INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 7, July 2026 | ISSN: 2456-4184 | IJNRD.ORG

Vegvison: A Computer Vision Framework for Real – Time Vegetable Quality Inspection and Defect Detection

[1]Ms. Deepika Dattatray Lalge, [2] Dr. Shankar S. Pujari

[1]Student, [2] Assistant Professor

[1]Department of Computer Science & Engineering (Data Science), D.Y. Patil Agriculture & Technical University, Talsande, Kolhapur, Maharashtra, India

[2]Department of Computer Science & Engineering (Data Science), D.Y. Patil Agriculture & Technical University, Talsande, Kolhapur, Maharashtra, India

Abstract : Quality assurance of fresh fruits and vegetables is a major issue in agriculture, whereby late detection leads to wastage of produce and financial losses. In this work, an automatic system called VegVision is presented, which uses object detection algorithms and deep learning models to detect the freshness status of the produce. Specifically, the YOLOv8 neural network algorithm was implemented for the identification and classification of different fruit and vegetable categories in both fresh and spoiled forms. The model was trained and tested with a dataset of annotated pictures consisting of 20,888 images belonging to twenty classes . The developed system is designed to operate in real-time mode as part of a web application, allowing for the uploading of pictures or performing the analysis based on the camera input. This way, a user is able to determine the quality of the produce without the need for direct examination and special skills. In addition to identification, the system estimates the freshness status in confidence terms, providing recommendations on the further course of action. Experimentation showed reliable results with satisfactory precision and recall scores, indicating its applicability for practical usage.

***IndexTerms* - Freshness Detection, Computer Vision, Deep Learning, Fruit Quality Assessment, Vegetable Quality Assessment, Object Detection, Food Waste Reduction, Real-Time Monitoring.**

INTRODUCTION

The maintenance of the quality of fruits and vegetables through its entire lifecycle is an issue that cannot be overlooked today. Fruits and vegetables constantly change biologically and environmentally from the moment of picking, thus posing the risk of damage to the produce when transporting, storing, and distributing them. As a consequence, a considerable amount of fruit and vegetable produce ends up being discarded, which entails both financial losses and extra pressures on food chains. Hence, there has been a need to develop an accurate method of assessing the state of fresh produce. As far as conventional approaches to testing fruit and vegetable freshness are concerned, they primarily rely on visual inspections carried out by humans. Surface color, texture, bruising, and other factors are often employed in estimating the quality of produce. While this approach is rather common, its results may vary based on personal experience of the inspector and conditions of the environment in which it takes place.

The fast development of AI has brought about some innovative ways of applying it in automating the visual inspection process. For example, computer vision has shown itself to be effective in acquiring useful information from digital images. These systems utilized hand-crafted features for color, shape, and texture. Although such solutions worked relatively well in laboratory conditions, it became challenging to apply them in practice due to the varying appearance and backgrounds of objects under observation. Deep learning has become highly successful at extracting representative features from data. Current object detection models can perform object recognition and localization at once. One such model that stands out among others is the set of the YOLO models which is known to produce accurate results while maintaining very little computational latency. The recently developed architecture YOLOv8 offers architectural adjustments that make it more suited for the task of detection. In spite of the numerous studies related to automated quality assessment of produce, the majority of the approaches have been tested using only several categories of fruits and vegetables, under artificial imaging scenarios, or just for classification purposes. Moreover, decision-support abilities of such systems remain unexplored, thus reducing their potential applicability in practical activities. Therefore, it becomes necessary to develop solutions that could solve both object identification and provide recommendations based on the detected characteristics of the produce to be analyzed. In order to cope with these obstacles, the present paper offers an intelligent system for automatic quality assessment based on the application of YOLOv8 object detection. The training data included 20,888 annotated images, featuring different classes of fresh and spoiled produce. Apart from the identification of objects and assessing their freshness, the proposed model was supplemented with a web platform, allowing to upload pictures and analyze images taken from the webcam in real time. The main contributions made by this research include: Developing an automated freshness inspection system able to classify various fruits and vegetables as fresh and spoiled in twenty different classes. Implementation and fine-tuning of a fruit & vegetable detection model based on YOLOv8 architecture with a large-scale training dataset consisting of 20,888 images labeled accordingly.

