Journal of Applied Research and Technology

www.jart.icat.unam.mx

Journal of Applied Research and Technology 23 (2025) 164-177

Original

Design of an iterative method for enhanced fruit detection in smart farming using YoLoV9s with transfer learning analysis

C. Bhole[a]* • C. Joshi[b] • P. Chakrabarti[b] • A. Sharma[c]

[a]*Department of Computer Science and Engineering, KJSIT, Mumbai and Research Scholar at Sir Padampat Singhania University, Udaipur, Rajasthan, India*
[b]*Department of Computer Science and Engineering, Sir Padampat Singhania University, Udaipur, Rajasthan, India*
[c]*Department of Mathematics, Sir Padampat Singhania University, Udaipur, Rajasthan, India*

Received 05 16 2024; accepted 10 07 2024
Available 04 30 2025

Abstract: Being able to accurately detect and classify fruits on trees would be important for smart farming in terms of the efficiency of yield management and labour cost savings. Conventional methods are prone to inconsistency in environmental conditions, resulting in lower precision in detection accuracy. This paper proposes a state-of-the-art fruit detection technique through an approach to image augmentation enhanced by a genetic algorithm and advanced deep learning models. The proposed approach works for the detection of chickoo, mango, sweet lime and tomato fruit. This approach employs a genetic algorithm to optimize image augmentation, which enhances the diversification of training data and also increases the adaptability of the model towards various conditions. Fruit detection will be performed using the newest You Only Look Once (YoLoV9s) framework, which provides the possibility of real-time detection of objects. Classification of the fruits will be done with the help of a transfer learning-based VGGNet16 model. A colour thresholding step is applied for further confirmation of fruit types; it gives the least number of wrong classifications. It can be shown from the experimental results that the model that I have proposed accords a very good level of accuracy (97.9%), precision (97.0%), recall (97.5%), and an F1-score of 97.3%, which is much beyond the different methodologies produced in the literature up till now. The model's real capability, that is, its ability to accurately predict fruits under various environmental conditions reflects its potential to increase productivity in the precision agriculture process.

Keywords: Genetic algorithm, YoLoV9s, smart farming, fruit detection, transfer learning.

*Corresponding author.
E-mail address: chitra.bhole@spsu.ac.in (C. Bhole).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

Smart farming, through integration of artificial intelligence, has been a boon to agriculture. It includes monitoring, growing, and even harvesting of crops. Among the plethora of applications (Khan et al., 2024; McHenry & Aćimović, 2024; Navarro Soto et al., 2024), fruit detection using computer vision and deep learning is one of the very pivotal innovations. The same seeks to augment yield assessments and optimize harvesting processes. Despite significant advancements, the real-world environment in outdoor farm landscapes remains quite challenging to the efficacy and accuracy of most developed fruit detection systems. Such variations in lighting (Eun et al., 2024; Stefi et al., 2024; Xia et al., 2024), occlusions, and the physical diversity (Mazurek et al., 2024; Morrone et al., 2024; Xiao et al., 2024) of fruits often make it a task for many traditional methods of detection; hence, it requires better and more robust, adaptive solutions. This paper describes an advanced, iterative method to address this issue by incorporating the strengths of genetic algorithms (GA) for image augmentation with the state-of-the-art abilities of the You Only Look Once version 9 (YoLoV9s) detection system. Genetic algorithms are put into operation to artificially augment a dataset of fruit images, hence simulating a much-wider range of environmental variants for the purpose of improving the model's generalization capabilities. Genetic algorithms thereby make the system more adaptive and less susceptible to overfitting, especially in a highly variable environment such as agriculture.

Immediately after the augmentation phase of the image is where YoLoV9s, one of the fastest and most accurate systems for detecting objects in images, is put into operation. The reason is that its high speed makes it very ideal for real-time applications, which is the very reason why it is very important to be used in smart farm scenarios, which are at the forefront of every action with timely information. In order to further improve the accuracy of the classification process, a transfer learning-based VGGNet16 model is used to process the detected objects. The reason behind this is that the model is therefore famed for its depth and robustness in image classification for adding an extra layer of verification through colour thresholding methods for making sure identification is both precise and reliable for different scenarios. The combination of the augmentation of the training data with a dual model detection and classification system provides marked improvements in key performance metrics, including accuracy, precision, recall, area under the curve, and specificity. A reduction in the time delay associated with fruit detection and classification also ensures more efficient farming operations, which the proposed system infuses with. The proposed methodology introduces an innovative process of iteration into the field of agricultural AI that not only extends the scope of such an application but also lays groundwork for future explorations into the use of hybrid artificial intelligence techniques in real-life environments. As a workable, improved alternative to the existing fruit detection systems, this work adds to the aim of achieving greater productivity and sustainability in agriculture.

