International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.552

e ISSN: 2584-2854
Volume: 02
Issue:12 December 2024
Page No: 3718-3724

Robust YOLO Approach for Mango Fruit Detection Using Computer Vision

Vidhya N P[1], Priya R[2]

[1]Assistant Professor, Sri. C. Achutha Menon Government College, Kuttanellur, Thrissur, University of Calicut, Kerala, India.

[2]Assistant Professor, Government College Nedumangad, Thiruvananthapuram, University of Kerala, Kerala, India.

Emails: *npvidu@rediffmail.com[1], priyanil2007@gmail.com[2]*

Abstract

In precision agriculture, on-tree fruit detection is a solution for automating many tasks which governs yield estimation, harvesting and quality monitoring. Among all tropical fruits, mangoes are the most widely consumed fruit and is a member of the cashew family. Mangoes have been cultivated in South and Southeast Asia for thousands of years. There are several cultivars of mango, which vary in size, shape, sweetness, skin colour and flesh colour. Mango is India's most important commercial fruit crop, accounting for over 54% of all mangoes produced globally. This paper is a study conducted to detect the fruit using a simple non-destructive YOLOv7 deep architecture, a highly accurate object detection method with less error rate. A robust dataset, MangoNet of mango images which are annotated using labelImg tool, is used in the study. The detection of on-tree mangoes will be beneficial for the yield estimation of fruits. The YOLOv7 deep model is employed, incorporating transfer learning to enhance model efficiency and accuracy. The performance of the model is evaluated using the metrics mean average precision (mAP) and Intersection over Union (IoU). 99.5% accuracy is met in mAP@0.5. The results show the feasibility of on-tree mango detection with high precision and recall for automated agricultural systems. Python programming language with pyTorch library is used for the transfer learning. This work highlights the potential of YOLO approach to optimize the mango farming practices and contribute to the smart agriculture technologies.

Keywords: *Computer Vision, Mango Net, Labeling, YOLOv7, Mean Average Precision (MAP).*

1. Introduction

Mango is a tropical fruit produced by the mango (Mangifera indica Linn.), originated in Indo-Burmese region and belonging to the Anacardiaceae family. The fruits are in round, oval or kidney-shaped ranging from 5-25 centimeters in length and from 140 grams to 2 kilograms in weight per individual fruit [1]. Mangi Ferin has potent antioxidant, anti-lipid peroxidation, immunomodulation, wound healing, cardiotonic, hypotensive, antidegenerative, and antidiabetic properties. It is a polyphenolic antioxidant and glucosyl xanthone [2]. It is mostly produced in tropical climate in developing nations with an estimated global production is 15.06 million tonnes [3]. For more than 4000 years, Ayurvedic and indigenous medical systems have valued the mango fruit. Ayurveda attributes various medicinal properties to various parts of the mango tree. Among all tropical fruits, mangoes are the most widely consumed. Mango pulp's chemical makeup varies depending on the variety, maturity stage, and cultivation location. A soft, edible, ripe fruit with desirable qualities develops as a result of a series of physiological, biochemical, and organoleptic changes that occur during the fruit-ripening process [4]. Mangoes are prized for their hardiness and capacity to thrive in a variety of conditions in addition to their fruits. Mango trees have a long lifespan—they can live to be 100 years old and still produce fruit. It is highly popular in market and has gradually developed into the second most cultivated and consumed tropical fruit in the world [5]. Henceforth, production of mangoes and their usage

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.552
e ISSN: 2584-2854
Volume: 02
Issue:12 December 2024
Page No: 3718-3724

