

## Smart Ripeness Assessment: Real-Time Image-Based Classification of Papaya Maturity Stages

Mahalaxmi Kalhal<sup>1</sup>, Sadashiv Bellubbi<sup>2</sup>, Pavan Kunchur<sup>3\*</sup>, Arati Shahapurkar<sup>4</sup>

<sup>1,3,4</sup> Department of Computer Science and Engineering, Gogte Institute of Technology, Belagavi– 590008, Karnataka, India

<sup>2</sup>Department of Mechanical Engineering, Jain College of Engineering, Belagavi– 590008, Karnataka, India

<sup>1</sup>mahalaxmibellubbi@gmail.com, <sup>2</sup>bellubbisadashiv@gmail.com, <sup>3\*</sup>pnkunchur@git.edu, <sup>4</sup>asshahapurkar@git.edu

**Abstract**— The classification of papaya maturity is vital for ensuring fruit quality, reducing waste and improving supply chain efficiency. Traditionally, farmers assess papaya ripeness manually by inspecting physical traits like colour, texture and softness, which is labour-intensive, subjective, time-consuming, and prone to errors. The CNN model improved accuracy but often lacked the speed needed for real-time use. This study proposes a real-time, automated papaya maturity classification system using the YOLOv8 model, which captures images using web-cam, extracts feature and classifies papayas as mature, semi- mature, immature, or rotten. The results show that the YOLOv8 model gave better accuracy than the CNN model

**Keywords**— *Papaya; YOLO v8; Maturity classification; DIP; Real time;*

### I. INTRODUCTION

Papaya is known as one of the delicious tropical fruits and it grows in the tropics and is widely spread around the world. Its scientific name is *Carica papaya*, and the fruit of the papaya grows on an evergreen tree, its height ranges from 2 to 10 meters. In addition, papaya fruit gives many benefits to human body such as, minimizing the risk of heart diseases, diabetes or cancer and contributes to maintaining healthy skin and hair.

Papayas ripen within a few days at room temperature, even faster in a paper bag. Once ripe, the fruit will quickly turn very mushy, if not stored properly. The maturity level of papaya directly affects its taste, texture, shelf life and marketability. Overripe or underripe papayas can lead to wastage or dissatisfaction among consumers, making the identification of the right harvesting time vital for ensuring optimal product quality. Usually, observer can easily identify and depict the state of papaya by its image. Machines can't perform this task as compared to human beings. Its timely and accurate classification based on maturity is crucial for optimizing its harvest, storage, and distribution. However, one of the major challenges in the commercialization of papaya is determining its optimal maturity stage for harvesting, transportation and sale.

Traditionally, the assessment of papaya maturity has been done manually, with farmers and vendors relying on visual and tactile inspections, which are often inconsistent, time-consuming and are prone to inaccuracies due to factors such as operator fatigue, bias, and subjectivity. Identification of papaya stage varies from person to person, which in-turn it difficult to maintain supply chain quality control. However, manual evaluation is not only time- consuming but often destructive to the samples. Consequently, there is a pressing need for rapid, intelligent and non-destructive techniques to enhance the efficiency, accuracy, and reliability of fruit quality assessment.

The investigation of contemporary technology, especially image processing and machine learning, has been prompted by the need for a more accurate and efficient way to evaluate papaya ripeness. The advent of modern technology, particularly real-time image processing, offers a solution to these challenges. This work mainly aims to utilize camera devices to automatically assess the maturity of papayas. The system will process visual features such as colour, texture and shape to classify papayas into distinct maturity stages green, ripe, or overripe. Image processing serves as a non-invasive, economical and scalable approach that can be effectively incorporated into the agricultural supply chain. Webcams are an affordable, easily accessible and versatile tool that can be integrated into the agricultural sector for real-time monitoring. The image processing algorithms will retrieve relevant characteristics from the captured images and use them to classify the papayas. This approach ensures faster, more consistent and objective maturity assessment, which helps in making informed decisions about harvesting, packaging and marketing to farmers, retailers and distributors.

