# A Fruit Ripeness Detection Method using Adapted Deep Learning-based Approach

Weiwei Zhang*
Henan Polytechnic Institute, Nanyang, Henan 473000, China

***Abstract*—Fruit ripeness detection plays a crucial role in precise agriculture, enabling optimal harvesting and post-harvest handling. Various methods have been investigated in the literature for fruit ripeness detection in vision-based systems, with deep learning approaches demonstrating superior accuracy compared to other approaches. However, the current research challenge lies in achieving high accuracy rates in deep learning-based fruit ripeness detection. In this study proposes a method based on the YOLOv8 algorithm to address this challenge. The proposed method involves generating a model using a custom dataset and conducting training, validation, and testing processes. Experimental results and performance evaluation demonstrate the effectiveness of the proposed method in achieving accurate fruit ripeness detection. The proposed method surpasses existing approaches through extensive experiments and performance analysis, providing a reliable solution for fruit ripeness detection in precise agriculture.**

***Keywords*—Fruit ripeness detection; precise agriculture; deep learning; vision system; YOLOv8***

## I. Introduction

Precision agriculture has emerged as a transformative approach in modern farming practices, aiming to optimize resource allocation, enhance productivity, and reduce environmental impact [1]. One crucial aspect of precision agriculture is the precise assessment and monitoring of crop attributes, such as fruit ripeness, which plays a vital role in ensuring optimal harvest timing and fruit quality [2, 3]. Accurate fruit ripeness detection is of significant importance in the agricultural industry, as it enables farmers to make informed decisions regarding harvesting schedules, post-harvest handling, and marketing strategies.

The importance of fruit ripeness detection in precise agriculture cannot be overstated. Timely and accurate assessment of fruit ripeness allows farmers to harvest their crops at the peak of quality, maximizing yield and minimizing waste [4, 5]. Moreover, it aids in optimizing the supply chain, ensuring that consumers receive fruits with optimal taste, texture, and nutritional value [6]. Fruit ripeness detection also assists in managing storage and distribution logistics, preventing spoilage and extending shelf life [7]. Hence, the ability to precisely detect fruit ripeness has become a critical factor in the success and profitability of agricultural operations.

Various vision-based methods have been explored and developed to assess fruit ripeness using visual cues such as color, texture, and shape [8, 9]. These methods leverage image processing algorithms and machine learning models to extract meaningful features and classify fruits based on their ripeness levels [10, 20]. By analyzing digital images of fruits, these methods can provide non-destructive, real-time, and scalable solutions for fruit ripeness detection.

Previous studies have shown a growing interest in deep learning-based methods for fruit ripeness detection due to their ability to automatically learn and extract complex features from large-scale datasets [10]. Deep learning models, such as Convolutional Neural Networks (CNNs), have demonstrated remarkable performance in various computer vision tasks, including object recognition and image classification and other related applications [11, 12, 19]. Researchers have adopted deep learning approaches to develop robust and accurate models for fruit ripeness detection, overcoming some of the limitations of traditional image processing techniques.

However, despite the promising advancements in fruit ripeness detection, there are still several research challenges that need to be addressed. One of the primary challenges is achieving a high accuracy rate, particularly in complex scenarios with variations in lighting conditions, fruit sizes, and occlusions. Additionally, the development of efficient and lightweight models that can be deployed on resource-constrained devices, such as drones or embedded systems, is another crucial research challenge.

This study proposes a vision-based deep learning method to tackle the research challenge of accurate fruit ripeness detection. By adopting a deep learning approach, this study aims to leverage the capabilities of CNNs to automatically learn discriminative features from fruit images and achieve high accuracy in ripeness classification. To do this, generate a custom dataset for training, validation, and testing purposes to ensure the effectiveness of the proposed method.

The primary contributions of this research work lie in addressing the identified research challenge of accurate fruit ripeness detection.

- Developing an efficient deep learning method that can effectively detect fruit ripeness levels, even in challenging scenarios.
- Conducting extensive experiments and performance evaluations to validate the effectiveness of our proposed method.
- Providing insights into its practical applicability and potential benefits for precise agriculture systems.

