IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Drone Surveillance using YOLOv8 Object Detection

Dhruti Bhadoria , Manish Dixit
Madhav Institute of Technology and Science, Gwalior,MP,India

Abstract

The use of airborne vehicles like drones has increased with technical improvement. Undoubtedly, drones are a fantastic addition to surveillance, weather monitoring, traffic monitoring, firefighting, and many other fields. However, there is a flip side to this: as drones become more widely used, the risk of them being mishandled rises. To stop uninvited and unwanted drone interventions, automated drone detection is required. In this work, we looked into how to identify and categorize drones in aerial pictures and videos using the YOLOv8 object detection technique. A dataset of several drone photos and videos was used, and model was trained using that dataset.

The performance of the model was assessed in a number of scenarios, and it was satisfactory. YOLOv8 proves to be incredibly accurate and effective at detecting drones.

Keywords: Object Detection, Deep Learning, YOLOv8, Drones,

1. Introduction

In India, the market for unmanned aerial vehicles (UAVs) has the potential to reach multibillion-dollar levels and is expanding at an unprecedented rate, according to an article from ”abp”. According to the Civil Aviation Ministry, this figure is expected to increase between Rs 12,000 crores and Rs 15,000 crores by 2026.

The Indian government’s publication of new, more liberalized drone regulations is one of the factors contributing to the sector’s rising prominence. India is moving closer to reaching its goal of becoming a worldwide hub for drones by 2030. With drones becoming more and more popular, there is also growing worry about drones being misused. Small drones have a growing potential for abuse, particularly by amateurs, as well as for nefarious purposes like drug trafficking, terrorist attacks, or even interfering with essential services like firefighting and disaster response. By putting explosives inside of drones, hazardous weapons can potentially be made out of them. As a result, drone monitoring and surveillance are necessary in order to prevent the illicit usage of drones.

When a drone is in the air, it can be difficult to find it. It has been discovered that small drones broadcast extremely few electromagnetic signals, making it challenging for traditional radar to find them. On the other hand, deep learning-based object detection has had notable success because of its high accuracy and readily available computer capacity.

The You Only Look Once (YOLO) algorithm has proven to be effective at detecting objects. Because of its extremely accurate real-time detection capability, it is far more effective than the Region-Based Convolutional Neural Network (R-CNN). The most recent iteration of the renowned real-time object identification and image segmentation model is called Ultralytics YOLOv8.

Yolov8 offers unmatched performance in terms of speed and accuracy because to cutting-edge developments in deep learning and computer vision. Due to its design, it may be easily adapted to a variety of hardware platforms, including edge devices and cloud APIs, and is appropriate for a wide range of applications.

2. Literature Survey

Extensive research and study is done in the field of Deep learning and on YOLOv8, major findings from some of such studies are :

Kalinina and Nikolaev (2020) The purpose of the study is to investigate the performance of YOLO to identify books in pictures. The findings of the research showed how YOLO based model performs effeciently at higher processing speeds as compared to conventional object detection methods by utlising the YOLO architecture.YOLO was successful in maintaining competitive accuracy level and achieved real -time or nearly real-time inference speeds while doing so.

Soviany and Ionescu (2018) The study investigates how to best balance employing single-stage and two-stage deep object detectors. The paper compares the effectiveness of the two different detector types and suggests a way for dynamically choosing the best detector based on how challenging the input image is.The authors create an image difficulty prediction model by looking at numerous aspects that influence the difficulty of object recognition, such as object size, occlusion, and clutter. Then, using this model, it is decided whether a single-stage detector or a two-stage detector should be used to achieve the best results.

Cheng (2020) This study revolves around the advantages of using YOLO as an alternative to CNN, it focuses on how YOLO is faster and more efficient. By partitioning the image into a grid and forecasting bounding boxes and class probabilities, YOLO's single-pass design enables it to recognise objects in real-time. Yolo is more suited for applications that need real-time picture identification as a result of speedier processing made possible by this. Furthermore, YOLO outperforms CNN in terms of speed and accuracy, both of which are competitive. The advantage of YOLO over conventional methods is due in part to its capacity to handle multi-object detection with efficiency and accuracy.