Integration of the trained model in a web application for analyzing static images and capturing video frames from the camera for making predictions. Implementation of additional functions, such as freshness score assessment and advice, apart from the basic detection functionality. Validation of the proposed system with the standard metrics used in object detection problems like Precision, Recall, F1 - score, and mAP.

METHODOLOGY.

This chapter provides an overview of the procedure utilized during the creation of the VegVison model, including the preparation of the dataset, model training, evaluation methodology, and deployment process..

2.1Dataset Description

The dataset employed in the experimental analysis is based on the Roboflow Universe dataset and consists of photos depicting various fruits and veggies captured in varying environmental conditions. All the images in the dataset were labeled using bounding boxes that indicated the fruit or vegetable type, as well as their freshness condition according to the YOLO model standard.The fruits and veggies dataset used in the experiment comprised ten categories: apples, bananas, bell peppers, carrots, cucumbers, mangos, oranges, potatoes, strawberries, and tomatoes. Within each produce category, two quality conditions – fresh and rotten – were included, which made for a total of twenty detection classes. The model development was performed with the help of 20,888 images labeled in accordance with the above criteria. The dataset was divided into training (18,333 images), validation (1,704 images), and test datasets (851 images).

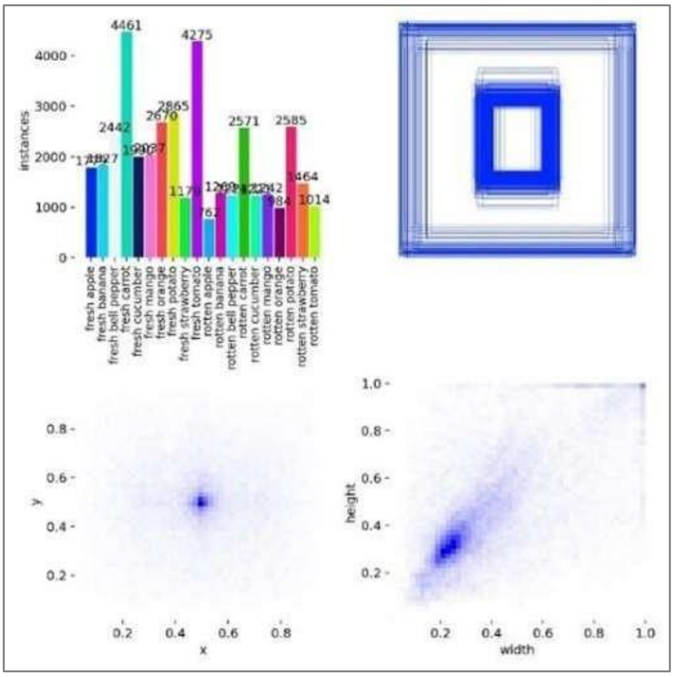

Fig. 2.1. (upper-left) Distribution of instances per class in the VegVision training set for all 20 categories, (upper-right) An example of bounding box labels, (lower-left) Spatial distribution of bounding box centers, and (lower-right) The distribution of the aspect ratio of the bounding box.

As shown in Figure 1, the dataset has some distinct features in terms of class distribution, label illustration, the pattern of object locations, and bounding box size. The detected variability in the scales and orientation of objects, along with their placement, enhances the realism of the dataset in practice.

2.2 Image Pre-Processing and Data Augmentation

All images were converted to 640 × 640 before training, ensuring compatibility with the chosen detector architecture. In order to make the network more robust, several augmentations were performed in the course of training. They included the application of such transformations as flipping, scaling, translating, changing the colour and brightness of an image, and random erasing. Additionally, mosaic augmentation was employed to compose training samples from different images into a single one. Through image pre-processing, the training set became more visually diverse and allowed exposing the model to a larger variety of objects' appearances.