A more sophisticated system for detecting fruits is inspired by the growing need of agriculture to adopt more sustainable and efficient ways of farming. Precision agriculture is a strategic move to monitor and control the variability inherent in farming operations. The variability in this case includes optimizing returns on inputs while conserving resources. Fruit detection systems, in particular, play a key role in many of the labor-intensive aspects of fruit farming, which includes yield estimation and harvest planning. Most of these processes currently require intensive human inspection and decision-making, though such methods are just not adequate in terms of time and are prone to error, which often results in significant losses in terms of yield and profitability. Moreover, most of the existing automated systems cannot handle such complex scenarios due to variation in light conditions, occlusions, and the inherent diversity in fruit appearances. All of these challenges point to the need for a far more advanced fruit detection system that can robustly handle the complexities of real-world farm environments.

The current studies on fruit detection conducted within the smart farming environment present with a number of shortcomings that hinder their effectiveness when applied to real applications. The main limitation is that classical models lack efficiency in handling the rich and complex visual conditions brought into the agricultural settings through varying lighting conditions, occlusions, overlapping fruits, and natural deformities. Most existing methods rely mainly on static image processing techniques or shallow machine learning models that fail to generalize properly on several environmental factors. Most of these models are, in addition, restricted by the small and limited datasets, which are central for training robust systems to detect fruits under challenging circumstances. This results in low precision and recall, with a very high rate of false positives and false negatives in fruit-related tasks.

Another serious limitation is the inability of existing detection systems to operate in real-time applications, which is critical for any practical deployment in smart farming. Many approaches lack the needed processing speed, or they sacrifice accuracy to achieve this speed. Conventional methods require multiple stages for image processing and object detection, which essentially increase the overhead of computation and reduce the system applicability for usage directly in the field for real-time applications. Most studies lack state-of-art image augmentation or optimization techniques like genetic algorithms, increasing variability in data and further aiding model performance. This gap in adaptively using learning mechanisms, the fact that state-of-the-art deep learning frameworks are underused, limits the scalability and efficiency of existing fruit detection systems in dynamic and uncontrolled farming environments.

The challenges involved in using such a system include false positives or the recognition of uninteresting objects by the detector or the difficulty of distinguishing the presence or absence of fruit. The current paper introduces a new detection and classification system, applied to increase significantly the accuracy and efficiency of fruit detection in smart farming applications. The main contribution of this paper are as follows,

- The application of a genetic algorithm with the YoLoV9s object detection system and the VGGNet16-based classification model.
- The GA is applied for augmentations in the training dataset in order to create many synthetic images with a wide range that mimic the potential types of variations and difficulties actually experienced in real farm scenarios. The method extends the size of the dataset and reduces overfitting, providing the model with a better capability for generalization in different environmental conditions.
- The use of YoLoV9s, which is known for its superior speed in detection and accuracy, allows for real-time processing that is needed for timely decisions in agriculture.
- This is followed by a classification with the VGGNet16 model, enhanced with colour thresholding techniques that guarantee a double verification process to provide a substantial improvement in model confidence in the accuracy of fruit type identification process.

Together with these advanced methodologies, significant improvements are obtained in the key performance metrics, such as accuracy, precision, recall, the AUC, specificity, and time delays commonly observed when using traditional fruit classification methods. This dual approach strongly consolidates the model's effectiveness in real-world deployment scenarios, and with this, a new benchmark has been set for the field of agricultural technology because hybrid artificial intelligence techniques are successfully applied to solve real-world high-impact problems in smart farming. The paper will be structured into the following sections: Section 1: This section would be "Introduction," in which the problem of fruit detection in smart farming would be introduced and outlined by thinking on the challenges that may appear using current methods. Section 2: Review literature of techniques around fruit detection. Section 3: In this section, we illustrate the proposed method from the integration of genetic algorithm-based image augmentation and YoLoV9s framework for real-time fruit detection along with fruit classification through VGGNet16 with transfer learning. Section 4 consists of the experimental setup, dataset, and evaluation metrics. Section 5 is a summary of results and takes into consideration adding an analysis of how well the proposed model performs as compared to available methods in terms of improvements in accuracy, precision, recall, and F1 score. To bridge this gap, Section 6 will present a working scenario of the model by applying it to real-world conditions. To summarize, Section 7 presents the conclusion of the paper, with the overall key findings and future directions to enhance the scalability and adaptability of the system in smart farming environments.