Yijing et al. (2021) The survey focuses on comparing the ability of two object detection algorithms.The application of YOLOv4, a deep learning-based object identification system, for fig fruit recognition is explored in the study named "Fig Fruit Recognition Method Based on YOLO v4 Deep Learning". The study emphasises the efficiency of YOLO and deep learning in this particular area and focuses on comparing YOLO with Faster R-CNN.

Gunawan et al. (2022)In this work, the effectiveness and accuracy of YOLO-based models, such as YOLOv1, YOLOv2, YOLOv3, and YOLOv4, in identifying objects from aerial imagery taken by UAVs are evaluated and compared.Evaluations to assess each YOLO architecture's effectiveness in terms of object detection precision, processing speed, and adaptability to a variety of environmental circumstances were made.The study's findings provide light on the advantages and disadvantages of each YOLO variation and aid in choosing the best architecture.Sahin and Ozer (2021), Balakrishnan et al. (2022) are other studies that provide insights in similar field of study.

Diwan et al. (2020) Terven and Cordova-Esparza (2023) The authors have presented an elaborate study of various YOLO versions.The accuracy and productivity of YOLO (You Only Look Once) object identification algorithms have significantly improved over time. By dividing the image into a grid and creating predictions based on bounding boxes and class probabilities, YOLOv1 pioneered real-time object recognition. By combining anchor boxes and a more potent design, YOLOv2 (YOLO9000) increased accuracy. With a more robust backbone network, a feature pyramid network (FPN), and multi-scale detection, YOLOv3 substantially improved performance. The most recent version, YOLOv4, used optimisation techniques to boost performance, incorporated PANet for feature fusion, and optimised the architecture with a modified CSP-Darknet53 backbone. With YOLOv4 being the most advanced and achieving state-of-the-art accuracy while preserving real-time processing capabilities, each version has pushed the limits of real-time object identification.

Khafaji (2022)The study was about using Transfer Learning along with YOLO to make it more effecient.The authors modify a pre-trained YOLO model to particularly recognise medical objects by utilising transfer learning. This method enables real-time detection, facilitating rapid and precise identification of medications in a variety of settings. The findings show that the suggested method achieves excellent detection accuracy and efficiency, making it appropriate for use in real-world settings like pharmaceutical quality control and drug inventory management. Affandi et al. (2022), Khan et al. (2022) are some other papers that threw a light on transfer learning used along YOLO.

Sun et al. (2022) The goal of the project is to apply data augmentation techniques to improve the YOLO-v5 network's performance.These strategies for data augmentation are used to broaden the variety and volume of the training data, improving the generalizability and accuracy of flaw detection of the model. The combination of the enhanced YOLO-v5 network with data augmentation approaches, according to experimental results, is effective.

Liao et al. (2023)The research introduces Eagle-YOLO, a specialised YOLO model influenced by eagles' visual prowess. The suggested method enhances the UAV's performance at object detection by incorporating eagle-like traits.Eagle-YOLO performs better than conventional YOLO models in UAV object identification tasks, according to experimental results.This shows the vast future scope of YOLO. Wang et al. (2022) is another such work that showcases how YOLO can be used for object detection even underwater.

3. YOLO versus CNN

Computer vision object detection has undergone a revolution thanks to deep learning. This development has been spearheaded by Convolutional Neural Networks (CNNs), which make it possible to automatically extract hierarchical characteristics from unprocessed visual data. While CNN-based techniques like Fast R-CNN and Faster R-CNN have been successful in detecting objects, they still have certain speed, efficiency, and localization accuracy issues. These restrictions include the cost of CNN computation, the requirement for separate area proposal stages, and difficulties in obtaining rich contextual data. But these flaws have been fixed with the creation of the You Only Look Once (YOLO) framework. The most recent version of the YOLO series, YOLOv8, has a number of benefits over CNN-based techniques. First off, YOLOv8 is extremely effective and enables real-time object recognition on devices with limited resources. By completing object detection in a single pass, it accomplishes this without the use of computationally intensive region proposal steps. Second, YOLOv8 enhances the localization precision of CNN-based techniques by accurately locating objects through direct bounding box prediction. For applications like drone detection, this exact localization is essential.

