Scope
Volume 13 Number 4 December 2023

Automated On-Tree Mango Fruit Detection and Counting Through Computer Vision

Vidhya N P
Research Scholar in Computer Science
University of Kerala, Thiruvananthapuram, India

Priya R
Assistant Professor
Department of Computer Science
Government College Kariavattom, Thiruvananthapuram, India

Abstract

Using imaging and computer vision to precisely identify and quantify fruits at different stages of plant development is important not only to optimize labour-intensive manual measurements of phenotypic data but also to make significant progress towards task automation. The estimation of fruit yield plays a pivotal role in Precision Agriculture, aiding growers in more precisely forecasting market planning, workforce scheduling, procurement of suitable equipment, and other related considerations. There is also a high demand for automated methods to estimate fruit yield accurately and reliably in orchards. The advancements in Deep Learning models have been a great boon in this regard. In this work, YOLOv5 model is employed to detect and count mangoes on trees. The model has achieved a mean Average Precision@0.5 (mAP@0.5) as 99.5% and the MAP across the range of 0.5 to 0.95 as 81.9%, with a Mean Absolute Error (MAE) of 1.5 during testing.

Keywords: Computer Vision, labelImg, YOLOv5, Mean Average Precision (MAP), Mean Absolute Error (MAE).

1.0. Introduction

The Indo-Burmese region is the origin of the mango (Mangifera indica Linn.), a dicotyledonous fruit belonging to the Anacardiaceae family. It is mostly produced in tropical developing nations; the estimated global production is 15.06 million tonnes [1]. 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. Mangiferin has potent antioxidant, anti-lipid peroxidation, immunomodulation, cardiotonic, hypotensive, wound healing, antidegenerative, and antidiabetic properties. It is a polyphenolic antioxidant and glucosyl xanthone [2]. The mango is India's most important fruit crop in terms of commerce, accounting for over 54% of all mangoes produced globally. There are more than thirty varieties of mangoes grown. 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 [3]. 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.

Using cutting-edge computer vision algorithms to analyse images and determine the presence and quantity of mangoes is a key component of deep learning techniques for mango detection and counting in trees. A review by V. 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[4].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. 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 [5]. L. Yuzhi et al. utilized the bibliometric software Cite Space to perform visualization analysis on literature within the core database of Web of Science [6]. 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[7].

Ball pepper plant leaf bacterial spot disease was identified by M. P. Mathew et al. using YOLOv5 with mAP 90.7%. [8]. 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 [9].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 [10]. 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% [11]. 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 [12].

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 [13]. 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% [14].H. H. C. Nguen 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 [15]. 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[16].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% [17].

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 [18].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 [19]. N. Hani et al. attained high yield estimation accuracies 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 [20].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 [21].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% [22].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[23]. 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 [24].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[25].

Literature reveals that YOLOv5 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 YOLOv5 methodology, applying post-processing methods for accurate counting.

2.0. Materials and Methods

The study involves a comprehensive assessment of the Computer Vision based YOLOv5 deep learning technique aimed at addressing the task of detecting and counting mangoes on trees. The dataset utilized for this evaluation was curated using images gathered from internet sources. The Fig. 1 shows the proposed methodology of detection and counting.

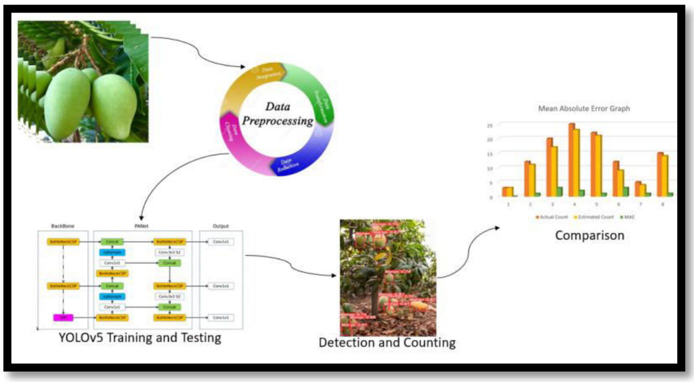

Fig. 1. Proposed methodology

2.1. Dataset

The dataset utilized in this study is compiled from various internet sources, including the Kaggle dataset (https://www.kaggle.com/datasets/warcoder/mangofruitdds/), the Freepik website (https://www.freepik.com/), and freely available images obtained through Google searches. This diverse compilation of data enhances the robustness and inclusivity of the dataset, providing a comprehensive foundation for the research. Exclusively healthy mangoes were incorporated into the dataset. The dataset comprises a diverse selection of mango varieties, encompassing various types 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 conditions for fruit detection. The deliberate variation in lighting, angles, and distances within the dataset contributes to a more comprehensive training environment, equipping the model to effectively adapt to real-world variations and challenges in mango detection scenarios. Fig. 2 illustrates one of the images included in the dataset.