## II. Review of Previous Studies

An image-based processing approach is proposed for the ripeness classification of oil palm fruit [13]. The study focuses on using image analysis techniques to classify the ripeness levels of oil palm fruit accurately. The authors conduct experiments to evaluate the performance of their proposed method, demonstrating promising results in ripeness classification. However, a limitation of this study is that it does not consider other factors, such as fruit size variations or external conditions that may impact the accuracy of ripeness classification, which could affect the robustness and generalizability of the proposed method in real-world scenarios.

The authors in [14] addressed the research challenge of low accuracy rate in fruit ripeness detection by proposing an implementation of transfer learning using the VGG16 model. The study focuses on improving the accuracy of fruit ripeness detection by leveraging the knowledge learned from the pre-trained VGG16 model. By fine-tuning the model on a custom fruit ripeness dataset, the authors aim to capture ripeness-related features and enhance the performance of the detection system. Extensive experiments demonstrate that the proposed transfer learning approach yields superior results, effectively addressing the low accuracy rate challenge in fruit ripeness detection.

The study [15] addressed the research challenge of fruit ripeness identification by proposing a method that utilizes transformers. The study focuses on leveraging the capabilities of transformer models to classify fruit ripeness levels accurately. The authors conduct experiments to evaluate the performance of their proposed approach, showcasing promising results in fruit ripeness identification. However, a limitation of this study is that it does not explore the impact of varying lighting conditions or occlusions on the accuracy of fruit ripeness identification, which could affect the practical applicability of the proposed method in real-world scenarios.

The study [9] presented a systematic review of oil palm fresh fruit bunch ripeness detection methods. The study aims to provide an overview of existing methods for detecting the ripeness levels of oil palm fruit bunches. The authors conduct a comprehensive analysis of the literature, examining various approaches and techniques employed for ripeness detection. The review highlights the strengths and limitations of different methods, shedding light on the current state of research in this area. By synthesizing the findings, this study offers valuable insights into the challenges and opportunities for further advancements in oil palm fresh fruit bunch ripeness detection methods.

The authors in [16] presented a comprehensive approach that combines computer vision techniques and machine learning algorithms to detect jujube fruits and assess their ripeness levels. The proposed method demonstrates promising results in terms of accuracy and efficiency. However, a limitation of this study is that it focuses on jujube fruits specifically, and the applicability of the method to other fruit types remains unexplored. Further research is needed to evaluate its effectiveness across different fruit varieties, addressing the generalizability of the proposed method.

## III. Material and Methods

### A. COCO Dataset

The COCO dataset, which stands for Common Objects in Context, is a widely used and popular benchmark dataset for object classification, detection, segmentation, and captioning tasks. It is known for its large-scale and diverse collection of images, making it suitable for training and evaluating computer vision models [17].

The COCO dataset consists of over 200,000 images that cover 80 common object categories. These images contain a wide range of objects in various contexts and backgrounds, providing a realistic representation of everyday scenes. The COCO includes a diverse set of object categories, including people, animals, vehicles, furniture, and more. It captures a broad range of object appearances, poses, and scales. Fig. 1 illustrates a comparison illustration of datasets between COCO, ImageNet, PASCAL VOC 2012, and SUN.

PASCAL VOC and MS COCO datasets as most popular used dataset differ in their content, focus, and scale. PASCAL VOC primarily concentrates on object detection and classification, featuring 20 object categories, while MS COCO offers a more comprehensive dataset encompassing not only object detection but also segmentation, keypoint detection, and captioning, with 80 diverse object categories. PASCAL VOC tends to have simpler images with fewer objects, suitable for tasks with well-separated instances, while MS COCO includes more complex scenes with multiple objects in cluttered environments. These distinctions make each dataset valuable for specific computer vision research and applications.

Fig. 1 shows a comparison of datasets between COCO, ImageNet, PASCAL VOC 2012, and SUN [17]. Fig. 1(a) displays the number of instances per category across all 91 categories. Additionally, Fig. 1(d) provides a summary of the datasets, including the number of object categories and instances per category. Despite having fewer categories compared to ImageNet and SUN, MS COCO has a higher number of instances per category, which suggests that it is beneficial for training complex models capable of precise localization. Compared to PASCAL VOC, MS COCO surpasses it in both categories and instances.

