Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

Literature Survey on YOLO Models for Face Recognition in Covid-19 Pandemic

Aseil Nadhum Kadhum[1], Aseel Nadhum Kadhum[2*]

[1,2]Imam Al-Kadhim College, Iraq.*

Email: [1]easeel18@yahoo.com
Corresponding Email: [2]easeel18@gmail.com*

Received: 15 March 2023 **Accepted:** 30 May 2023 **Published:** 29 July 2023

Abstract: Artificial Intelligence and robotics the fields in which there is necessary required object detection algorithms. In this study, YOLO and different versions of YOLO are studied to find out advantages of each model as well as limitations of each model. Even in this study, YOLO version similarities and differences are studied. Improvement in the YOLO (You Only Look Once) as well as CNN (Convolutional Neural Network) is the research study present going on for different object detection. In this paper, each YOLO version model is discussed in detail with advantages, limitations and performance. YOLO updated versions such as YOLO v1, YOLO v2, YOLO v3, YOLO v4, YOLO v5 and YOLO v7 are studied and showed superior performance of YOLO v7 over other versions of YOLO algorithm.

Keywords: YOLO, Object Detection, Bounding Boxes, Neural Network, YOLO V1, YOLO V2, YOLO V3.

1. INTRODUCTION

Object detection is very important nowadays which consists of detecting different objects from input video or input image, things like human, building, animals, car, birds, etc. The goal is to detect Objects that use You Only Look Once (YOLO). This method has several advantages over other detection algorithms such as CNN, YOLO is a single stage deep learning algorithm which uses convolution neural network for object detection. It is popular due to its speed and accuracy. There are various deep learning algorithms, but they are unable to detect an object in a single run but YOLO, on the other hand, makes the detection in a single forward propagation through a neural network which makes it suitable for real-time application. This property has made YOLO algorithm popular among the other deep learning algorithms [1].The algorithm will not look at the image completely but in YOLO the algorithm looks the image completely by predicting the bounding boxes using convolutional

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

network and the class probabilities for these boxes and detects the image faster as compared to other algorithms. In real time our algorithm process 45 frames per second. YOLO algorithm makes localization errors but predicts less false positives in the background [2]. You Only Look Once: Unified, Real-Time Object Detection, by Joseph Redmon. Their prior work is on detecting objects using a regression algorithm. To get high accuracy and good predictions they have proposed YOLO algorithm in this paper, Finally, YOLO learns very general representations of objects. It outperforms other detection methods, including DPM and R-CNN, when generalizing from natural images to other domains like artwork [3][4].

1.1 What is YOLO for Object Detection?

'You Only Look Once' in short called as YOLO, this technique which works in real time for finding object class and its location in either from image or video input. YOLO algorithm uses neural network concept for object detection. Detecting object location and for detected objects applying bounding boxes. As the name suggest 'You Only Look Once', complete image objects are identified in single run.

YOLO (You Only Look Once) is the novel algorithm for object detection in real time environment using neural networks. YOLO algorithms are very popular as it provides higher accuracy and less time for object detection. Different applications of YOLO include animal detection, parking slot detection, people detection, traffic signal detection, etc. In YOLO algorithm, neural network is single forward propagation for accurate object detection. It has been used in various applications to detect traffic signals, people, parking meters, and animals [5].

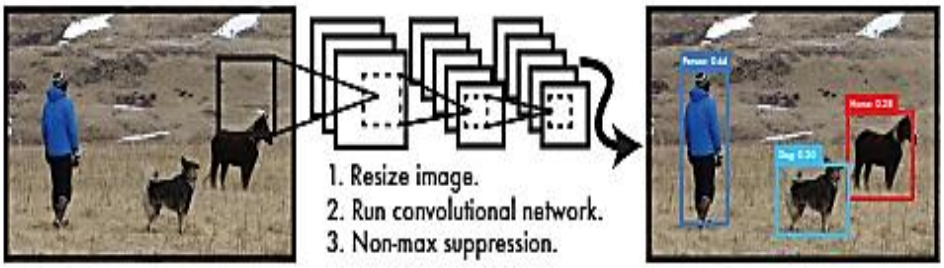

Fig.1.1 Object Detection System using YOLO [6]