Fig. 2. Example of dataset

2.2. Pre-processing

The dataset has undergone pre-processing methods, incorporating both image size equalization and data augmentation. In order to accommodate variations in image dimensions arising from diverse sources, a standardized dimension of 640 x 640 has been uniformly applied. Additionally, to combat the potential issue of model overfitting, data augmentation techniques have been applied to each image. This strategic approach introduces variability into the training data, aiming to enhance the model’s generalization. Precise annotations are performed on a total of 559 images using the labelImgtool, which facilitates defining bounding boxes and also outlining the corresponding pixels for each fruit. Subsequently, the dataset was randomly partitioned into an 80% training set and 20% validation set. Each image underwent manual labelling to specify the location and boundaries of every mango on the tree. This meticulous annotation process serves as the ground truth for training the deep learning model, providing crucial reference points for accurate detection and evaluation.

2.3. Computer Vision and YOLO Training

In this paper, a framework is introduced for detecting and counting mangoes on a tree in images captured by a camera. The primary objectives of this framework are accuracy and computational efficiency, focusing on detecting a single class, the mango fruit. 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 is the YOLOv5 model. According to Redmon, one-level object recognition architectures treat object recognition as a regression problem [26]. YOLOv5 adopts CSPDarknet as its backbone architecture by incorporating the cross-stage partial network (CSPNet) into Darknet. This algorithm ensures consistent measurements by minimizing variances arising from human error. In this study, the state-of-the-art real-time object detector YOLOv5 is employed to train a dataset comprising images of mangoes. The model simultaneously determines both the coordinates of bounding boxes and class probabilities on the input images, utilizing the PyTorch library in Python. The pretrained model is used to train the training set with an image size of 640 and a batch size of 16. The model undergoes 100 epochs of training to fine-tune and optimize the weights. The model's output includes a bounding box outlining the fruit's location, accompanied by a label and the associated probability for the class. The sample of the training batch images are depicted in Fig. 3.

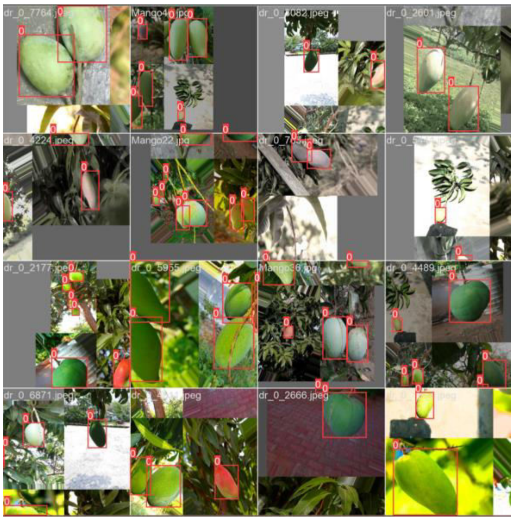

Fig. 3. Training batch images

Following training, the model's performance is assessed using the validation set, and subsequently, it is tested using the testing set. Fig. 4 depicts the sample of the validation batch images.

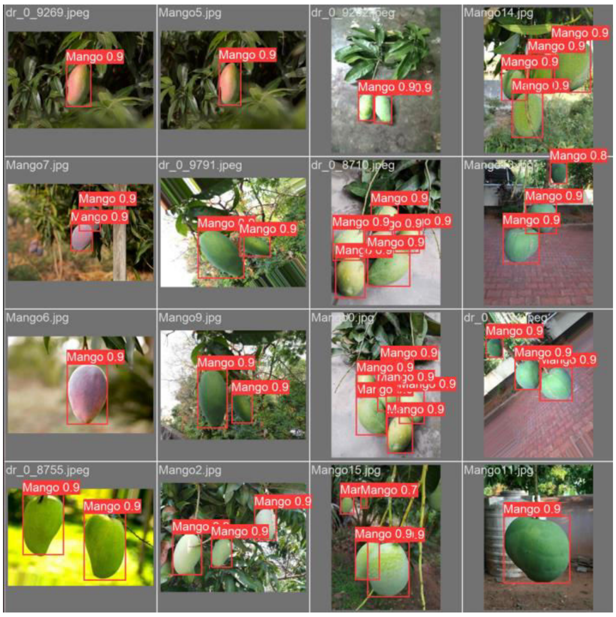

Fig. 4. Validation batch images with label and prediction value

3.0. Results and Discussion

The model's performance was evaluated on a separate validation dataset. The model was fine-tuned to improve its accuracy and generalization to new data. The fine-tuned trained model was deployed to analyse new images. The model detects and localizes mangoes in the tree, providing bounding boxes and probability scores around each mango.