2.3 Detection Model Configuration

In this regard, the VegVison detection approach employs the YOLOv8n version of the YOLO neural network as the main detector. The choice of the above detector was mainly influenced by its lightweight nature and compatibility with limited computational hardware.

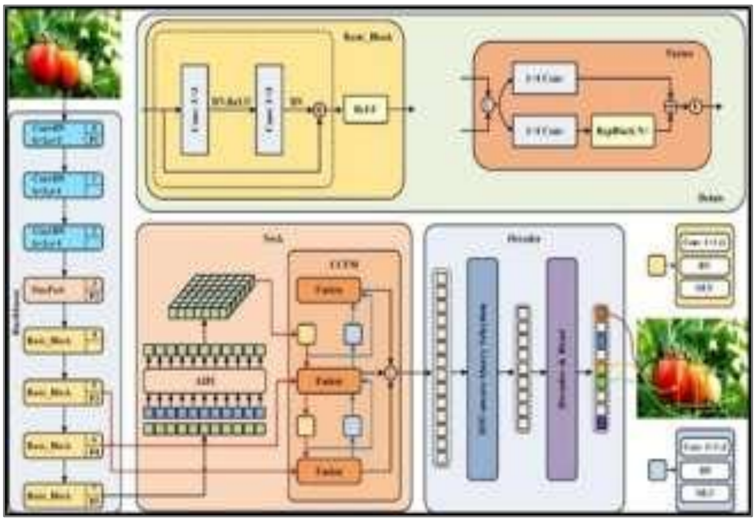

Fig. 2.2. Architecture of the YOLOv8n Model Used in VegVison.

The input images to the model undergo processing through numerous feature extraction phases. After that, the obtained information on the various scales is then merged to ensure accurate detection for the various object sizes. Object coordinates, probabilities, and associated freshness levels are eventually determined by the prediction layers. As opposed to training from scratch, the model makes use of transfer learning and incorporates pre-trained YOLOv8n parameters into the VegVison detector. In this case, fine-tuning on the VegVison dataset helped the model learn features associated with object freshness.

2.4 Training Procedure

Model training was done using Python programming language and deep learning framework PyTorch. The computation was optimized using GPU through CUDA to accelerate experimentation processes. Model training was done using the following training parameters:

- Input image size: 640 × 640 pixels
- Epochs: 30
- Batch size: 2
- Learning rate: 0.0004
- Optimizer: Adaptive optimizer
- Loss functions: Localization loss function, classification loss, bounding box regression loss

The Auto Mixed Precision(AMP) was used to reduce memory usage for training without impacting the efficiency. Early stop technique was also implemented to prevent unnecessary computations once the efficiency stabilizes.

2.5 Performance Assessment Criteria

Performance evaluation of the proposed object detection model was conducted based on common evaluation measures. In the evaluation process, the following criteria were considered:

- Precision
- Recall
- F1 score
- Mean Average Precision(mAP) with Intersection Over Union (IoU) equal to 0.5 (mAP@0.5)

2.6 Deployment Framework

The deployment framework comprises five core elements:

- User interaction layer
- Backend processing module
- YOLOv8-based inference engine
- Freshness assessment module
- Result visualization interface

The trained model was deployed into a web application developed using Django. Users may submit images either via the upload feature or through an attached camera for live detection. Once the image is received, the backend conducts inference, detects objects, assigns probabilities of freshness, and visualizes the output on the dashboard.Deployment provides a nearly real-time process that offers further recommendation-based information about the quality of produce and its storage and use potentials.