This innovation has the potential to enhance product quality, reduce wastage and improve economic efficiency in the papaya industry. By

adopting this real-time classification system, the agricultural industry can minimize fruit waste, reduce manual labour and increase product quality. This work demonstrates real-time image processing using a normal webcam, it can revolutionize the way papayas are evaluated for maturity, leading to better economic outcomes for farmers and ensuring that consumers receive fresh, high-quality fruit. These techniques can contribute to the advancement of agricultural practices, improve the quality of food products and create economic opportunities for farmers and businesses involved in the papaya industry.

Image classification and identification are challenging tasks for machines which includes deep learning concepts. Early image classification relied on raw pixel data. Identifying two pictures of same object are difficult due to their different backgrounds, angles, poses etc. Further, it is challenge for computers to correctly 'see' and categories images. Image classification has a few uses and vast potential as it grows in reliability. Accurate determination of fruit maturity is crucial for assessing eating quality and predicting optimal storage duration.

### 1.1 Process Distribution of Fruit Classification:

Fruit detection is a crucial component of modern agriculture, as it directly impacts the efficiency and effectiveness of farming operations. The ability to precisely identify fruits and assess their ripeness is essential for maximizing crop yields, improving the quality of produce, and streamlining harvesting practices. In recent years, advancements in fruit detection technologies have gained considerable attention for their potential to transform the agricultural sector, offering new opportunities to optimize operations and ensure better outcomes for farmers.

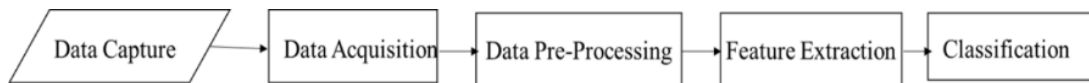


Figure 1 Process distribution of fruit classification.

**Data Capture:** The process of gathering visual data using a variety of sensors, including cameras, webcams, drones, or specialized imaging equipment, is known as data capture in image processing. Following processing, this data is used for tasks including monitoring, analysis, and object detection. Fruit identification, maturity evaluation, insect detection, and crop health monitoring are all accomplished through data gathering in agriculture. It entails taking pictures or videos in the best possible lighting, resolution, and angle, and it could require preprocessing techniques like image enhancement or noise reduction to boost quality prior to analysis.

**Data Acquisition:** In image processing, the term "data acquisition" describes the procedure of obtaining and capturing picture data for subsequent processing and analysis. To produce high-quality photos suitable for a variety of uses, it uses a number of techniques and technologies.

**Data Preprocessing:** Data preprocessing in image processing refers to the series of techniques and steps taken to prepare raw image data for analysis and processing. This step is essential since the efficacy of any further image analysis or feature extraction is directly impacted by the quality of the input data.

**Feature Extraction:** In image processing and computer vision, feature extraction is a crucial procedure that entails locating and measuring important features or qualities in images. By using these attributes, the image is represented in a form that facilitates object and pattern analysis, classification, and recognition. Successful feature extraction can greatly improve the results of subsequent tasks like object detections, segmentation and classification.

**Classification:** In image processing, classification is the process of giving images names or categories according to their content. Numerous applications, including as object recognition, medical diagnosis, and scene comprehension, depend on this stage. Traditional machine learning methods or more sophisticated deep learning techniques, such as the Yolov8 model, CNN model, can be used for classification. In the field of fruit identification, image processing is crucial because it allows automated systems to recognize and classify different fruit varieties according to their visual traits. This application combines methods from computer vision, machine learning, and artificial intelligence to enhance accuracy and efficiency in tasks that traditionally relied on manual inspection.

## II. LITERATURE SURVEY

This section describes few research works done by many investigators.

Masawabe et al.[1] presented a deep learning model based on VGG16 for classifying papaya maturity stages, achieved 100% accuracy in just 1m and 52 sec of training. It removes the need for manual feature extraction and despite a small dataset, proves highly effective and reliable for automating maturity assessment. Nada, AM Abu, et al. [2] presented age, gender and the recency of personal photographs play an increasingly vital role in identification, security, and business processes within various organizations and government agencies. To minimize errors in user-submitted information, enterprise systems incorporate form validation techniques. This paper introduces a method to enhance these validations by predicting age and gender from a single photograph and comparing the results against the provided details. The method achieves high accuracy in gender prediction however; precise age estimation remains a challenging aspect.