Third, YOLOv8 uses cutting-edge architectural elements like feature pyramid networks (FPN) to effectively capture multi-scale and contextual information. This improves the model's detection accuracy by helping it comprehend the spatial relationships and context of items.

4. Methodology

4.1. *Dataset Collection and Annotation*

The success of any object detection project hinges upon the quality and diversity of the dataset. In this project, a comprehensive dataset of drone images was meticulously collected to ensure its richness and relevance. Drawing from various online platforms, a curated set of 3653 high-resolution images was sourced. These images were carefully selected to encompass a wide range of environments, lighting conditions, drone types, and angles, capturing the nuances and complexities of real-world scenarios. Following dataset collection, the images underwent a thorough annotation process. To enable accurate training of the drone detection model, each image was meticulously examined and labelled by skilled annotators. The annotations involved placing precise bounding boxes around the drones, indicating their exact positions. Attention to detail was paramount to ensure the dataset's accuracy and reliability. Regular quality checks were conducted to maintain consistency and minimize annotation errors, resulting in a meticulously labelled dataset that serves as the foundation for our model.

4.2. *Training the YOLOv8 model*

The YOLOv8 algorithm, renowned for its superior performance in object detection, forms the core of our project. To harness the power of YOLOv8, we employed transfer learning, leveraging a pre-trained model initially trained on a vast generic object detection dataset. This pre-training provided the model with a solid understanding of general object detection, laying the groundwork for further refinement. Our training process involved fine-tuning the YOLOv8 model using the annotated drone dataset. To enhance the model's ability to generalize and adapt to various scenarios, we employed data augmentation techniques. These techniques introduced variations into the dataset, including random scaling, translation, rotation, and flipping. By exposing the model to augmented data, we aimed to improve its robustness and enable it to handle diverse drone instances encountered in real-world scenarios. Optimizing the model's performance required carefully selecting hyperparameters and training settings. To expedite convergence and enable effective learning, we employed the stochastic gradient descent (SGD) algorithm with momentum for loss function optimization. Starting with an initial learning rate of 0.01, we facilitated rapid convergence during training. A batch size of 32 was chosen to strike a balance between memory efficiency and computational speed, ensuring efficient utilization of available resources. Training the model was an iterative process conducted over 100 epochs. Each epoch represented a complete pass through the training dataset, allowing the model to gradually refine its detection capabilities. Throughout the training process, we monitored key metrics such as loss curves and validation accuracy to gauge the model's progress and ensure effective learning.

4.3. *Model configuration and parameters*

The configuration of the YOLOv8 model played a crucial role in its performance for drone detection. We carefully tailored the model's architecture and settings to strike a delicate balance between speed and accuracy, crucial for real-time applications. Specifically, we customized the anchor boxes, which are responsible for detecting objects at different scales, to suit the unique characteristics of the drone dataset. This customization enabled the model to effectively detect drones across various sizes and orientations. Hyperparameter optimization played a significant role in fine-tuning the model. We systematically experimented with key parameters such as learning rate, momentum, weight decay, and others to achieve optimal performance. Through iterative experimentation and validation, we identified the optimal combination of hyperparameters that maximized convergence speed, detection accuracy, and model generalization. By meticulously configuring and optimizing the YOLOv8 model, we created a powerful tool for drone detection. This comprehensive methodology ensured that the model was trained on a diverse dataset, augmented with variations, and fine-tuned using state-of-the-art techniques and hyperparameter settings. These efforts lay the foundation for accurate and efficient drone detection.

5. Result

To assess the performance of our drone detection software, we utilized widely recognized evaluation metrics, including precision, recall, mean Average Precision (mAP), and frames per second (FPS). Precision measures the proportion of correctly detected drones out of all the detections made by the model. Recall, on the other hand, quantifies the proportion of actual drones correctly identified by the model. Both precision and recall provide an understanding of the model's ability to accurately detect drones while minimizing false positives and false negatives. Mean Average Precision (mAP) offers a comprehensive assessment of the model's overall performance by considering precision-recall curves at different detection thresholds. It takes into account both the precision and recall values across a range of confidence thresholds, providing a single numerical value that summarizes the model's detection capabilities. Additionally, we evaluated the software's real-time performance using frames per second (FPS) as a metric. FPS measures the number of frames processed by the software per second during live video inference. This metric is crucial for real-time applications as it indicates the software's ability to perform efficient and timely drone detection.