highly decides the commercial and economic growth of a nation. To estimate the yield of mangoes, it is highly beneficial for farmers, if detection of mangoes on-tree and during harvesting could be done automatically. Using cutting-edge computer vision algorithms to analyse images is a key component of deep learning techniques, which can be employed to reach the objective. A review by Athanasios et al. highlights the dominance of deep learning methods, including Convolutional Neural Networks, Deep Boltzmann Machines, Deep Belief Networks, and Stacked Denoising Autoencoders, over traditional machine learning techniques in computer vision applications [6]. In the context of computer vision applications, J. Chai et al. identified eight emerging deep learning techniques: AlexNet, VGGNet, GoogLeNet& Inception, ResNet, DenseNet, MobileNets, Efficient Net, and RegNet [7]. The analysis emphasizes the performance of these techniques in each task, categorizing recent developments into three stages and outlining future research directions in terms of both applications and techniques. L. Yuzhi et al. utilized the bibliometric software Cite Space to perform visualization analysis on literature within the core database of Web of Science [8]. I. Guillermo et al. presented a comprehensive overview of Generative Adversarial Networks (GANs), encompassing the latest architectures, optimizations of loss functions, validation metrics, and application domains of widely acknowledged variants [9]. Ball pepper plant leaf bacterial spot disease was identified by M. P. Mathew et al. using YOLOv5 with mAP 90.7%. [10]. A. Sarda et al. demonstrated the application of YOLO to categorize road objects into distinct groups, achieving a mean Average Precision (mAP) of 74.6% in their study [11]. G. Dai et al. introduced a model for detecting and grading sprouted potatoes, achieving an accuracy of 90.14% and a mean Average Precision (mAP) of 88.1% through enhanced YOLOv5 model training [13]. In efforts to enhance productivity, I. Ahmad et al. devised a model employing the YOLOv5 algorithm for classifying and identifying crop-damaging insect pests. Furthermore, they introduced a smartphone-based automatic system with an impressive accuracy of 98.3% [12]. J. Ma et al. employed an improved YOLOv5 model for lotus seed pod detection, revealing a 0.7% accuracy improvement compared to the traditional YOLOv5s model [14]. W. S. Qureshi et al. employed K-nearest neighbour pixel classification and support vector machine classification to assess the quantity of fruits in images depicting mango trees [15]. O. E. Apolo-Apolo et al. applied deep learning techniques to compute the yield and size of citrus fruits using a UAV, with a standard error of 7.22% [16]. H. H. C. Nguyen et al. successfully delved into a portion of deep learning algorithms, uncovering both their strengths and weaknesses. Through this exploration, they acquired knowledge in deep learning and constructed a model capable of recognizing fruits from images [17]. A. Koirala et al. reported that deep learning models excel in fruit-on-plant detection compared to pixel-wise segmentation techniques involving traditional machine learning, shallower CNNs, and neural networks [18]. M. Horea et al. explored deep learning algorithms to recognize fruits from images, leading to a comprehensive understanding of their strengths and weaknesses. The team developed a software capable of fruit recognition with accuracy 96.3% [19]. J. P. Vasconez et al. evaluated the performance of two widely used architectures, Faster R-CNN with Inception V2 and Single Shot Multibox Detector (SSD) with MobileNet, for fruit detection. The testing involved three types of fruits—Hass avocado and lemon from Chile, and apples from California, USA— across diverse field conditions. The results indicated that the system achieved high fruit counting accuracy, with Faster R-CNN and Inception V2 reaching up to 93% overall for all fruits, and SSD with MobileNet achieving 90% overall accuracy for all fruits [20]. I. Sa et al. presented a new method for fruit detection employing deep convolutional neural networks. They introduced multi-modal Faster R-CNN model showcased state-of-the-art performance, specifically excelling in sweet pepper detection. The model demonstrated an improvement in the F1 score, increasing from 0.807 to 0.838 compared to prior methods, highlighting enhanced precision and recall capabilities [21]. N. Hani et al. attained high yield estimation accuracies

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.552
e ISSN: 2584-2854
Volume: 02
Issue:12 December 2024
Page No: 3718-3724