2. In-depth review of existing models

In recent years, significant advancements have been made in the development of species-specific detection techniques. McHenry & Aćimović (2024) introduced new species-specific real-time PCR assays for detecting *Colletotrichum* species responsible for bitter rot in apples. This method highlights the precision and sensitivity required for accurate pathogen detection, crucial for effective disease management in agriculture. The specificity of the PCR assays ensures that accurate identification of the *Colletotrichum* species can lead to targeted interventions, thereby reducing crop losses and improving yield quality.

The assessment of food quality has seen notable improvements with the integration of electronic nose technology. Navarro Soto et al. (2024) explored the influence of fruitiness on the quality assessment of virgin olive oils using electronic noses. This technology mimics the human olfactory system and has shown promise in providing consistent and objective quality assessments. The study underlines the potential of electronic noses in the food industry, particularly in quality control processes where traditional sensory evaluation methods may be subjective and variable in process. The dual role of herbal spices as food and medicine has been comprehensively studied by Khan et al. (2024). Their work on the microscopic authentication of commercial herbal spices provides a detailed analysis of the authenticity and quality of these products. Given the increasing demand for natural and organic products, ensuring the authenticity of herbal spices is critical for consumer safety and satisfaction. The study's findings are significant for both food safety and pharmacognosy, emphasizing the need for stringent quality control measures in the spice industry sets.

Multi-omics approaches have revolutionized our understanding of metabolic pathways in plants. Xia et al. (2024) conducted a multi-omics analysis on giant pumpkins, revealing distinct features in metabolism pathways that influence fruit size and colour. This integrative approach combines genomics, transcriptomics, proteomics, and metabolomics, offering a comprehensive understanding of the biological processes underlying phenotypic traits. Such insights are invaluable for breeding programs aimed at improving crop yield and quality. The impact of genetic mutations on fruit characteristics was investigated by Eun et al. (2024), who characterized a new citrus mutant with a unique fruit shape induced by gamma irradiation. By identifying specific selection markers using allele-specific PCR, the study provides a foundation for future breeding programs targeting desirable traits. This research underscores the potential of induced mutations in developing new fruit varieties with improved attributes. Stefi et al. (20204) presented eco-anatomical data for *Saponaria jagelii*, a species on the brink of extinction. Understanding the anatomical adaptations of this species to its environment is crucial for conservation efforts. The study highlights the importance of integrating anatomical and ecological data to develop effective conservation strategies for endangered plant species.

Morrone et al. (2024) examined the effect of chabazite zeolite foliar applications on olive fruit fly control, photosynthesis, and the quality of extra virgin olive oil. Their findings suggest that such applications can influence volatile organic compound emissions and enhance olive oil quality. This study contributes to sustainable agriculture by providing insights into alternative pest control methods that can improve crop quality and yield sets. The propagation methods of highbush blueberry plants were compared by Mazurek et al. (2024), who analyzed plants propagated in vitro and conventionally. The comprehensive analysis revealed differences in growth and development between the two methods, offering valuable information for optimizing propagation techniques in blueberry cultivation. The field of robotic harvesting has seen significant advancements, with Xiao et al. (2024) reviewing the latest developments in fruit and vegetable harvesting robots. These robots incorporate sophisticated algorithms and sensors to enhance harvesting efficiency and accuracy. Arikapudi and Vougioukas (2021) further explored the use of telescoping arms in robotic tree-fruit harvesting, highlighting the importance of linear fruit reachability under geometric constraints. These innovations are paving the way for increased automation in agriculture, reducing labor costs and improving productivity.