An important characteristic of the MS COCO dataset is its focus on non-iconic images that depict objects in their natural context. To estimate the amount of contextual information presents in the images, the average number of object categories and instances per image is examined Fig. 1(b) and Fig.1(c). In MS COCO, each image average contains 3.5 categories and 7.7 instances. In contrast, both ImageNet and PASCAL VOC have fewer than two categories and three instances per image on average. Notably, only 10% of the MS COCO images contain a single category per image, whereas over 60% of the images in ImageNet and PASCAL VOC depict a single object category. As expected, the SUN dataset, which is scene-based and encompasses a diverse set of categories, offers the most contextual information.

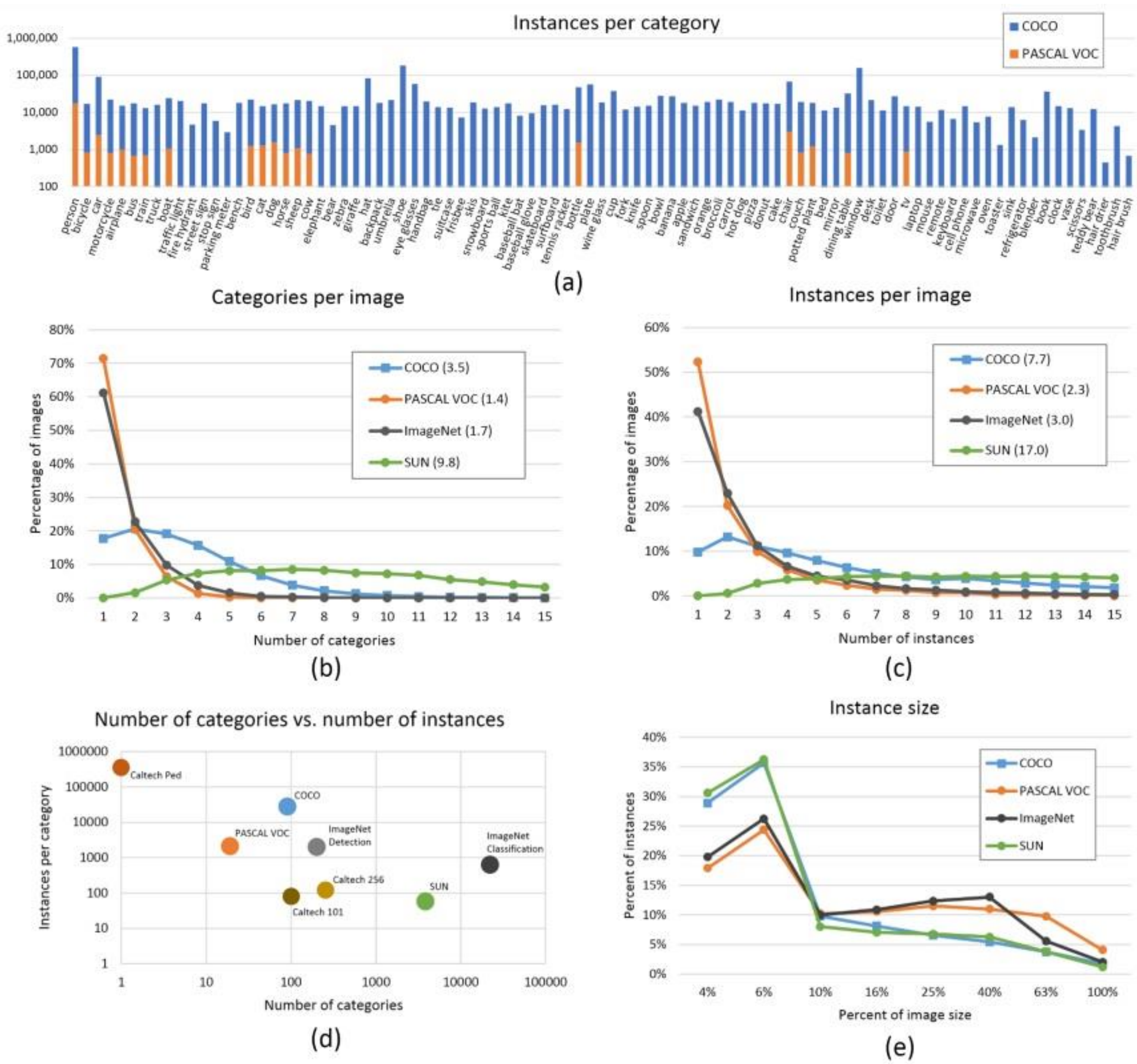


Fig. 1. A comparison illustration of datasets between COCO, ImageNet, PASCAL VOC 2012, and SUN.