ranging from 95.56% to 97.83% in apple orchards. Notably, the fruit detection results revealed that the semi-supervised method, relying on Gaussian Mixture Models, outperformed the deep learning-based approach across all datasets [22]. M. Afonso et al. reported results in their study on detecting tomatoes in greenhouse images using the Mask RCNN algorithm. This algorithm not only identifies objects but also outlines the corresponding pixels for each detected object [23]. N. Mamdouh et al. introduced a framework that demonstrated notable performance metrics, including a precision of 0.84, a recall of 0.97, an F1-score of 0.9, and a mean Average Precision of 96.68% [24]. In the study conducted by H. Mirhaji et al., the YOLO-V4 model demonstrated superior performance for orange detection over test images, achieving precision, recall, F1-score, and mean Average Precision (mAP) of 91.23%, 92.8%, 92%, and 90.8%, respectively [25]. A. I. B. Parico et al. achieved a remarkable Average Precision (AP) at an intersection over union (IoU) of 0.50, with an impressive value of 98%, designating the YOLOv4-CSP model as the optimal choice in terms of accuracy of real time pear fruit detection and counting [26]. Y. Ge et al. computed the bounding box error by comparing the predicted bounding box with the one generated by the object detection network. This information was then used to update the parameters of the Kalman filter, and the process was iterated to achieve accurate tracking of both tomato fruits and flowers [27]. Literature reveals that YOLO architecture has been employed in different domains and these models show good accuracy. Hence, in this work it is proposed to create a model for region-based fruit detection of on-tree mangoes using the deep learning YOLOv7 methodology. This study is conducted to detect mangoes using a simple non-destructive YOLOv7 deep learning technique, a highly accurate object detection method with less error rate. YOLOv7 surpasses all known object detectors in both speed and accuracy [28]. Figure 1 depicts the proposed methodology of detection of mango fruits.

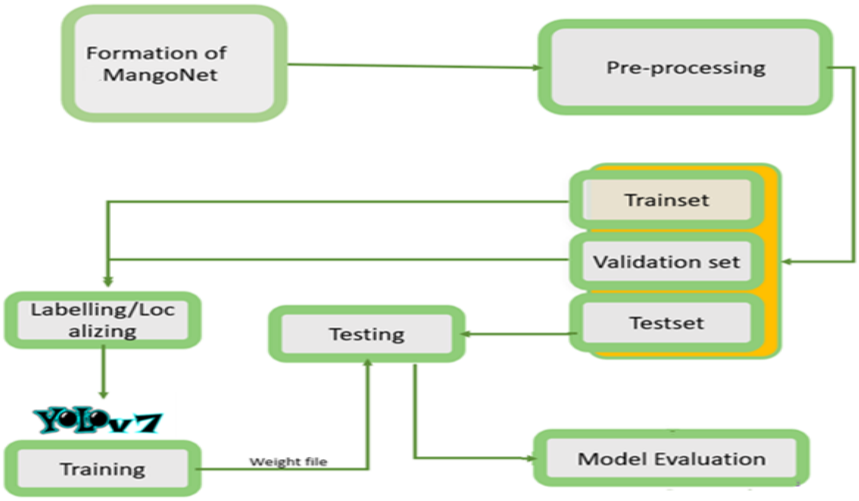

Figure 1 Overview of the Proposed Method

2. Dataset

The images of on-tree mangoes are manually collected using mobile camera. The dataset comprises of a diverse selection of mango varieties, at various angles of capture to ensure the model's ability to generalize and recognize different characteristics across mango variations. The meticulous inclusion of only healthy mangoes in the dataset enhances the model's focus on identifying optimal features for fruit detection. The deliberate variation in lighting, angles, and distances within the dataset contributes to a more comprehensive training environment. The dataset has undergone pre-processing methods, comprising both image size equalization and data augmentation. This strategic

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.552
e ISSN: 2584-2854
Volume: 02
Issue:12 December 2024
Page No: 3718-3724

approach introduces variability into the training data, aiming to enhance the model's generalization. This created dataset is named as MangoNet. Precise annotations are performed on a total of 473 images using the labelImg tool, which facilitates defining bounding boxes and also outlining the corresponding pixels for each fruit. Each image underwent manual labelling to specify the location and boundaries of every mango fruit. This meticulous annotation process serves as the ground truth for training the deep learning model, providing crucial reference points for accurate detection and evaluation. Figure 2 shows the labelled image of the fruit. The same annotation process is applied to all images in the dataset before training.

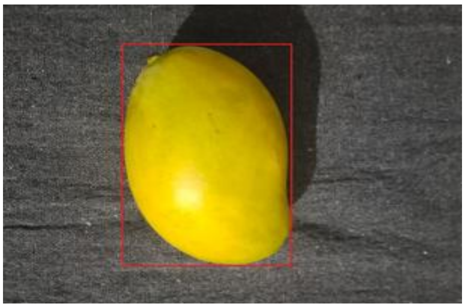

Figure 2 Localizing Fruit in an Image

The dataset is randomly partitioned into an 80% training set, 10% validation set, and 10% testing set.