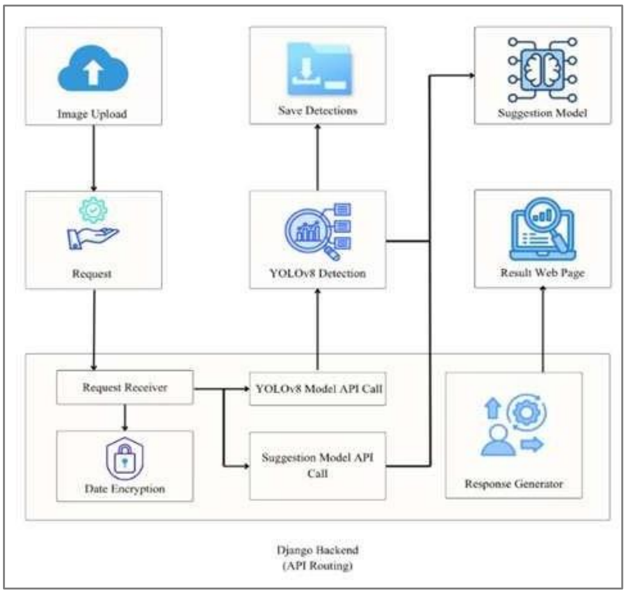

Fig. 2.3**.** System architecture of VegVison.

RESULTS AND DISCUSSION

Here are the results acquired by applying VegVison on training, validation, and testing phases. The results of the experiment were evaluated via conventional object detection metrics and performance measurements to assess the practical efficiency of the algorithm.

3.1Learning Behavior During Training

As seen from Figure 4, the process of learning went well through all the 30 epochs of the optimization. There is a consistent decline in the loss values for training and validation sets, which shows that the neural network managed to extract discriminative features necessary for proper classification of fresh vegetables. The validation results showed a steady improvement in the beginning and stabilization after achieving convergence. The major performance increase happened in the first and middle epochs with minor improvements during the last ones. This result proves that the used training regime worked effectively with no signs of overfitting.

.

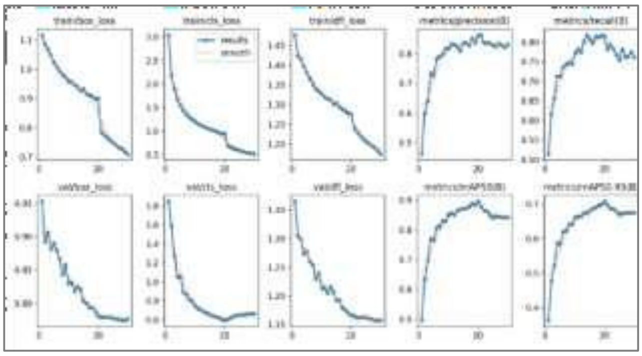

Fig. 3.1. Loss values in training and validation and performance metrics progress during 30 epochs of VegVison YOLOv8n.

3.2 Overall Detection Performance

As seen from Figure 4, the process of learning went well through all the 30 epochs of the optimization. There is a consistent decline in the loss values for training and validation sets, which shows that the neural network managed to extract discriminative features necessary for proper classification of fresh vegetables. The validation results showed a steady improvement in the beginning and stabilization after achieving convergence. The major performance increase happened in the first and middle epochs with minor improvements during the last ones. This result proves that the used training regime worked effectively with no signs of overfitting.

Table 3.2: Descriptive Statics Overall VegVison Model Performance on Validation Set

Metric	Value
Images evaluated	1,704
Total instances	3,721
Precision	0.859
Recall	0.817
mAP@0.5	0.896
mAP@0.5–0.95	0.708
Preprocessing time	2.3 ms/image
Inference time	59.4 ms/image
Post-processing time	2.7 ms/image
Total processing time	~64.4 ms/image
Model size	6.2 MB

In regard to the final object detection model, it showed excellent object detection efficiency for all twenty classes of freshness categories. The precision of 0.859 and recall of 0.817 have been obtained using the validation dataset. These results reveal that the object detector was able to classify fruits and vegetables effectively and reliably. The achieved mAP@0.5 of 0.896 reveals the effectiveness of both localization and classification abilities. In addition, the mAP@0.5-0.95 metric of 0.708 reveals that the model is capable of stable behaviour at different thresholds. In respect of model usability, it is necessary to note that the mean inference time is approximately 64.4 ms per frame. Such processing rate is equivalent to nearly 15 frames per second. Therefore, it means that the framework will allow implementing the applications in a quasi-realistic environment. Finally, the small size of 6.2 MB allows deploying this object detection model on constrained devices.