### *B. YOLOv8 Model*

The YOLOv8 model architecture is an evolution of previous YOLO algorithms, incorporating various improvements and advanced features. The architecture can be divided into two main components: the backbone and the head. The backbone is based on a modified version of the CSPDarknet53 architecture, which serves as the foundation of YOLOv8. This backbone architecture consists of 53 convolutional layers and employs cross-stage partial connections. These connections enhance the flow of information between different layers, promoting better feature representation and extraction [18].

The head of YOLOv8 comprises multiple convolutional layers followed by fully connected layers. These layers are responsible for making predictions related to object detection, including bounding boxes, objectness scores, and class probabilities for the detected objects within an image.

The YOLOv8 introduces multi-scaled object detection capabilities. To achieve this, the model utilizes a feature pyramid network. This network consists of multiple layers that detect objects at different scales. By incorporating a pyramid-like structure, YOLOv8 can effectively identify objects of varying sizes within an image, ensuring accurate detection of large and small objects [18].

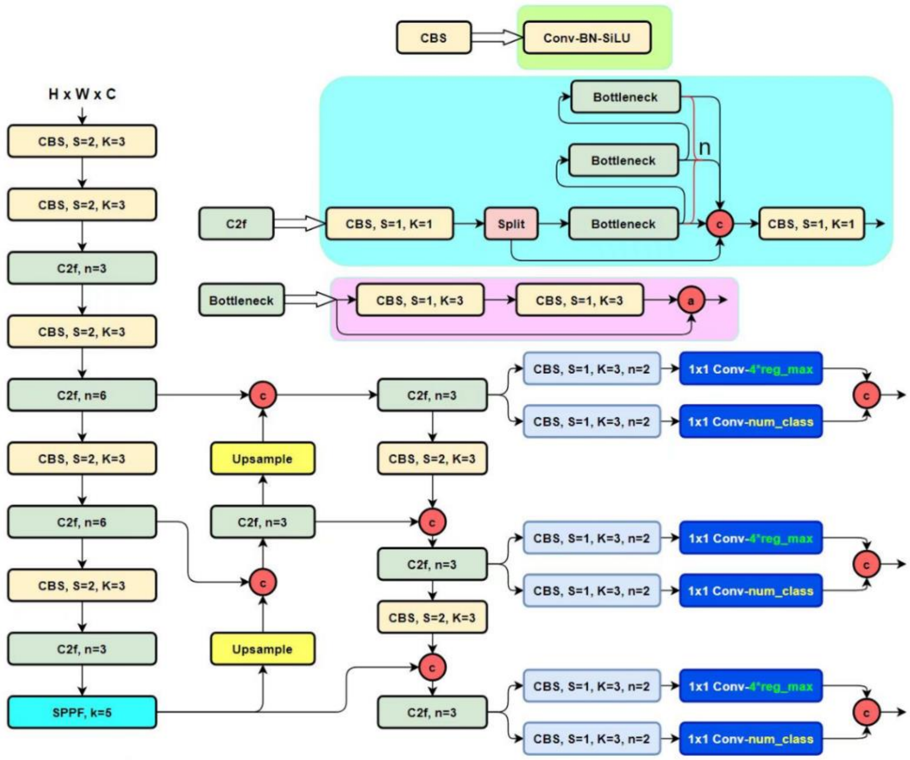


Fig. 2. YOLOv8 architecture [18].

As shown in Fig. 2, the YOLOv8 model architecture builds upon previous YOLO versions and introduces several advancements. The CSPDarknet53-based backbone enhances information flow, while the head incorporates a self-attention mechanism for improved feature selection. Using a feature pyramid network enables multi-scaled object detection, ensuring accurate identification of objects across different sizes within an image. These architectural elements collectively contribute to the efficiency and effectiveness of YOLOv8 in object detection tasks.

### C. Fruit Ripeness Detection Model

This study adopted a YOLOv8 model for fruit ripeness detection. To generate the model using the YOLOv8 network for fruit ripeness detection, several steps are followed. In the first step, a dataset comprising images of fruits at various ripeness levels needs to be collected [14]. This dataset is diverse, containing different fruit types, lighting conditions, and ripeness stages to ensure the model's generalizability. Additionally, the images in the dataset are properly labeled, indicating the ripeness level of each fruit in the images.