3. YOLOv7 In Mango Net

The primary objective of this study is detecting mango fruit, a single class, accurately and with high computational efficiency. Convolutional Neural Networks are extensively utilized for object detection tasks, leveraging their efficacy in capturing spatial hierarchies in images. A notable and robust architecture specifically crafted for object detection, the YOLOv7 model is fine-tuned on the annotated MangoNet using transfer learning. The model learns to localize the mango fruits and simultaneously determines both the coordinates of bounding boxes and class probabilities on the input images, utilizing the PyTorch library in python. The YOLOv7 deep learning pretrained model is used to perform inference on new images of fruits. The model's output includes a bounding box outlining the fruit's location, accompanied by a label and the associated probability for the class. A sample batch of labelled fruits during training is depicted in Figure 3. The fruits in the images are labelled with the class name, represented by '0'.

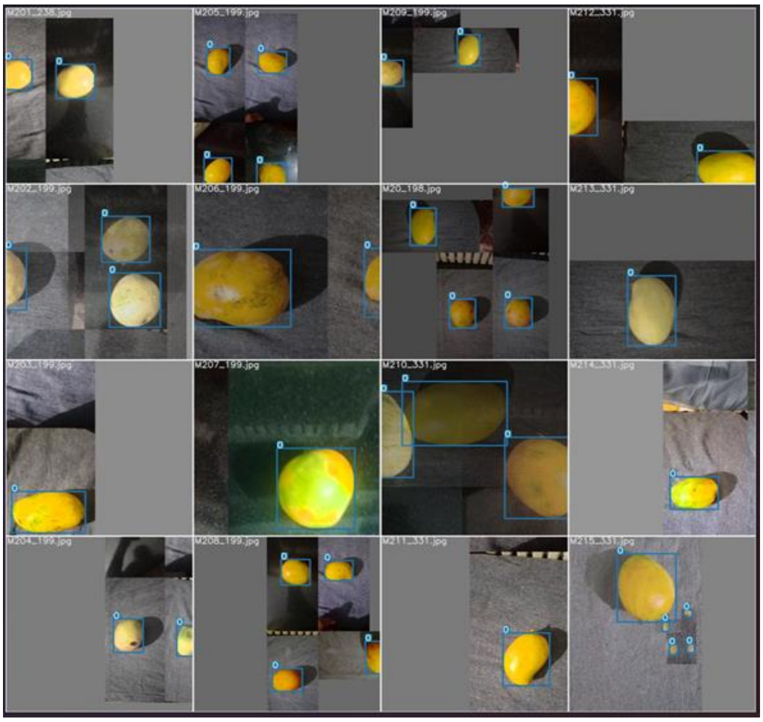

Figure 3 Sample of Train Batch

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.552
e ISSN: 2584-2854
Volume: 02
Issue:12 December 2024
Page No: 3718-3724

Following training, the model's performance is assessed using the validation set, and subsequently, it is tested using the testing set. The sample validation batch is depicted in Figure 4. The fruits in the images are labelled with the class name, 'Mango'.

Figure 4 **Sample of Validation batch**

4. Results and Discussions

The YOLOv7 pretrained model is applied on training and validation set for fruit detection with learning rate of 0.001, image size of 640x640 and batch size of 16. Python programming language is utilized for implementing the transfer learning. The model is trained for 100 epochs and achieved an accuracy of 99.5%. The mAP@0.5 is the average precision of the fruit detection when IoU as 0.5 and in mAP0.5:0.95, IoU value ranges from 0.5 to 0.95. The accuracy of mAP@0.5 is 99.5% and that of mAP@0.5-0.95 is 81.9% in the training period. Figure 5 and 6 shows the graphs which are generated during the training period.