3.3 Class Level Prediction Analysis

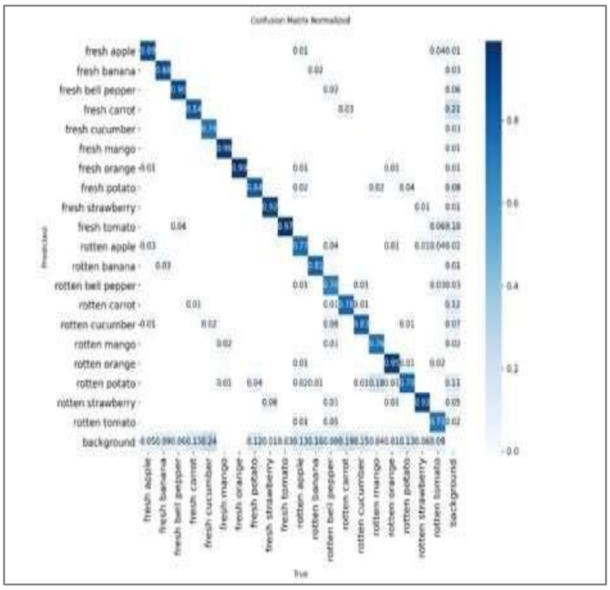

Fig. 3.2. Confusion matrix normalization for VegVison in all 20 classes and background in the validation set.

The confusion matrix depicted in Fig. 5 gives an in-depth look at the behaviour of predictions in all classes. Most classes show high values along the diagonal, meaning that most images have been classified in their respective categories.

Errors have mainly been committed while categorizing classes with visually similar freshness states, especially those with slight differences between fresh and spoiled classes. Despite the difficulties posed, it is clear that the model was able to maintain good separation of classes. The effectiveness of the training approach can be demonstrated through the confusion matrix, where classes are able to be categorized correctly despite the multiple fruit and vegetable classes.

3.4 Precision-Recall Characteristics and Threshold Optimization

This section elaborates the proper statistical/econometric/financial models which are being used to forward the study from data towards inferences. The detail of methodology is given as follows.

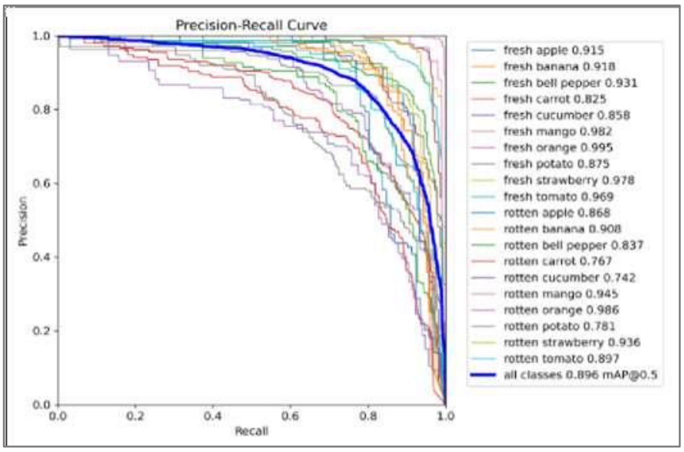

Fig. 3.3**.** Precision-Recall curve with AP@0.5 values for all 20 classes and macro-averaged mAP@0.5 equal to 0.896.