Abu-Saqer et al. [3] accurately classified different types of grapefruit, achieving an accuracy of 100%. A deep learning model based on convolutional neural networks (CNNs) was developed, utilizing a network architecture comprising four layers with a dropout rate of 0.2. The model effectively predicts the type of previously unseen grapefruit images. This system has the potential to support various applications, enabling users to automate processes in restaurants, factories, and other industries where grapefruit classification is required.

Abu-Saqer et al. [4] demonstrated remarkable accuracy in identifying and highlighting targets in images, even when they were partially or fully obscured. Its swift processing speed of approximately 20 milliseconds per image ensures efficiency for real-time applications. With its combination of speed and precision, the model proves to be a promising solution for diverse image recognition tasks.

Santi Kumari Behera et al. [5] represented a papaya fruit maturity classification model using machine learning and transfer learning. Among machine learning models, Weighted KNN with HOG features achieved 100% accuracy with the shortest training time (0.0995s). In transfer learning, VGG19 outperformed other models, also achieving 100% accuracy with the fastest training time (1m 52s). VGG19 is preferred for its efficiency and no need for feature extraction. The ` approach surpasses previous methods (94.7% accuracy) and can be extended for prototype development. Jairo Lucas de Moraes et al. [6] addressed a critical need in papaya cultivation by proposing a publicly available, multi-class annotated image database and a disease detection system based on convolutional block attention modules (CBAM). The proposed detector outperforms several state-of-the-art models, achieving an average mAP of 86.2%, even in challenging classes with high intra-class variation. Beyond its technical contributions, this work offers valuable potential for advancing efficient deep learning models, supporting industrial automation, enabling precision agriculture, and providing accessible solutions for small rural producers, thereby promoting sustainable and efficient fruit quality control practices.

Chenglin Wang et al. [7] presented an improved YOLOv8+ model capable of accurately identifying strawberry ripening stages in complex environments. Key enhancements include integrating the ECA mechanism to better capture strawberry features and employing focal-EIOU loss to address sample imbalance. The proposed model showed an accuracy of 97.81%, a recall of 96.36%, and an F1 score of 97.07%. These results confirm the model's effectiveness in reliably classifying ripeness stages, providing a foundation for developing optimized, path-planning strategies for fruit-picking robots to enhance harvesting efficiency. Yongkuai Chen et al. [8] addressed the challenge of balancing accuracy and performance in strawberry ripeness detection for automated harvesting robots. It introduces an improved CES-YOLOv8 model, which incorporates ConvNeXt V2 modules and ECA attention to enhance feature diversity and generalization. Data was collected under varied lighting, occlusion, and angle conditions to ensure real- world applicability. The enhanced model achieved 88.20% accuracy, 89.80% recall, 92.10% mAP50, and 88.99% F1 score—improvements over the original YOLOv8.

### III. EXISTING METHODOLOGY

This chapter describes an overview of existing methods for object detection, highlighting their strengths and limitations. In addition, this chapter explain the suggested techniques and process, emphasizing improvements over traditional techniques.

**Deep Learning:** Neural networks are used in deep learning, a subfield of machine learning, to model intricate data patterns. In this work, a transfer learning approach is applied by fine-tuning a pre-trained CNN model-VGG16-for classifying the maturity stages of papayas. VGG16 is well-known for its straightforward architecture, comprising sequential 3×3 convolutional layers and max pooling for dimensionality reduction. CNNs are commonly used in image classification, object detection, recommendation systems and even natural language

processing. A typical CNN consists of four main types of layers: convolution, activation, pooling and layer that is completely connected.

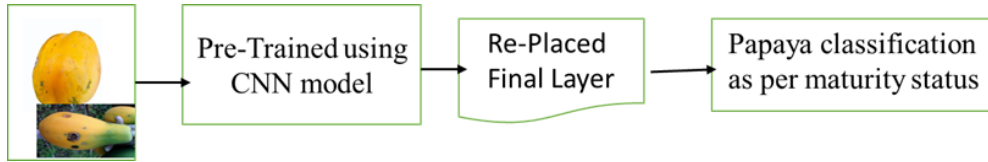


Figure 2 Papaya classification step using CNN model.