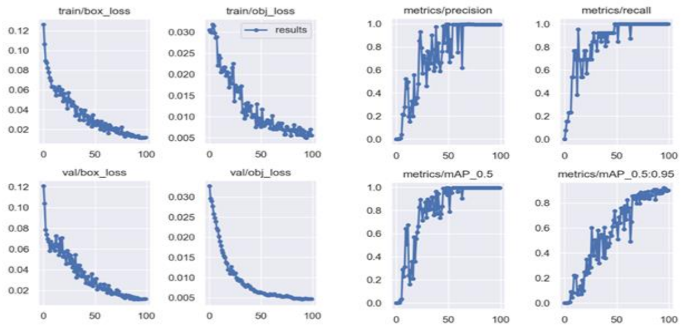

Figure 5 **Plot of Loss Function Convergence, Precision, Recall, And Map After Training the Model**

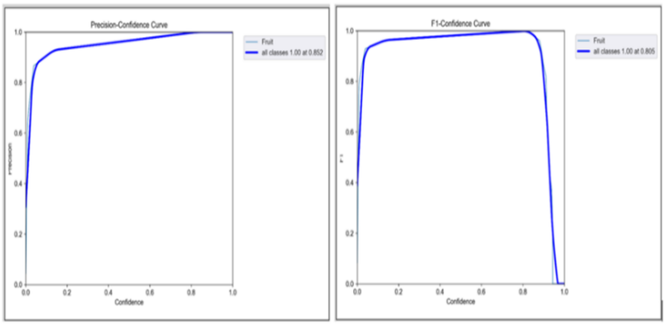

Figure 6 **Precision-Confidence and F-Confidence Curves**

The model's performance is evaluated on a separate testing dataset. The model is fine-tuned to improve its accuracy and generalization to new data. The fine-tuned trained model is deployed to analyse new images. The model detects and localizes mangoes on the tree and after harvesting, generates bounding boxes and probability scores for each mango. Figure 7 shows the sample results obtained with fruit name and confidence score.

Figure 7 **Testing Images with Bounding Box and Confidence Score**

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.552
e ISSN: 2584-2854
Volume: 02
Issue:12 December 2024
Page No: 3718-3724

Conclusion

This paper discusses the on-tree mango detection using YOLOv7 pretrained model precisely. Compared to conventional techniques, using deep learning for mango detection offers a more automated and effective solution. Its use in agricultural environments enables crop management support, harvesting process optimization, and fruit yield monitoring. It is important to remember that the model's performance depends on the variety of the training data in addition to the careful adjustment of hyperparameters throughout the training phase. In the long run, the accuracy of the model should be greatly improved by applying background removal techniques to the images. The results demonstrate that computer vision and deep learning can precisely detecting the fruit.

References

[1]. Morton, Julia Frances (1987). Mango. In: Fruits of Warm Climates. NewCROP, New Crop Resource Online Program, Center for New Crops & Plant Products, Purdue University. pp. 221–239. ISBN 978-0-9610184-1-2.

[2]. Shah, K. A., Patel, M. B., Patel, R. J., and Parmar, P. K (2010). Mangifera indica (mango). Pharmacogn Rev.4(7):42-8.

[3]. Lizada, C (1993). Mango. Biochemistry of Fruit Ripening. Springer, Dordrecht, pages 255-271.

[4]. Tharanathan, R. N., Yashoda, H. M., and Prabha, T. N (2006). Mango (Mangifera indica L.), “The King of Fruits” – An Overview, Food Review International, Volume 22, Pages 95-123.

[5]. Liu, B.; Xin, Q.; Zhang, M.; Chen, J.; Lu, Q.; Zhou, X.; Li, X.; Zhang, W.; Feng, W.; Pei, H.; et al. Research Progress on Mango Post-Harvest Ripening Physiology and the Regulatory Technologies. Foods 2023, 12, 173. https://doi.org/10.3390/foods12010173.

[6]. Athanasios, V., Nikolaos, D., Anastasios, D., and Eftychios, P (2018). Deep Learning for Computer Vision: A Brief Review, Computational Intelligence and Neuroscience, Hindawi, 1687 5265.

[7]. Chai, J., Hao, Z.,Anming, L., and Eric, N. W. T (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios, Machine Learning with Applications.

[8]. Yuzhi, L (2021). Research progress and trend analysis of computer vision based on cite space, Journal of Physics: Conference Series.

[9]. Guillermo, I., Edgar, T., and Alberto, D (2023). A survey on GANs for computer vision: Recent research, analysis and taxonomy, Computer Science Review, Volume 48.

[10]. Mathew, M. P., and Mahesh, T. Y (2022). Leaf-based disease detection in bell pepper plant using YOLOv5, Signal Image Video Process. 16, 841–847.