The relationship between the detection precision and recall as shown in Fig. 6 represents a balance between the quality and extent of object detections. The macro-averaged mAP@0.5 value of 0.896 demonstrates consistency in results throughout the tested classes. The nature of the Precision–Recall curve demonstrates high levels of detection precision, combined with good recall, providing a good compromise suitable for the purposes of detecting fresh objects.

3.4.1 F1-Score vs Confidence Analysis

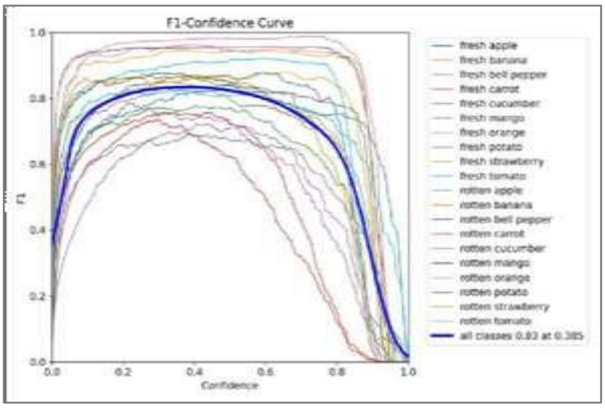

Fig. 3.4.1**.** F1-score-Confidence relationship demonstrating the peak macro-averaged F1-score of 0.83 when a confidence threshold equals 0.385.

The analysis of the F1-score and confidence level allows determining the best confidence level to obtain a good compromise between precision and recall. From Fig. 7, the peak of the curve, representing the best F1-score, corresponds to the threshold of approximately 0.385.

3.4.2 Precision Trend Across Confidence Levels

As seen from Figure 8, the more limiting the confidence threshold is, the greater the precision in prediction will be. When the confidence is very high, the system achieves absolute precision, which means that the predicted items are highly trustworthy. This approach could prove helpful for systems where reduction of false positives is especially important.

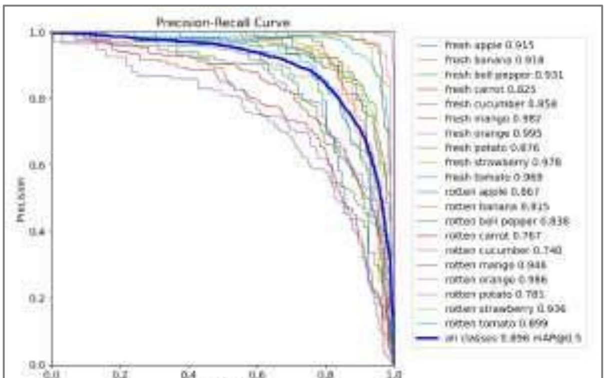

Fig. 3.4.2 Precision-Confidence chart illustrating precision of 1.00 achieved at highest possible confidence threshold of 1.000.

3.4.3 Recall Behavior Across Confidence Levels

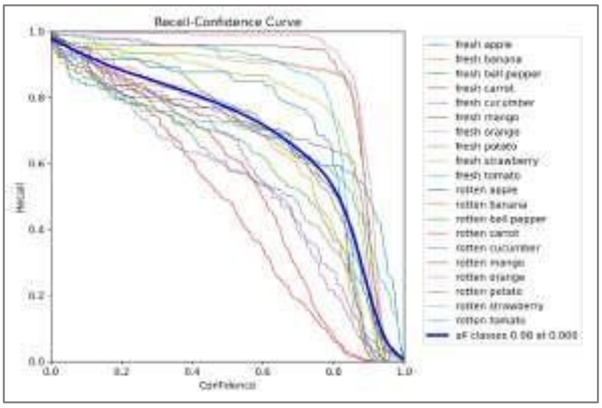

Fig. 3.4.3. Recall-Confidence chart of VegVison on all classes of product freshness

As can be seen from Figure 9, recall steadily decreases with increasing confidence threshold. With lower thresholds, the detection of target items will be easier, while higher thresholds result in more accurate predictions.Such a compromise helps choose the mode of operation depending on application requirements.