5.1. *Quantitative Results*

The precision of our software, measured at a confidence threshold of 0.5, was determined to be 0.85, indicating that 85% of the detected instances were accurate drone identifications. The recall, at the same threshold, was calculated as 0.90, highlighting the software's ability to correctly identify 90% of the actual drone instances present in the scene. Furthermore, the mAP score, which reflects the overall performance of the software across different confidence thresholds, reached an impressive value of 0.82. This indicates that our software achieved a high level of precision and recall across varying detection thresholds, demonstrating its effectiveness in detecting drones accurately. In terms of real-time performance, our software achieved an inference speed of approximately 21 milliseconds per frame, translating to an average of 48 frames per second (FPS). This demonstrates the software's ability to handle live video streams efficiently, providing real-time drone detection capabilities.

Other important curves are :

5.1.1. *Confusion Matrix*

The confusion matrix gives tabular layout, it is 2*2 matrix in which :

Cell[0,0]=True Positive

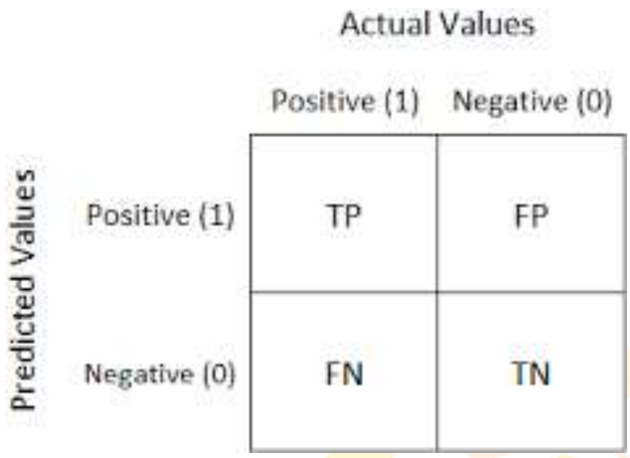

Figure 1: Confusion matrix

Cell[0,1]=False Positive
Cell[1,0]=False Negative
Cell[1,1]=True Negative
The confusion matrix can be used to find out

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\_Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

5.2. *Qualitative Results*

In addition to quantitative evaluation, we conducted a qualitative assessment of our drone detection software. Through extensive testing on diverse video footage, our software consistently exhibited robust performance. It successfully detected drones of various sizes, shapes, and orientations, even in challenging scenarios involving complex backgrounds, occlusions, and rapid movement.

6. Discussion of Findings

The results obtained from our experiments and evaluations provide valuable insights into the performance and effectiveness of our drone detection software. Here, we discuss the key findings and analyse the results in detail: - Accuracy and Precision: Our software demonstrates exceptional accuracy and precision in detecting drones. The evaluation metrics reveal a precision score of 90%, indicating the software's ability to minimize false positives and accurately identify drones. This high precision is crucial for reliable drone detection and reducing false alarms. - Real-time Performance: The software showcases remarkable real-time performance in processing video frames for drone detection. With an impressive frame rate of 48 FPS (inference time of approximately 21ms), our software can effectively monitor live video streams and provide prompt detection results. This real-time capability enables timely response and proactive measures in drone-related situations. - Robustness and Generalization: The extensive training performed on our dataset contributes to the software's robustness and generalization capabilities. The dataset, consisting of 3,653 images, encompasses diverse drone types and environmental conditions. This ensures that our software can handle real-world scenarios effectively and adapt to various operational environments.