[11]. Sarda, A., Dixit, S., and Bhan, A (2021). Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm, In Proceedings of the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 4–6; pp. 1370–1374.

[12]. Dai, G.,Hu, L., Fan, J., Yan, S., and Li, R (2022). A Deep Learning-Based Object Detection Scheme by Improving YOLOv5 for Sprouted Potatoes Datasets, IEEE Access. 10, 85416–85428.

[13]. Ahmad, I., Yang, Y., Yue, Y., Ye, C., Hassan, M., Cheng, X., Wu, Y., and Zhang, Y (2022). Deep Learning Based Detector YOLOv5 for Identifying Insect Pests, Appl. Sci. 12, 10167.

[14]. Ma, J., Lu, A., Chen, C., Ma, X., and Ma, Q (2023). YOLOv5-lotus an efficient object detection method for lotus seedpod in a natural environment, Computers and Electronics in Agriculture, Volume 206.

[15]. Qureshi, W. S., Payne, A., Walsh, K. B., Linker, R., Cohen, O., & Dailey, M. N (2017). Machine vision for counting fruit on mango tree canopies. Precision Agriculture, 18, 224-244.

[16]. Apolo-Apolo, O. E., Martínez-Guanter, J., Egea, G., Raja, P., & Pérez-Ruiz, M. (2020).

Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. European Journal of Agronomy, 115, 126030.

[17]. Nguyen, H. H. C., Luong, A. T., Trinh, T. H., Ho, P. H.,Meesad, P., Nguyen, T. T (2021).Intelligent Fruit Recognition System Using Deep Learning, Recent Advances in Information and Communication Technology, Lecture Notes in Networks and Systems, vol 251. Springer.

[18]. Koirala, A., Walsh, K. B., Wang, Z., and McCarthy, C (2019). Deep learning – Method overview and review of use for fruit detection and yield estimation. Computers and Electronics in Agriculture, 162, 219–234.

[19]. Horea, M., and Mihai, O (2018). Fruit recognition from images using deep learning, Acta Uni. Sapientiae, Informatica10, 1, 26-42.

[20]. Vasconez, J. P., and Delpiano, J., Vougioukas, S., and Cheein, F. A (2020). Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation, Computers and Electronics in Agriculture, volume 173.

[21]. Sa, I., Ge, Z., Dayoub F., Upcroft, B., Perez, T., McCool, C (2016).DeepFruits: A Fruit Detection System Using Deep Neural Networks. Sensors. 16(8):1222.

[22]. Häni, N., Roy, P., and Isler, V (2019). A comparative study of fruit detection and counting methods for yield mapping in apple orchards. Journal of Field Robotics.

[23]. Afonso, M.,Fonteijn, H.,Fiorentin, F. S., Lensink, D.,Mooij, M., Faber, N., Polder, G., and Wehrens, R (2020). Tomato Fruit Detection and Counting in Greenhouses Using Deep Learning, Front. Plant Sci., Sec. Technical Advances in Plant Sci., Volume 11. 267 www.scope-journal.com Scope Volume 13 Number 4 December 2023.

[24]. Mamdouh, N. and Khattab, A (2021). YOLO-Based Deep Learning Framework for Olive Fruit Fly Detection and Counting, in IEEE Access, vol. 9, pp. 84252-84262.

[25]. Mirhaji, H., Soleymani, M., Asakereh, A., and Mehdizadeh, S. A (2021). Fruit detection and load estimation of an orange orchard using the YOLO models through simple approaches in different imaging and illumination conditions, Computers and Electronics in Agriculture, Volume 191.

[26]. Parico, A. I. B. and Ahamed, T (2021). Real Time Pear Fruit Detection and Counting Using YOLOv4 Models and Deep SORT. Sensors, 21 (14).

[27]. Ge, Y., Lin, S., Zhang, Y., Li, Z., Cheng, H., Dong, J., Shao, S., Zhang, J., Qi, X., Wu, Z (2022). Tracking and Counting of Tomato at Different Growth Period Using an Improving YOLO Deepsort Network for Inspection Robot. Machines, 10, 489.

[28]. Wang, C., Bochkovskiy, A., Liao H. M (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, Computer Vision and Pattern Recognition, arXiv:2207.02696v1.