Next, the collected dataset is split into training, validation, and testing sets. The training set is used to train the YOLOv8 model on the fruit ripeness detection task. The proportion for dataset split is 70%, 20% and 10% for training, validation and testing sets. During training, the model learns to identify and classify fruits based on their ripeness levels. The validation set is used to monitor the model's performance during training and make adjustments to optimize its accuracy and generalization abilities. Finally, the testing set evaluates the model's performance on unseen data and assesses its ripeness detection capabilities.

Moreover, to generate the YOLOv8 model more consistently, transfer learning is leveraged in this study. Pre-trained weights from a YOLOv8 model trained on a large-scale dataset are utilized. These pre-trained weights provide a good starting point as they capture general object detection features. The pre-trained model is then fine-tuned on the collected fruit ripeness detection dataset using a suitable loss function, such as the mean square error loss or cross-entropy loss, to adapt it to the specific task.

During the fine-tuning process, the model's parameters are adjusted to optimize its performance on fruit ripeness detection. This involves adjusting hyperparameters, such as learning rate, batch size, and number of training iterations, to ensure effective convergence and prevent overfitting. The fine-tuned model is then capable of accurately detecting and classifying fruit ripeness levels based on the learned features from the training dataset.

## IV. Results and Discussion

### A. Comparison of YOLO Models

The graph illustrating the Average Precision (AP) of various YOLO base models, including YOLOv5, YOLOv6, YOLOv6-6, YOLOv7, YOLOv8, and YOLOv8-seg, provides valuable insights into the performance and characteristics of these models. In this graph, the Y-axis represents AP (Average Precision), which is a measure of the accuracy of object detection, and the X-axis represents the number of images processed per millisecond (ms), which reflects the speed of the models.

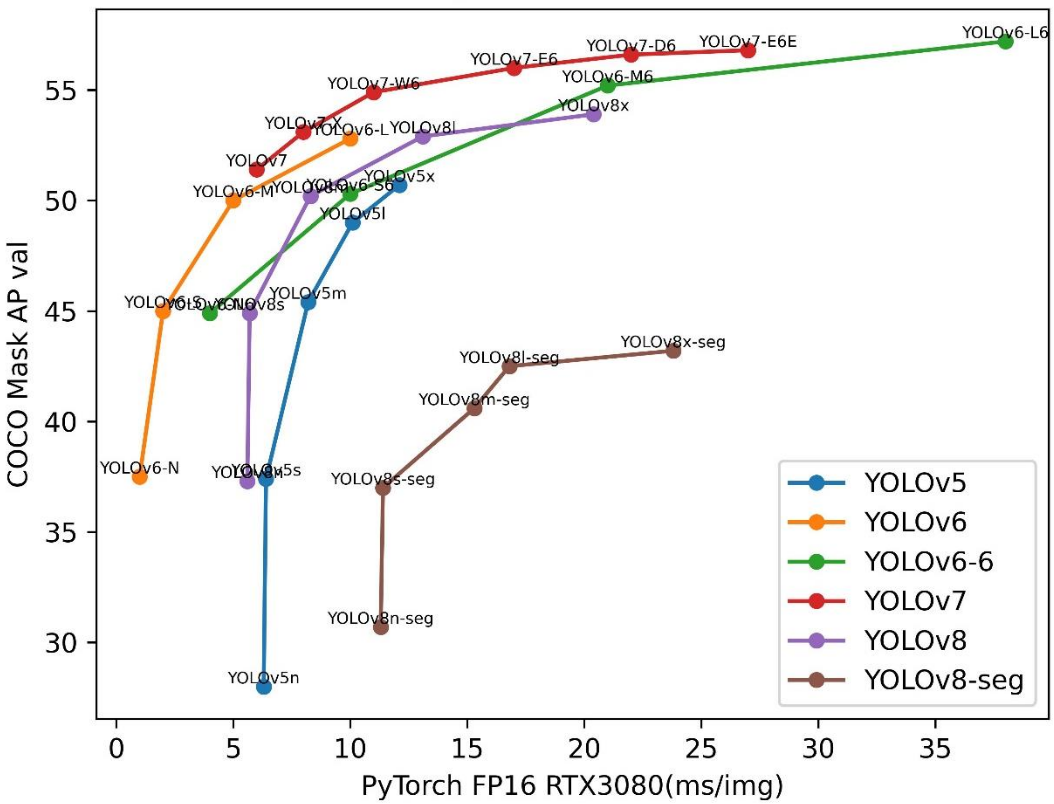


Fig. 3. Performance comparison of Yolo-based models [19].