3.5 Practical Application Assessment

This framework has possibilities of application in various aspects of the agriculture and food supply chain.
Some of them are listed below:

- Automatic testing of products in stores.
- Ensuring the quality of goods before they enter the cold storage.
- Monitoring the freshness of food products in local markets and small-scale distribution canters.
- Help to vendors in getting quick information about the quality of their products without any additional tools.

3.6 Deployment Validation

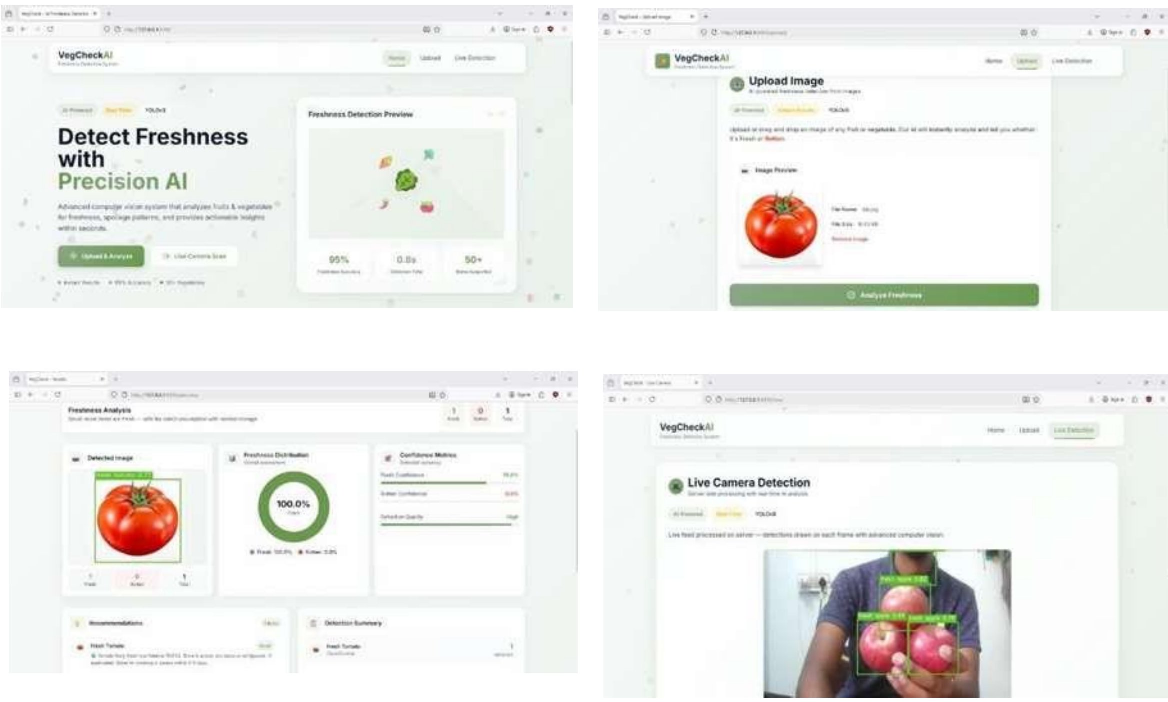

Fig. 3.5 VegVison web application interfaces: (a) Landing page (b) Upload interface (c) Results dashboard (d) Live detection interface

To test the usability of the developed system, it was implemented into a web application able to handle both uploaded images and live video streams. It has been shown that the deployment system can carry out the process of object detection, freshness estimation, and visualization of results in one sequence of actions. It is possible for the users to receive results on product freshness along with confidence level information and recommendation-related data. The response time has shown that the system is able to work effectively in real-life settings.Thus, it has been proved that VegVison is not just an academic prototype but a full-fledged solution for checking the freshness of products.