**Convolution Layer:** This layer applies filters (kernels) over the input image to extract essential features. It performs dot product operations between the filter and image pixels, enabling the detection of patterns such as edges and textures.

**ReLU Activation Layer:** The Rectified Linear Unit (ReLU) introduces non-linearity to the model. It retains only positive values in the feature maps, setting negative values to zero, which helps the model learn more complex patterns.

**Pooling Layer:** Pooling reduces the spatial dimensions of the feature maps while preserving important information. It helps lower computational cost, reduces parameters, and mitigates overfitting without affecting depth.

**Fully Connected Layer:** In this layer, every neuron is connected to all outputs from the previous layer. It applies a linear transformation using a weight matrix and is typically used at the final stage for classification.

#### IV. PROPOSED METHODOLOGY

##### 4.1 YOLO V8 suite for papaya image processing

YOLOv8 (You Only Look Once version 8) is an advanced real-time object detection model that offers improved accuracy and faster processing compared to earlier versions. Its high performance makes it well-suited for a wide range of applications, including fruit detection, surveillance, robotics, autonomous driving, and beyond.

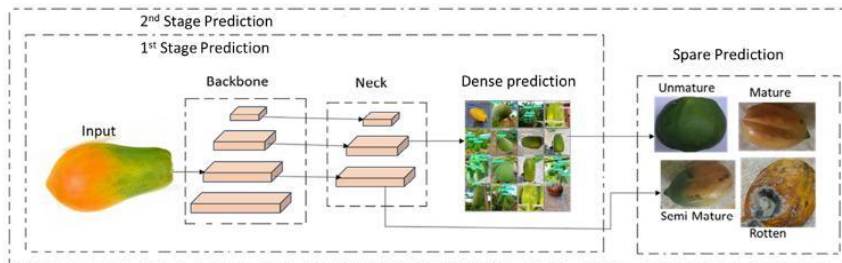


Figure 3 Object detector anatomy.

The diagram [Fig. 3] depicts the architecture of the YOLOv8 object identification model as applied to papaya detection and classification. It describes a detailed overview of the model’s key components, including the backbone, neck, first-stage predictor, second-stage predictor, as well as the concepts of dense prediction and sparse prediction.

##### First stage predictor

The first stage predictor generates initial, coarse predictions by thoroughly scanning the entire image for possible object locations. It uses dense prediction, evaluating every spatial point (such as each pixel or region) to identify potential objects. This ensures that even small or obscured objects in complicated backgrounds are detected. The first stage outputs preliminary bounding box proposals, which are further refined in later stages.

##### Second Stage Predictor

The second stage predictor builds upon the first stage by focusing on specific regions of interest and performing sparse prediction. It analyses smaller, filtered subset of regions identified as potential object candidates in the first stage. This selective focus enhances bounding box accuracy and improves classification by concentrating computational resources on high-confidence areas. By refining the predictions and reducing false positives, it ensures that precise identification of papaya types. This architecture allows YOLOv8 to accurately detect and

classify papayas, differentiate healthy from diseased fruit and manage objects of varying sizes and complexities within the scene.

#### 4.2 Process flow diagram

YOLO models are designed for object detection tasks and very effective for real-time object detection. Dynamic training of a Papaya detection system using YOLOv8 with a webcam can be achieved by following a systematic set of steps.

1. Capturing Images from the Webcam
2. Labelling and Annotating the Data in Real-Time
3. Training the YOLOv8 Model on the Captured Data
4. Detecting papayas in Real-time using the trained model
5. Improving the model continuously (Dynamic Training)

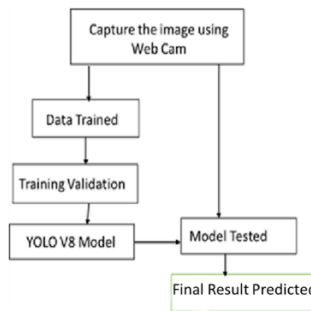


Figure 4 Block Diagram of papaya fruit maturity classification.