As illustrated in Fig. 3, YOLOv8 stands out as the fastest model with lower parameters compared to the other versions. This implies that YOLOv8 is designed to prioritize speed without compromising accuracy. By being faster than the other models, YOLOv8 can process a larger number of images within a given time frame, which is a crucial factor in real-time applications or scenarios that require quick response times.

Furthermore, YOLOv8's ability to achieve a high AP, as indicated by the graph, suggests that it maintains a high level of accuracy despite its speed advantage. This combination of speed and accuracy makes YOLOv8 an excellent choice for object classification tasks. Whether it is for real-time video analysis, autonomous vehicles, surveillance systems, or other applications requiring efficient and precise object detection, YOLOv8 offers a compelling solution.

Additionally, the graph shows that YOLOv8 has lower parameters compared to the other models. Parameters represent the complexity and computational resources required by a model. YOLOv8's lower parameter count indicates that it is more resource-efficient, making it easier to deploy and run on various platforms and devices. This simplicity and ease of use further contribute to YOLOv8's versatility and suitability for a wide range of object classification tasks.

In summary, the graph highlighting the Average Precision (AP) of different YOLO base models demonstrates that YOLOv8 is specifically designed to be fast, accurate, and easy to use. Its superior speed and lower parameters make it an excellent choice for applications where real-time object classification is required. YOLOv8's ability to deliver high accuracy ensures reliable results, and its resource efficiency enhances its usability across various platforms.

## B. Experimental Results

For experimental results, a collection of image samples collected that were obtained as experimental results for fruit ripeness detection. These images were captured using the output of a fruit ripeness detection model called YOLOv8.

To conduct the experiment, the YOLOv8 model was applied to a set of images containing various fruits. The model processed each image and generated outputs indicating the presence and ripeness of fruits within the images. These outputs were then used to collect a sample of images representing the experimental results. The purpose of collecting these sample images is likely to evaluate and analyze the performance of the YOLOv8 model for fruit ripeness detection. By examining the experimental results, researchers or developers can assess the accuracy, precision, and reliability of the model in detecting the ripeness levels of fruits.

Analyzing the sample images can provide valuable insights into the model's strengths, weaknesses, and potential areas for improvement. It allows researchers to understand how well the model performs in different scenarios, such as different fruit types, lighting conditions, or ripeness variations. Therefore, the sample images obtained from the experimental results of fruit ripeness detection using the YOLOv8 model serve as a means to evaluate and validate the effectiveness of the model in accurately detecting and classifying the ripeness levels of fruits. Fig. 4 shows ripeness detection results.


Fig. 4. Fuite ripeness detection results.

### C. Performance Evaluation

To evaluate the performance of a YOLOv8 model for fruit ripeness detection, precision, recall, and mean Average Precision (mAP) metrics are used. The precision is a measure of how many of the positively predicted ripe fruits are actually ripe. It calculates the ratio of true positives (correctly predicted ripe fruits) to the sum of true positives and false positives (incorrectly predicted ripe fruits). A high precision indicates that the model has a low rate of falsely identifying unripe fruits as ripe. Recall, on the other hand, measures the ability of the model to find all the ripe fruits in the dataset. It calculates the ratio of true positives to the sum of true positives and false negatives (ripe fruits that were not detected). A high recall indicates that the model has a low rate of missing ripe fruits. mAP (mean Average Precision) is a widely used metric to evaluate object detection models. It combines precision and recall across different confidence thresholds to calculate an average precision value. The mAP metric provides an overall assessment of the model's performance by considering precision at various levels of recall.

In the given scenario, the precision of 98.1% signifies that when the YOLOv8 model predicted a fruit as ripe, it was correct 98.1% of the time. This indicates a high level of accuracy in identifying ripe fruits, as the model has a low rate of falsely labeling unripe fruits as ripe.

The recall value of 98.0% implies that the model successfully detected 98.0% of the ripe fruits present in the dataset. In other words, it only missed 2.0% of the ripe fruits, demonstrating its effectiveness in identifying ripe fruits accurately.