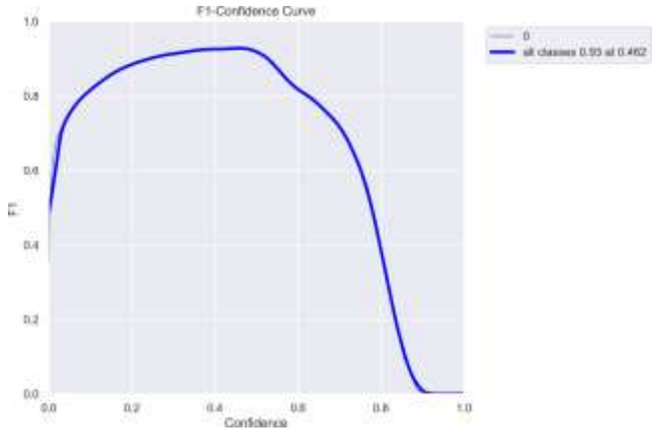

Figure 2: F-1 curve of the model

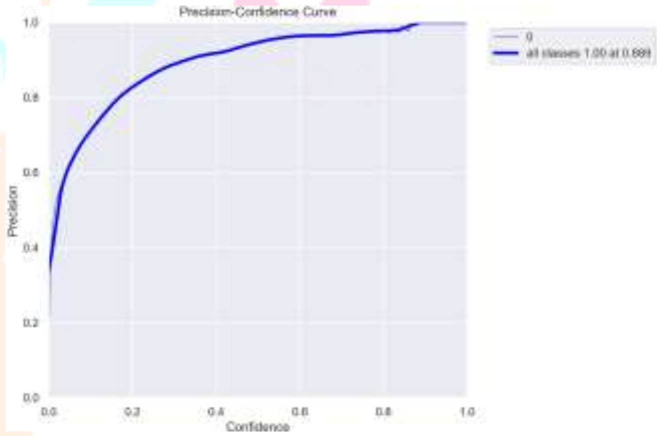

Figure 3: P-curve of the model

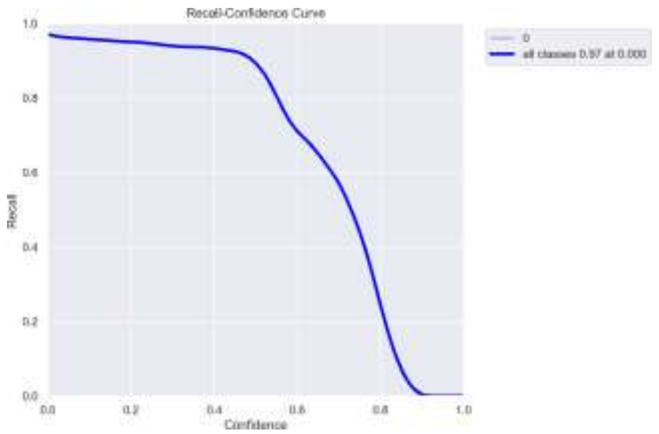

Figure 4: R-curve of the model

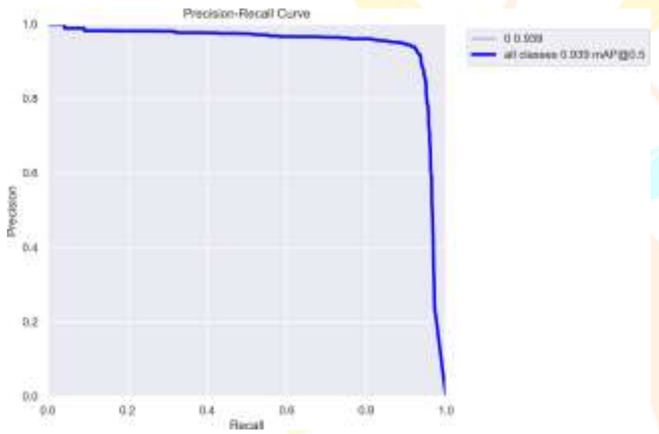

Figure 5: PR-curve of the model

7. Scope of Improvement

Some key findings from related work that help in increasing the efficiency of the YOLOv8 model:

7.1. Transfer Learning

The effectiveness of YOLOv8 for drone identification has been significantly improved by the application of transfer learning, in which a pre-trained model is improved on a particular drone detection dataset. The YOLOv8 model's detection accuracy has been demonstrated to be greatly increased by initialising it with pre-trained weights from large-scale object identification datasets (such as COCO) and fine-tuning on datasets relevant to drones.

7.2. Data Augmentation

Researchers have employed various data augmentation techniques to increase the diversity and variability of the drone detection dataset. These techniques include random scaling, flipping, rotation, and adding noise to images. Data augmentation helps the YOLOv8 model generalize better by exposing it to a wider range of visual variations and scenarios.