The mAP@0.5 is the average precision of the fruit detection when IoU (Intersection over Union) as 0.5 and in mAP0.5:0.95, IoU value ranges from 0.5 to 0.95. The IoU can be calculated using the equation Eq. (1).

$$IoU = \frac{Area\ of\ overlap}{Area\ of\ union} \qquad (1)$$

$mAP = \frac{\sum AP_k}{n}$, where AP is the average precision and n is the number of classes.

The accuracy of mAP@0.5 is 99.5% and that of mAP@0.5-0.95 is 81.9% in the training period. Fig. 4 depicts the graphs generated during training period.

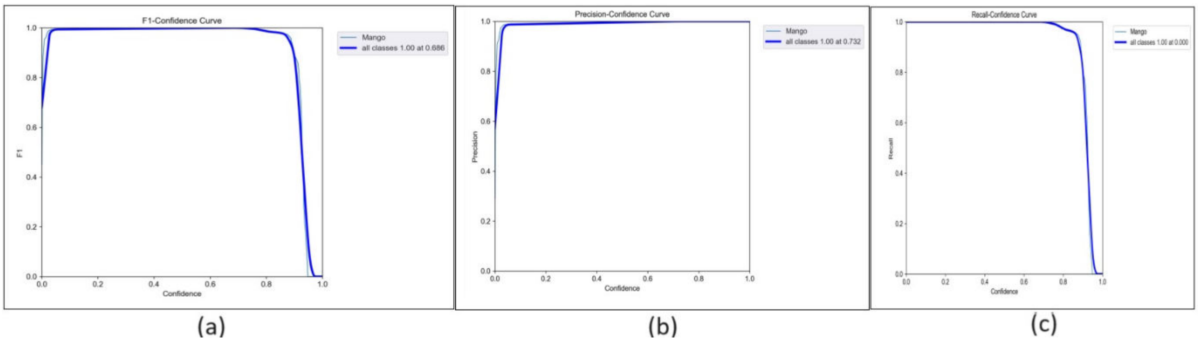

Fig. 4. (a) F-Confidence Curve (b) Precision-Confidence Curve (c) Recall-Confidence Curve

Implementing the trained model in real-world scenarios involves deploying it to process images of mango trees, ensuring accurate detection and counting of mangoes. Post-processing techniques are utilized to optimize the count by eliminating duplicate labels. The test image with labelling box and probability score shows in the Fig. 5.

Fig. 5. Original and labelled image after testing

The relative error is characterized as the ratio between the number of labels assigned to the image during testing and the actual count of fruits on the tree. This calculation of error quantifies the magnitude of the absolute error concerning the true number. The relative error serves as a measure of the accuracy of the labelled count relative to the actual quantity of fruits being assessed. The computation of MAE can be expressed using Eq. (2). It is found that the MAE is nearly about 1.5.

$$MAE = \frac{\sum_{k=0}^{n}(number\ of\ actual\ fruits - number\ of\ labels)}{n} \qquad (2)$$

4.0. Conclusion and Future work

This paper discusses the on-tree mango detection and counting using YOLOv5 pretrained model precisely. Compared to conventional techniques, using deep learning for mango detection and counting 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.

References

1. Shah, K. A., Patel, M. B., Patel, R. J., and Parmar, P. K (2010). Mangifera indica (mango). Pharmacogn Rev.4(7):42-8.
2. 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.
3. 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.
4. 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.
5. Yuzhi, L (2021). Research progress and trend analysis of computer vision based on cite space, Journal of Physics: Conference Series.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. Horea, M., and Mihai, O (2018). Fruit recognition from images using deep learning, Acta Uni. Sapientiae, Informatica10, 1, 26-42.
17. 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.
18. 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.
19. 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.
20. 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.

21. 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.
22. 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.
23. Parico, A. I. B. and Ahamed, T (2021). Real Time Pear Fruit Detection and Counting Using YOLOv4 Models and Deep SORT. Sensors, 21 (14).
24. 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.
25. Redmon, J.,Divvala, S.,Girshick, R., and 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.