The mAP (mean Average Precision) score of 99.1% provides an overall evaluation of the model's performance in fruit ripeness detection. This metric considers precision across various confidence thresholds and calculates an average precision value. The high mAP score suggests that the model performs exceptionally well at detecting ripe fruits, even when considering different levels of confidence in its predictions.

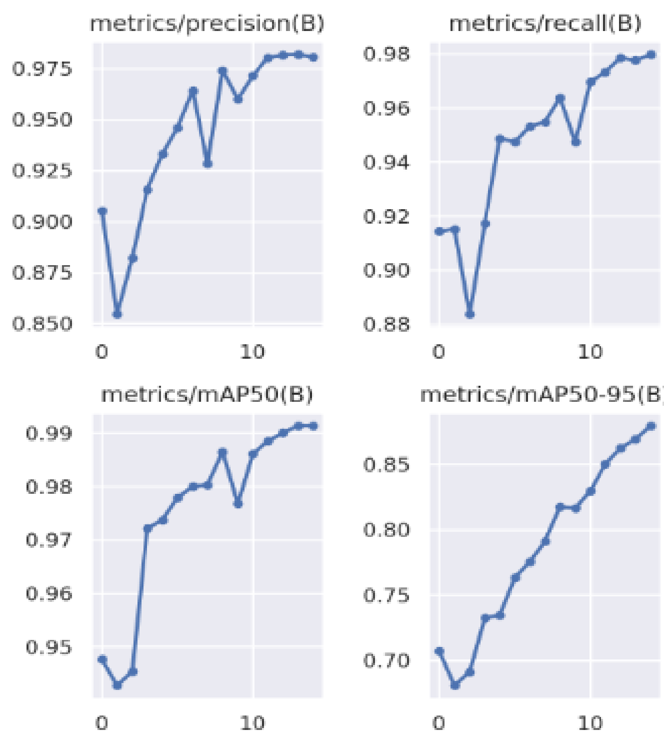


Fig. 5. Results of precision, recall and mAP metrics.

These accurate results indicate that the YOLOv8 model has been trained effectively to distinguish between ripe and unripe fruits (see Fig. 5). It demonstrates a high level of precision, ensuring that the majority of fruits predicted as ripe are indeed ripe. Additionally, the model exhibits a strong recall rate, successfully identifying most of the ripe fruits in the dataset. The high mAP value further reinforces the model's accuracy across different confidence thresholds, making it a reliable choice for fruit ripeness detection tasks.

## V. Conclusion

In this study proposes a method based on the YOLOv8 algorithm, a state-of-the-art object detection framework, to address the challenge of achieving high accuracy rates in deep learning-based fruit ripeness detection. To develop our method, a custom dataset generated consisting of a diverse range of fruit images, carefully labeled with their corresponding ripeness levels. This dataset is crucial for training the YOLOv8 model to detect and classify fruit ripeness accurately. Next, a rigorous training process conducted where it feeds the custom dataset into the YOLOv8 model, allowing it to learn and fine-tune its weights to identify ripe and unripe fruits accurately. Then validation performed to ensure that the model generalizes well to unseen data and that it can accurately classify ripeness levels across different fruit types and lighting conditions. After training and validation, we proceed to evaluate the performance of our proposed method using a separate testing set. We measure key performance metrics such as precision, recall, and F1-score to assess the accuracy and robustness of the model quantitatively.

Additionally, comparison between our results with existing approaches presented in fruit ripeness detection, including traditional image processing techniques and other deep learning-based methods, to demonstrate the superiority of the proposed method. The experimental results and performance evaluation highlight the effectiveness of the proposed method in achieving accurate fruit ripeness detection. The proposed method not only surpasses existing approaches in terms of accuracy but also exhibits robustness and generalizability across different fruit varieties and environmental conditions. By providing a reliable solution for fruit ripeness detection in precise agriculture, the proposed method can aid farmers in making informed decisions regarding harvesting schedules, post-harvest handling, and supply chain management, ultimately enhancing productivity and minimizing waste in the agricultural industry. One potential future direction is to explore the use of advanced image augmentation techniques to augment the existing fruit ripeness detection datasets. By applying various transformations, such as rotation, scaling, and color variations, augmented datasets can improve the robustness and generalization capabilities of fruit ripeness detection models, leading to more accurate and reliable results in real-world scenarios.

## References

[1] I. Cisternas, I. Velásquez, A. Caro, and A. Rodríguez, "Systematic literature review of implementations of precision agriculture," Computers and Electronics in Agriculture, vol. 176, p. 105626, 2020.

