded object detection in video." arXiv preprint arXiv:1709.05943 (2017).

- [14] **Journal article:** Chiang, Holly, Yifan Ge, and Connie Wu. "Multiple Object Recognition with Focusing and Blurring." Lectures from the Course (2016).
- [15] **Conference:** Du, Juan. "Understanding of object detection based on CNN family and YOLO." Journal of Physics: Conference Series. Vol. 1004. No. 1. IOP Publishing, 2018
- [16] **Article:** Joseph Redmon*, Santosh Divvala, Ross Girshick Ali Farhadi* University of Washington*, Allen Institute for AI†, Facebook AI Research, "You Only Look Once: Unified, Real-Time Object Detection"
- [17] **Conference:** Santosh Divvala, Redmon, Joseph, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788. 2016
- [18] **Journal article:** Wong, Alexander, et al. "Yolo nano: a highly compact you only look once convolutional neural network for object detection." 2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMCC2-NIPS). IEEE, 2019
- [19] **Journal article:** Wei H, Kehtarnavaz N (2019) Semi-supervised faster RCNN-based person detection and load classification for far field video surveillance. Mach Learn Knowl Extraction 1(3):756–767
- [20] **Article:** Tsang S-H (2018) Review: Inception-v4 - Evolved From GoogLeNet, Merged with ResNet Idea (Image Classification), towards data science
- [21] **Conference:** Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [22] **Journal article:** H. Deshpande□ , A. Singh, H. Herunde, "Comparative Analysis on YOLO Object Detection with OpenCV"
- [23] **Journal article:** Stauffer, C., & Grimson, W. E. L. (1999, June). Adaptive background mixture models for real-time tracking. Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149) (Vol. 2, pp. 246-252). IEEE
- [24] **Journal article:** Liu, Y., Ai, H., & Xu, G. Y. (2001, September). Moving object detection and tracking based on background subtraction. Proc. SPIE 4554, object detection, classification, and tracking technologies (Vol. 4554, pp. 62-66).
- [25] **Journal article:** Sungandi, B., Kim, H., Tan, J. K., & Ishikawa, S. (2009). Real time tracking and identification of moving persons by using a camera in outdoor environment. International journal of innovative computing, information and control, 5, 1179-1188
- [26] **Journal article:** Jacques, J. C. S., Jung, C. R., & Musse, S. R. (2005, October). Background subtraction and shadow detection in grayscale video sequences. XVIII Brazilian symposium on computer graphics and image processing (SIBGRAP'05) (pp. 189-196). IEEE
- [27] **Conference:** Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In proceedings of the IEEE conference on computer vision and pattern recognition, pp 779-788
- [28] **Article:** Long, Xiang, et al. "PP-YOLO: An effective and efficient implementation of object detector." arXiv preprint arXiv:2007.12099 (2020)
- [29] **Journal article:** Zhang, Zhi, et al. "Bag of freebies for training object detection neural networks." arXiv preprint arXiv:1902.04103 (2019)
- [30] **Conference:** Yin, Xuanyu, et al. "YOLO and K-Means Based 3D Object Detection Method on Image and Point Cloud." The Proceedings of JSME annual Conference on Robotics and Mechatronics (Robomec) 2019. The Japan Society of Mechanical Engineers, 2019
- [31] **Conference:** Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [32] **Journal article:** Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems (pp. 1097-1105).
- [33] Ghosh, H., Tusher, M.A., Rahat, I.S., Khasim, S., Mohanty, S.N. (2023). Water Quality Assessment Through Predictive Machine Learning. In: Intelligent Computing and Networking. IC-ICN 2023. Lecture Notes in Networks and Systems, vol 699. Springer, Singapore. https://doi.org/10.1007/978-981-99-3177-4_6
- [34] Alenezi, F.; Armghan, A.; Mohanty, S.N.; Jhaveri, R.H.; Tiwari, P. Block-Greedy and CNN Based Underwater Image Dehazing for Novel Depth Estimation and Optimal Ambient Light. Water 2021, 13, 3470. <https://doi.org/10.3390/w13233470>