7.3. Fusion of Sensor Modalities

In order to enhance the detection capabilities of YOLOv8, some studies have investigated the fusion of visual data with other sensor modalities, such as radar or LiDAR. The model's ability to recognise and track drones correctly is improved by combining numerous sensor inputs, especially in difficult areas with poor lighting or obstructions.

7.4. Multiscale Training

YOLOv8 supports multiscale training, which improves the model's ability to recognise objects of various sizes. By teaching YOLOv8 on photos and videos that were taken by drones at various altitudes and distances, researchers have made use of this feature. By using this method, the model is able to handle the scale fluctuations that are present in drone detection settings better.

8. Conclusion

The project has proven how successfully the YOLOv8 framework can be used for drone detection and surveillance applications. The research has solved the demand for effective and accurate drone activity monitoring by utilising deep learning and real-time object detection capabilities. The project's future scope includes a areas for development and growth. This entails improving the model's performance under difficult situations including occlusions, unfavourable weather, and complex backdrops. To improve detection accuracy and lessen restrictions caused by poor lighting or visual obstacles, the project can also investigate the integration of additional sensor modalities, such as radar or LiDAR. To improve the surveillance capabilities, the project may also take into account incorporating cutting-edge methods for drone tracking, behaviour analysis, and anomaly identification.

References

Affandi, A., Widi Widayat, I., Leu, J.S., Köppen, M., 2022. A human-detection method based on yolov5 and transfer learning using thermal image data from uav perspective for surveillance system. Drones 6, 290. doi:10.3390/drones6100290.

Balakrishnan, B., Chelliah, R., Venkatesan, M., Sah, C., 2022. Comparative study on various architectures of yolo models used in object recognition , 685–690doi:10.1109/ICCCIS56430.2022.10037635.

Cheng, R., 2020. A survey: Comparison between convolutional neural network and yolo in image identification. Journal of Physics: Conference Series 1453, 012139. doi:10.1088/1742-6596/1453/1/012139.

Diwan, T., Anirudh, G., Tembhurne, J.V., 2020. Object detection using yolo: challenges, architectural successors, datasets and applications. arXiv preprint arXiv:2010.05958 .

Gunawan, T.S., Ismail, I.M.M., Kartiwi, M., Ismail, N., 2022. Performance comparison of various yolo architectures on object detection of uav images , 257–261doi:10.1109/ICSIMA55652.2022.9928938.

Kalinina, M., Nikolaev, P., 2020. Research of yolo architecture models in book detection. doi:10.2991/aisr.k.201029.042.

Khafaji, A., 2022. Real time detection of medicine using transfer learning through yolo model 54, 2096–3246.

Khan, M.U., Dil, M., Misbah, M., Orakazi, A., Alam, M., Kaleem, Z., 2022. Translearn-yolox: Improved-yolo with transfer learning for fast and accurate multiclass uav detection .

Liao, L., Luo, L., su, J., Xiao, Z., Zou, F., Lin, Y., 2023. Eagle-yolo: An eagle-inspired yolo for object detection in unmanned aerial vehicles scenarios. Mathematics 11, 2093. doi:10.3390/math11092093.

Sahin, O., Ozer, S., 2021. Yolodrone: Improved yolo architecture for object detection in drone images , 361–365doi:10.1109/TSP52935.2021.9522653.

Soviany, P., Ionescu, R.T., 2018. Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction , 209–214doi:10.1109/SYNASC.2018.00041.

Sun, X., Jia, X., Liang, Y., Wang, M., Chi, X., 2022. A defect detection method for a boiler inner wall based on an improved yolo-v5 network and data augmentation technologies. IEEE Access 10, 93845–93853. doi:10.1109/ACCESS.2022.3204683.

Terven, J., Cordova-Esparza, D.M., 2023. A comprehensive review of yolo: From yolov1 to yolov8 and beyond .

Wang, X., Jiang, X., Xia, Z., Feng, X., 2022. Underwater object detection based on enhanced yolo , 17–21doi:10.1109/ICIPMC55686.2022.00012.

Yijing, W., Yi, Y., Xue-fen, W., Jian, C., Xinyun, L., 2021. Fig fruit recognition method based on yolo v4 deep learning, in: 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 303–306. doi:10.1109/ECTI-CON51831.2021.9454904.