In above figure, probability of class is predicted by using single network and detected objects are shown by bounding boxes. There are many benefits of YOLO model over state of art methods used for detection of objects. YOLO module [6] does follow three operations as,

a) Image resizing
b) Run single network for image
c) Using models confidence, non-max suppression

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

Object detection from images is straightforward as well as simple by using YOLO algorithm. YOLO is considered as regression model for finding objects in image or video using a single network. As there is single network, its performance can be analysed using prediction output and can be modified very easily to get superior performance. False positive prediction rate of YOLO model is very less compared to state of art techniques for object detection.
Object detection in real time scenario is challenging task as existing techniques has multiple limitations which can be overcome using YOLO model as it is more accurate as well as fast even in highly complex task.

1.2 Face Detection and Recognition by YOLO Models

Deep learning algorithms plays important role in face detection and recognition as deep learning can recognise specific pattern from massive amount of pattern data. Deep leaning can be used for different types of data classification such as images, music and text data. Combination of object pattern finding and localisation is nothing but object detection. Face pattern finding and localization is nothing but face detection. In pattern recognition, face pattern finding is one of the most challenging tasks. Deep learning uses mainly CNN (Convolutional Neural Network) for semantic segmentation and pattern recognition [7].
Face detection may include face clustering, recognition of face pattern, verification of face, etc. Traditional deep learning approaches not giving better performance for face detection in real time environment. So, to overcome the limitations of traditional deep learning approaches YOLO model is invented which mainly focus on taking only suitable values for consideration and finetuning required performance metrics. YOLO models also solved the performance in real time environment for face detection as well as solved the complex training models by reducing time and complexity of training for face detection [7].

1.3 Advantages of YOLO Algorithm [5]

There are following advantages of YOLO algorithm as below,

A. Capability of Learning Object Representation

YOLO model has superior learning ability of object representation and using this representation, it can easily detect the objects in real time scenario. Single network can be used for detection of object class as well as location of an object.

B. Higher Accuracy

YOLO algorithm gives higher accuracy in real time scenario of object detection. Objects with small size can be detected more precisely using YOLO models.

C. Higher Speed of Object Detection in Real Time Scenario

Object detection with very less time, is the main advantage of YOLO algorithm. Smaller objects such as traffic light, etc are detected with very high speed.

D. Single Network

Object classes are detected and their single location can be found using single network. So, prediction accuracy can be improved by only modifying the single network.

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

E. Generalized Model

YOLO is most generalized model as it gives better accuracy of object detection even in artworks or in natural/original domain image.

F. Frame Rate

Each frame is processed with the frame rate of, for smaller network 150 fps while larger network with 45 fps which shows that frame rate is very much suitable for real time applications.

1.4 Limitations of YOLO algorithm

Prediction limitations can be imposed due to strong spatial constraints. Small objects can be predicted using YOLO which has many challenges as there may be possibility of nearby objects has very less distance like bird's flocks, having very less distance between birds. Error in large bounding box as well as error in small bounding box are treated same. Small error in small box is may not be observable but small error in small bounding box may affect object detection much [6].

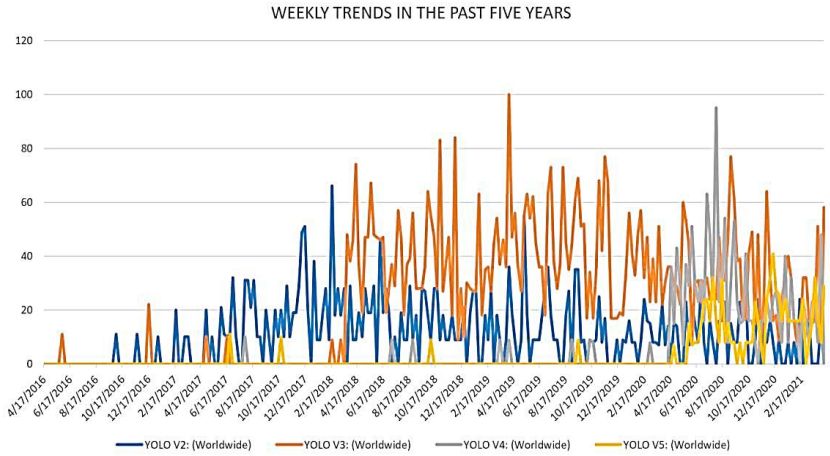

Fig. Trend of YOLO models in last few years [8]