Figure 4 illustrates the workflow of an image prediction system using the YOLO V8 model, commonly applied for object detection or classification. The process is as follows:

1. **Image Capture:** A live image is captured through a webcam to serve as input.
2. **Data Training:** Collected images are used to train the model, building its ability to recognize patterns.
3. **Validation:** The model’s performance is validated to ensure accuracy and prevent overfitting.
4. **YOLO V8 Model:** The YOLO V8 model is developed and refined through training.
5. **Model Testing:** The trained model is tested using new images, often also from the webcam.

**Prediction Output:** The system provides the final prediction, identifying or classifying objects within the image.

#### V. RESULTS AND DISCUSSION

This chapter describes the results drawn from proposed work and discussed the comparisons of various models. Table 1 shows the comparison between existing model and proposed model, whereas, Table 2 describes the comparison results of existing and proposed models.

**Table 1** Comparison between CNN and YOLO v8 model

| Model/Feature       | CNN (Existing Model)                                                          | YOLOv8 (Proposed Model)                                                                            |
|---------------------|-------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| Purpose             | Used for image classification and feature extraction.                         | Specifically designed for real-time object detection.                                              |
| Architecture        | Consists of multiple convolutional layers followed by fully connected layers. | Integrates backbone, neck and head for feature extraction, aggregation and detection respectively. |
| Speed               | Slower for detection tasks for localization.                                  | Fast and efficient, capable of processing images in real-time (up to 60 FPS or more).              |
| Output              | Produces class probabilities for entire images.                               | Provides bounding boxes, class probabilities and confidence scores for each object in an image.    |
| Training Complexity | Requires separate models for classification and detection.                    | End-to-end training for both detection and classification tasks in a single model.                 |
| Inference           | Needs additional processing to convert features into bounding boxes.          | Directly outputs bounding boxes and class predictions from a single pass network.                  |

|           |                                                                                           |                                                                                                         |
|-----------|-------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| Use Cases | Best suited for image classification tasks.<br>Ex: identifying objects in a single class. | Ideal for applications needed real-time object detection.Ex. surveillance, smart vehicles and robotics. |
|-----------|-------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|

Table 2 Comparison results of CNN model and YOLO v8 model

| Particulars        | Existing System | Proposed System |
|--------------------|-----------------|-----------------|
| Data Set           | Fruit 360       | Fruit 282       |
| Collected Data Set | Kaggle          | Kaggle          |
| Model              | CNN             | YOLO v8         |
| Image Size         | 150×150         | 640 ×640        |
| Training Accuracy  | 92.97%          | 95.30%          |
| Testing Accuracy   | 85.12%          | 99.81%          |

Precision, recall, precision-confidence and F1 score curves are vital for evaluating the efficiency of a YOLOv8 model in fruit detection. Precision measures the accuracy of positive predictions, recall gauges the model's ability to capture actual instances, precision- confidence reflects the trade-off between confidence thresholds and precision and the F1 score balances precision and recall. Table 3 depicts outcome obtained from proposed work and acquired the necessary outcomes.

Table 3 Results obtained

| Maturity Levels   | Accuracy Values |
|-------------------|-----------------|
| Mature            | 0.995           |
| Semi Mature       | 0.916           |
| Unmature          | 0.692           |
| Rotten            | 0.995           |
| Precision Recall  | 0.899           |
| Recall Confidence | 0.930           |
| F1 Confidence     | 0.880           |