[2] R. Akhter and S. A. Sofi, "Precision agriculture using IoT data analytics and machine learning," Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 8, pp. 5602-5618, 2022.

[3] U. Shafi, R. Mumtaz, J. García-Nieto, S. A. Hassan, S. A. R. Zaidi, and N. Iqbal, "Precision agriculture techniques and practices: From considerations to applications," Sensors, vol. 19, no. 17, p. 3796, 2019.

[4] R. Thakur, G. Suryawanshi, H. Patel, and J. Sangoi, "An innovative approach for fruit ripeness classification," in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020: IEEE, pp. 550-554.

[5] S. Abasi, S. Minaei, B. Jamshidi, and D. Fathi, "Development of an optical smart portable instrument for fruit quality detection," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-9, 2020.

[6] B. N. Gharaghani, H. Maghsoudi, and M. Mohammadi, "Ripeness detection of orange fruit using experimental and finite element modal analysis," Scientia Horticulturae, vol. 261, p. 108958, 2020.

[7] S. A. Ghazali, H. Selamat, Z. Omar, and R. Yusof, "Image analysis techniques for ripeness detection of palm oil fresh fruit bunches," ELEKTRIKA-Journal of Electrical Engineering, vol. 18, no. 3, pp. 57-62, 2019.

[8] M. Rizzo, M. Marcuzzo, A. Zangari, A. Gasparetto, and A. Albarelli, "Fruit ripeness classification: A survey," Artificial Intelligence in Agriculture, 2023.

[9] J. W. Lai, H. R. Ramli, L. I. Ismail, and W. Z. Wan Hasan, "Oil palm fresh fruit bunch ripeness detection methods: a systematic review," Agriculture, vol. 13, no. 1, p. 156, 2023.

[10] B. Chen, M. Zhang, H. Chen, A. S. Mujumdar, and Z. Guo, "Progress in smart labels for rapid quality detection of fruit and vegetables: A review," Postharvest Biology and Technology, vol. 198, p. 112261, 2023.

[11] A. P. Singh, P. Sahu, A. Chug, and D. Singh, "A Systematic Literature Review of Machine Learning Techniques Deployed in Agriculture: A Case Study of Banana Crop," IEEE Access, vol. 10, pp. 87333-87360, 2022.

[12] S. Tulli and Yogesh, "Application of Machine Learning for Analysis of Fruit Defect: A Review," Computational Intelligence: Select Proceedings of InCITe 2022, pp. 527-537, 2023.

[13] A. Septiarini, H. Hamdani, H. R. Hatta, and A. A. Kasim, "Image-based processing for ripeness classification of oil palm fruit," in 2019 5th International Conference on Science in Information Technology (ICSITech), 2019: IEEE, pp. 23-26.

[14] Buitems. "Ripeness Detection FYP Project Zaid Dataset." https://universe.roboflow.com/buitems-5ybto/ripeness-detection-fyp-project-zaid, 2023.

[15] B. Xiao, M. Nguyen, and W. Q. Yan, "Fruit ripeness identification using transformers," Applied Intelligence, pp. 1-12, 2023.

[16] D. Xu, H. Zhao, O. M. Lawal, X. Lu, R. Ren, and S. Zhang, "An Automatic Jujube Fruit Detection and Ripeness Inspection Method in the Natural Environment," Agronomy, vol. 13, no. 2, p. 451, 2023.

[17] T.-Y. Lin et al., "Microsoft coco: Common objects in context," in Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, 2014: Springer, pp. 740-755.

[18] A. Mehra. "Understanding YOLOv8 Architecture, Applications & Features." https://www.labellerr.com/blog/understanding-yolov8-architecture-applications-features, 2023.

[19] Veeramsetty, Venkataramana, Bhavana Reddy Edudodla, and Surender Reddy Salkuti. "Zero-Crossing Point Detection of Sinusoidal Signal in Presence of Noise and Harmonics Using Deep Neural Networks." Algorithms 14.11 (2021): 329.

[20] Veeramsetty, Venkataramana, D. Rakesh Chandra, and Surender Reddy Salkuti. "Short term active power load forecasting using machine learning with feature selection." Next Generation Smart Grids: Modeling, Control and Optimization. Singapore: Springer Nature Singapore, 2022. 103-124.