The author [8], compared different versions of YOLO, implementation and design details. First author discussed overall idea based on version and then by using public data author analysed data results in tabular format and plot format. YOLO model is famous because of its quick object detection and smaller size of model. Regression concepts are used in YOLO for object detection. There are three basic types of objects based on size, large, medium and small. So, design of generalized model which identifies all the three types of objects is required. YOLO model can be used for such detection using multiple layers with different sampling values. Main structure of YOLO has in all total 26 layers, among 26 layers 2 layers are FC (Fully Connected) layers and remaining 24 layers are convolution layers. YOLO model has two main limitations as recall rate is lower and positioning is inaccurate [8].

Localization and classification tasks are included in every object detection technique. YOLO is based on single network-based architecture. There are many existing state of art techniques

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

such as RCNN (Region-based Convolutional Neural Networks), Faster-RCNN, FPN (Feature Pyramid Networks), which has many limitations and less accuracy [9].

YOLO v2 is invented by the two researchers A. Farhadi and J. Redmon in the year 2016. It is showed that older object detection techniques such as SSD and R-CNN has lower performance than proposed YOLO v2 model for object detection. Architecture of YOLO v3 is used by YOLO 9000 for more than 9000 classes detection. In YOLO v2 model, inventor has solved problems of multiple parameters improving speed of YOLO v2 model. Training of this model is separated in two stages so its easy for training the model [10].

Existing face detection uses Viola-Jones algorithm detect face using Haar features and AdaBoost technique. For improving performance many DMP (Deformable Part) models have been used. Deep learning takes the features from training model and new test data features are classified with trained deep model features. Deep learning-based feature extraction method YOLO v3 found better performance of face detection compared to state of art techniques such as R-CNN and old YOLO models [11].

Human life had threatened in the COVID-19 pandemic because of coronavirus disease. Many researchers and doctors had worked on different technologies to provide solution for timely treatment and to provide good facilities in less time. To block the virus chain, wearing mask is suggested by medical doctors. medical organizations and government organisations. COVID-19 transmission is reduced by alerting people and warning them for wearing mask. In public, people are following social distancing or no, and everyone in the public is using mask or no, made automatic using different machine and deep learning models. A framework is designed which is fine grained and hierarchical model for face mask detection [12].

Face detection has many applications such as attendance management, unauthorized access detection, face mask detection, etc. Deep learning techniques plays major role in face detection and recognition. For face detection many state-of-art machine learning techniques were used which require lot of handcrafted features in training the model. Computational time as well as power required is very high providing lower accuracy. To improve accuracy and to get detection in real time environment YOLO model which is part of deep learning algorithms is used [7].

Literature Survey

Inaccurate positioning as well as lower recall rate problems can be overcome using YOLO v2 model [8]. YOLO v2 is very simple network which do not consider broaden or deepen network.

Improvements updated in YOLO model is as below [8],

Sr No	YOLO Version	Techniques Used in the Specific YOLO Version	Invented Month/Year	Limitations
1	YOLOv1	Confidence loss and detection is mainly based on grid division	2015 by Joseph	Smaller object detection is