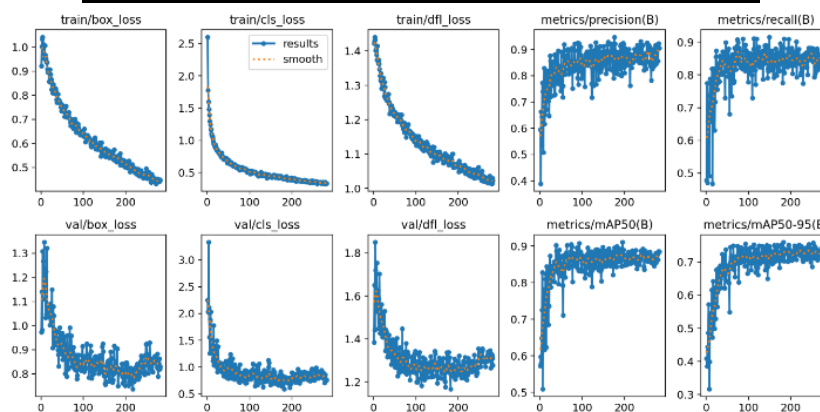


Figure 5 Performance analysis and validation

Figure 5 illustrates the performance analysis and validation obtained from proposed method. With rigorous training, validation, and testing on a real time dataset, the model successfully identified different ripeness levels, showcasing its potential for real-world agricultural applications.

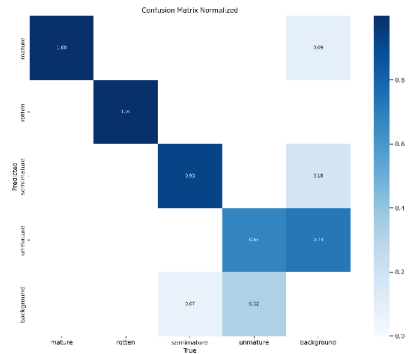


Fig 6(a): Confusion Matrix

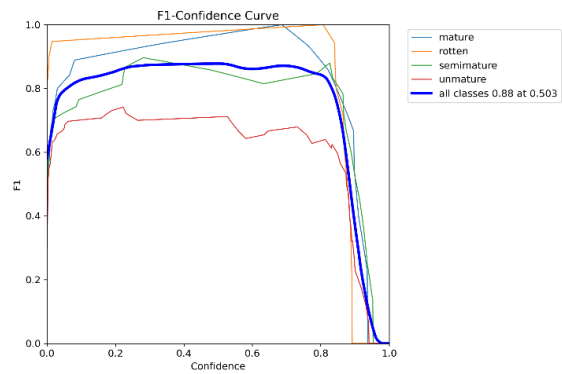


Fig 6(b): F1 Confidence curve

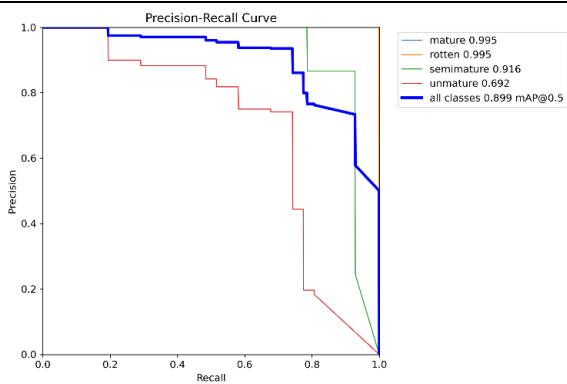


Fig 6(c): Precision Recall curve

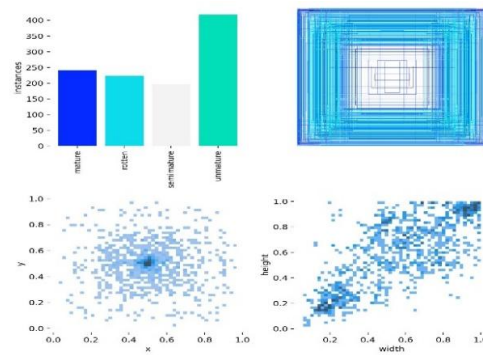


Fig 6(d): Scatter graph

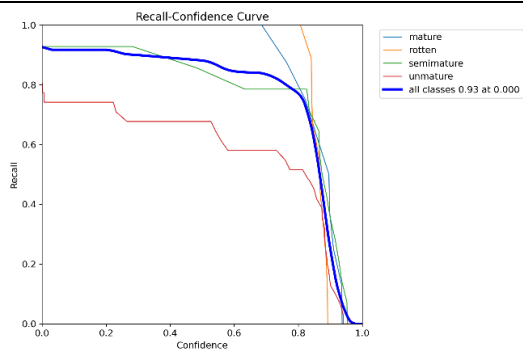


Fig 6(e): Recall Confidence curve

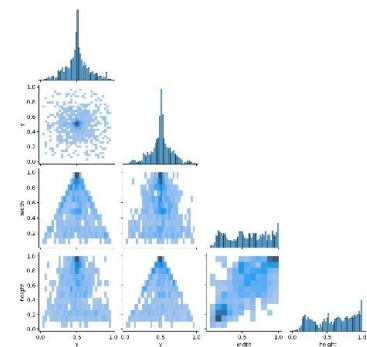


Fig 6(f): correlation graph

Figure 6 Graphs and analysis

While evaluating a classification model, several key performance metrics helped to understand its effectiveness of precision, recall, F1-score and the confusion matrix as shown in Figure 6.