IV. CONCLSUION

VegVison is an advanced machine learning-powered framework proposed in this study to address the problem of automated fruit and vegetable quality assessment. This paper investigated whether object detection technologies could be integrated with a web-based user interface to design a real-world application for quality monitoring purposes. Thanks to the availability of a sizable dataset featuring twenty freshness categories, the framework was capable of discriminating fresh produce from those experiencing decay with high detection accuracy. The experimental results show that the trained detector managed to detect various produce classes without compromising on the localization and classification capabilities. As evident from the calculated precision, recall, and mean average precision metrics, the framework was consistent and accurate enough to analyse different types of produce samples. Moreover, the moderate inference speed ensures applicability in fast-paced environments.

In addition to its predictive power, the constructed system brings about practical advantages by virtue of its capability to handle uploaded images and camera feeds under one platform. Combining confidence levels in the prediction task along with recommendations allows users to benefit from relevant insights useful for daily quality assessment activities. The results of the current research demonstrate the possibility of using intelligent computer vision-based systems to help minimize unnecessary food losses and increase the efficiency of quality assurance processes in retail, warehouse and logistics environments. With the further development of computer vision techniques, such tools might become available even for individual users. Further improvements will center around increasing the diversity of datasets, adding more types of produce, optimizing the operation of the proposed models under difficult visual conditions, and studying the feasibility of deploying such systems in an edge and mobile computing environment. Researching grading, shelf-life determination and other multimodal techniques might be helpful in this case.

I. ACKNOWLEDGMENT

The successful completion of this project became possible through the educational environment and academic assistance offered by the Department of Computer Science and Engineering (Data Science), School of Engineering and Technology, D.Y. Patil Agriculture and Technical University, Talsande. Laboratory assistance along with computing resources played a significant role in developing and analysing the VegVison system. The authors would like to extend their gratitude to all faculty members and project supervisors who assisted them in the process of developing and improving this system with their observations and helpful suggestions. It would not be an exaggeration to say that without the cooperation of those who have made direct and indirect contributions towards this project's success, this paper could not have come into existence.

REFERENCES

[1] H. Altaheri, M. Alsulaiman, and G. Muhammad, “Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning,” IEEE Access, vol. 7, pp. 117115–117133, 2019.

[2] C. Arce-Lopera, T. Masuda, A. Kimura, Y. Wada, and K. Okajima, “Luminance Distribution as a Determinant for Visual Freshness Perception: Evidence from Image Analysis of a Cabbage Leaf,” Food Quality and Preference, vol. 27,pp. 202–207, 2013.

[3] D. M. Barrett and B. Lloyd, “Advanced Preservation Methods and Nutrient Retention in Fruits and Vegetables,” Journal of the Science of Food and Agriculture, vol. 92, no. 1, pp. 7–22, 2012.

[4] J. Gopal, H. N. Abdelhamid, J.-H. Huang, and H.-F. Wu, “Non-destructive Detection of the Freshness of Fruits and Vegetables Using Gold and Silver Nanoparticle-Mediated Graphene-Enhanced Raman Spectroscopy,” Sensors and Actuators B: Chemical, vol. 224, pp. 413–424, 2016.

[5] A. Kazi and S. P. Panda, “Determining the Freshness of Fruits in the Food Industry by Image Classification Using Transfer Learning,” Multimedia Tools and Applications, vol. 81, no. 6, pp. 7611–7624, 2022.

[6] M. Mukhiddinov, A. Muminov, and J. Cho, “Improved Classification Approach for Fruits and Vegetables Freshness Based on Deep Learning,” Sensors, vol. 22, no. 18, art. no. 6959, 2022.

[7] S. Shegar and P. Bhambu, “Fruit Classification Using Deep Learning and CNN for Fresh and Rotten Categories,” International Journal of Engineering Research and Technology, vol. 12, no. 3, pp. 77–81, 2023.

[8] Y. Yuan and X. Chen, “Vegetable and Fruit Freshness Detection Based on Deep Features and Principal Component Analysis,Current Research in Food Science, vol. 8, art. no. 100656, 2024.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.

[10] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv:2004.10934, 2020.