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

			Redmon	challenging (objects in group)
2	YOLO v2	Training is of two stages, K-means concept added	2016 by A. Farhadi and J. Redmon	Whole image training is required
3	YOLO v3	Object detection using feature pyramid network (FPN)	2018 by Joseph Redmon	It is very fast than other algorithms but not accuracy is limited
4	YOLO v4	Slightly different strategy (SDT), Mosaic enhancement of data.	April 2020 by Alexey Bochkovskiy, etAl.	Not an adaptive Model, it can be applied in specific scenes
5	YOLO v5	Optimize the input data, the size of the model can be adaptive, this is the model that was invented for lightweight, training is very easy, multiple export options are available, very accurate and faster, and it is one of the YOLO models	2020 by a company named 'Ultralytics'	Slight Slower inference speed [14]
6	YOLOv6	Redesigned network backbone and neck to EfficientRep Backbone and Rep-PAN Neck. Network head is decoupled separating different features from the final head.	2022 by chuyi et al.	Improvement The detecting small objects, anchor free training of model. Less stable and flexible as compared to YOLOv5.
7	YOLOv7	YOLOv7 surpasses all known object detectors in both speed and accuracy in the range from 5 FPS to 160 FPS and has the highest accuracy 56.8% AP among all known real-time object detectors Layer aggregation using E-ELAN, trainable bag of freebies, 35% fewer network parameters. Model scaling for concatenation-based model	2022 (Wanget al.,)	Increase in speed and accuracy ,ease of training and inference[17]

Fastest object detection with very simple structure is presented by YOLO-v1 model and it is applicable for real time environment-based object detection. Even YOLO-v1 model found many limitations too like one class is available to each grid, even problems related generalization is present in YOLO-v1 model for object detection [9].

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

YOLO-v2 model has small size which gives poor performance and there is requirement of huge parameters to run an object detection model. In this model, k-means concepts get added to improve the performance of model. YOLO v1 limitations was identified and improved the model using YOLO v2. YOLO v2 model predicts bounding boxes using anchor boxes, even this model has advantages of classifier with higher resolution, batch normalization and backbone as Darknet-19[10].

YOLO v3 is invented by Redmon in 2018. YOLO v3 model can be used for face detection in complicated environment as in complicated environment faces are either blurred or blocked. To find suitable priori box, dataset is re-clustered in proposed YOLO v3 model. While prediction different images with few test images get different scores, optimal score is considered for finding nearby matches. Object detection in YOLO v3 is using feature pyramid network (FPN) for better performance. YOLO v3 has many advantages in complex environment over state of art deep learning or machine learning models [11].

YOLO v4 model gives better performance compared to older models and performance is shown using precision, recall, etc parameters [12]. This model uses SDT (Slightly Different Strategy) and mosaic enhancement of original data before training. It is invented by Alexey in April 2020. YOLO v4 is having mAP (mean average precision) 0.599 while mAP for YOLO v3 is 0.574. mAP of YOLO v4 give better performance over YOLO v3, faster RCNN with baseline, faster RCNN with FPN. Performance of YOLO v4 under different conditions like face detection, face mask detection, mask weared incorrect are considered in this approach [12].

Worldwide 5,049,374 lives were taken by SARS-CoV-19 in pandemic. Many safety measures were given to people in different countries to reduce the spread of disease. The key requirement of reduction in cases of covid-19 is social distancing and wearing face mask. Manual detection of face mask and social distancing in public places and warning the people who are not following rules is challenging task. YOLO v5 is designed for detecting weather a person wearing mask or no in public using CCTV cameras [13][14][15]. YOLO v5 based object detection is compound scaled. YOLO v5 has simple functionality, hyper parameter tuning, etc.

YOLOv6 model is a cutting-edge object detector that offers remarkable balance between speed and accuracy, making it a popular choice for real-time applications. This model introduces several notable enhancements on its architecture and training scheme, including the implementation of a Bi-directional Concatenation (BiC) module, an anchor-aided training (AAT) strategy, and an improved backbone and neck design for state-of-the-art accuracy on the COCO dataset [18].

YOLOv7 is a state-of-the-art real-time object detector that surpasses all known object detectors in both speed and accuracy in the range from 5 FPS to 160 FPS. It has the highest accuracy (56.8% AP) among all known real-time object detectors with 30 FPS or higher on GPU V100. Moreover, YOLOv7 outperforms other object detectors such as YOLOR,

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

YOLOX, Scaled-YOLOv4, YOLOv5, and many others in speed and accuracy. The model is trained on the MS COCO dataset from scratch without using any other datasets or pre-trained weights. Source code for YOLOv7 is available on GitHub [19][20].