## VI. CONCLUSION AND FUTURE SCOPE

### Conclusion

Real-time classification of papaya maturity stages using image processing can revolutionize agricultural practices by enabling precise harvesting decisions, minimizing waste and ensuring optimal fruit quality. This study introduced a deep learning-based approach leveraging the YOLOv8 architecture to classify papaya maturity stages accurately. With rigorous training, validation and testing on a custom dataset, the model successfully identified different ripeness levels, showcasing its potential for real-world agricultural applications. Obtained various

classification results from proposed method are **0.995**, **0.916**, **0.692** and **0.995** for mature, semi mature, unmaturing and rotten respectively.

However, challenges such as high computational demands and model generalization remain key concerns. Deploying deep learning models in real-time, especially in resource-limited agricultural environments, requires computationally efficient architectures. Additionally, ensuring the model's adaptability to varying lighting conditions, environmental factors, and different papaya varieties is crucial for practical implementation.

#### **Future scope**

According to the current proposed work, still more work to be done through advanced techniques available in market.

- It is recommended to identify diseases in papayas that affect both the plants and fruits. Common symptoms include mottled leaves, yellowing, and ring spots on the fruit.
- Further, research is recommended to explore the utilization of additional features in determining the quality of papaya fruit. This work may furthermore have extended with the development of a prototype and integrate in fruit industry to make fast sorting process.

#### **VII. REFERENCES**

- [1]. Al-Masawabe, M. M., Samhan, L. F., AlFarra, A. H., Aslem, Y. E., & Abu-Naser, S. S. (2021). Papaya Maturity Classifications using Deep Convolutional Neural Networks.
- [2]. Nada, AM Abu, et al. "Age and gender prediction and validation through single user images using CNN." *Int. J. Acad. Eng. Res. (IJAER)* 4 (2020): 21- 24.
- [3]. Mohammed M. Abu-Saqer, Mohammed O. Al-Shawwa (2020) "Grapefruit Classification Using Deep Learning" *International Journal of Academic Information Systems Research (IJAISR)* 4(1), 1-5
- [4]. Zihui Wei, Z., Chang, M., & Zhong, Y. (2023). Fruit Freshness Detection Based on YOLOv8 and SE attention Mechanism. *Academic Journal of Science and Technology*, 6(1), 195-197.
- [5]. Santi Kumari Behera, S. K., Rath, A. K., & Sethy, P. K. (2021). Maturity status classification of papaya fruits based on machine learning and transfer learning approach. *Information Processing in Agriculture*, 8(2), 244-250.
- [6]. Jairo Lucas de Moraes, J. L., de Oliveira Neto, J., Badue, C., Oliveira-Santos, T., & de Souza, A. F. (2023). Yolo-papaya: a papaya fruit disease detector and classifier using cnns and convolutional block attention modules. *Electronics*, 12(10), 2202.
- [7]. Chenglin Wang, Y., Xu, H., Chang, P., Huang, Y., Zhong, F., Jia, Q., & Liu, S. (2024). CES- YOLOv8: Strawberry Maturity Detection Based on the Improved YOLOv8. *Agronomy*, 14(7), 1353.
- [8]. Brucal, S. G. E., Colobong, J. B., De Guzman, D. B., De Leon, J. N. M., Samaniego, L. A., & Yong, E. D. (2021, July). Papaya Maturity Classification in MATLAB Platform using Lab Method and DHT11 Sensor. In *2021 Fifth World Conference on Smart Trends in Systems Security and Sustainability (WorldS4)* (pp. 171-175). IEEE.