2. CONCLUSION

There have been many models for object detection invented by different researchers but the main goal of YOLO invention is to achieve higher accuracy in a short time for multi object detection. The YOLO model has many advantages as discussed in the paper, so many applications use the YOLO model to detect objects with higher accuracy. In this study, different YOLO algorithms such as YOLO -v1, YOLO-v2, YOLO-v3, YOLO-v4, YOLO-v5 and YOLOv7 were studied. Each model has some advantages and limitations that are discussed in detail in this paper. But the YOLO v7 model is more accurate and requires less time to detect objects. YOLO v7 has an adaptive feel designed even for a lightweight model. found the training process very easy in YOLO v7 compared to other YOLO models for object detection. Therefore, YOLO v7 can be used to recognize faces by considering different faces as different objects.

3. REFERENCES

1. Nepal, U., & Eslamiat, H. (2022). Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. Sensors, 22(2), 464.
2. Srijanani, K., Rajeswari, S. K., & Murthy, S. V. B. REAL TIME OBJECT DETECTION USING YOLO ALIGORITHM.
3. 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).
4. Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).
5. 'Basics of YOLO' : https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/
6. Joseph Redmon, etAl.(2022), 'You Only Look Once: Unified, Real-Time Object Detection', http://pjreddie.com/yolo/
7. Dweepna Garg, etAl.(2018), 'A Deep Learning Approach for Face Detection using YOLO'
8. Peiyuan Jiang,etAl.(2021) , 'A Review of Yolo Algorithm Developments' ScienceDirect Available online at www.sciencedirect.com Procedia Computer Science 199 (2022) 1066–1073
9. YOLO v1 explanation, 'https://medium.com/mlearning-ai/object-detection-explained-yolo-v1-fb4bcd3d87a1'.
10. Rui Li (2018) etAl. 'Improved YOLOv2 Object Detection Model', School of Computer and Communication, LanZhou University of Technology, Lanzhou, China.

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 04, June-July 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.34.27.35

11. Lin Zheng Chun, etAl.(2018), 'YOLOv3: Face Detection in Complex Environments', https://www.researchgate.net/publication/343694111_YOLOv3_Face_Detection_in_Complex_Environments/link/5f3a7a31a6fdcccc43cfd5ad/download
12. PeishuWu, etAl., (2022) 'FMD-Yolo: An efficient face mask detection method for COVID-19 prevention and control in public', https://www.sciencedirect.com/science/article/pii/S0262885621002468
13. Inderpreet Singh Walia ,etAl (December 2021), 'An Integrated Approach for Monitoring Social Distancing and Face Mask Detection Using Stacked ResNet-50 and YOLOv5', Electronics 2021, 10(23), 2996; https://doi.org/10.3390/electronics10232996
14. Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs'
15. Hooman Samani, etAl. (Aug2021), 'Social and Safety Monitoring for Pandemic with YOLO'
16. Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7464-7475).
17. Olorunshola, O. E., Irhebhude, M. E., & Evwiekpaefe, A. E. (2023). A Comparative Study of YOLOv5 and YOLOv7 Object Detection Algorithms. Journal of Computing and Social Informatics, 2(1), 1-12.
18. https://docs.ultralytics.com/models/yolov6/
19. 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.
20. Zhao, H., Zhang, H., & Zhao, Y. (2023). Yolov7-sea: Object detection of maritime uav images based on improved yolov7. In Proceedings of the IEEE/CVF winter conference on applications of computer vision (pp. 233-238).
21. 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.
22. https://www.researchgate.net/publication/334070398_A_Deep_Learning_Approach_for_Face_Detection_using_YOLO.
23. https://www.researchgate.net/publication/356161868_Social_and_Safety_Monitoring_for_Pandemic_with_YOLO/link/618e345761f098772087e8af

The Author(s) 2023.This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)