Machine learning applications have been extensively explored for various agricultural purposes. Hassan et al. (2024) developed a machine learning approach for the automatic disease and colour classification of olive fruits, employing support vector machines and artificial neural networks. This approach enhances the accuracy and efficiency of disease detection and fruit classification, contributing to better crop management. Similarly, Aldakhil and Almutairi (2024) utilized transfer learning for multi-fruit classification and grading, demonstrating the potential of deep learning models in agricultural applications. Non-destructive testing techniques have become increasingly important in assessing fruit quality. Yogarajan et al. (2022) developed an apple fruit quality detector using fiber optics and colour sensors, providing a non-invasive method for determining ripeness. Kojić et al. (2022) investigated the use of electrical impedance spectroscopy with protein-based edible foils for detecting the freshness of fruits and vegetables for different scenarios. These techniques offer reliable and non-invasive options for quality assessment in the food industry scenarios. Advanced imaging and sensing technologies have been employed to improve fruit detection and classification. Sharafudeen and Chandra (2023), Patel and Patil (2023) proposed a multimodal Siamese framework for accurate grade and measure estimation of tropical fruits, utilizing object detection networks. Patel and Patil (2023) enhanced convolutional neural networks for fruit disease detection and grading, integrating spatial pyramid pooling and support vector machines. These technologies enhance the precision of fruit quality assessment, contributing to better post-harvest management. Feature selection and optimization techniques have been explored to improve the performance of fruit recognition systems. Huynh et al. (2022) proposed a two-stage feature selection approach using adaptive particle grey wolf optimization for fruit recognition. This method improves the accuracy and efficiency of fruit classification systems, providing a robust solution for agricultural applications. Real-time monitoring and detection systems have been developed to enhance agricultural productivity. Suharjito et al. (2023), Khan et al. (2022) introduced a real-time oil palm fruit grading system using a modified YOLOv4 model and smartphones. This system enables accurate and timely detection of fruit ripeness, facilitating better harvest management. Tian et al. (2024), (Aiadi et al., 2022) developed a lightweight detection method for real-time monitoring of tomato growth, demonstrating the potential of YOLOv5s in agricultural applications.

The integration of deep learning models in agriculture has been extensively reviewed by Espinoza et al. (2023), Zhong et al. (2022), Bhole and Joshi (2023), Hiren et al. (2023). Their systematic literature review highlights the advances and future scopes in the application of convolutional neural networks (Hussain et al., 2024; Hussain & Aslam, 2024) and visual transformers in agriculture operations. These models have shown significant potential in improving the accuracy and efficiency of agricultural practices. Despite the significant advancements in agricultural technologies, several challenges remain. The complexity of implementing these technologies in diverse agricultural settings, the need for large datasets for training machine learning models, and the high costs associated with advanced sensors and robotic systems are some of the key challenges. Future research should focus on developing cost-effective solutions, improving the robustness and scalability of existing technologies, and exploring the integration of multi-modal data for comprehensive agricultural analysis. The literature reviewed highlights the rapid advancements in agricultural technologies, particularly in species-specific detection techniques, food quality assessment, robotic harvesting, and machine learning applications. These innovations have the potential to significantly enhance agricultural productivity, quality control, and sustainability. However, addressing the existing challenges and exploring new research directions are crucial for the continued development and successful implementation of these technologies in agriculture.

3. Design of the proposed model process

To overcome issues of low efficiency and high complexity of fruit detection, which is present in existing models, this section discusses the design of an iterative method for enhanced fruit detection in smart farming using YoLoV9s with transfer learning analysis. Initially, as per Figure 1, In the proposed design the integration of a genetic algorithm (GA) with image augmentation techniques plays a pivotal role in enhancing the robustness and accuracy of the You Only Look Once (YoLoV9s) detection system. The GA is utilized to optimize the parameters and strategies for image augmentation, thus increasing the variability and representativeness of training data, crucial for improving the model's ability to generalize across different environmental conditions and fruit appearances. The genetic algorithm starts with an initial population of augmentation strategies, each represented as a chromosome. These chromosomes encode various parameters such as rotation angles, scaling factors, shear intensities, and colour adjustments. The fitness of each chromosome is evaluated based on how well the augmented images help the YoLoV9s model achieve higher performance metrics like precision, accuracy, and recall.

The fitness function is defined as a weighted sum of precision, accuracy, and recall obtained from the classifier's performance on a validation set, which is augmented using the parameters encoded by the chromosome. Mathematically, fitness function $\boldsymbol{F}$ is expressed via Equation 1:

$$F = w1 \cdot Precision + w2 \cdot Accuracy + w3 \cdot Recall \qquad (1)$$

Where $w1$, $w2$, and $w3$ are the weights that balance the importance of each metric according to specific detection needs. Crossover is implemented by combining pairs of chromosomes, through a single-point or uniform crossover method, to produce offspring that inherit traits from both parents.

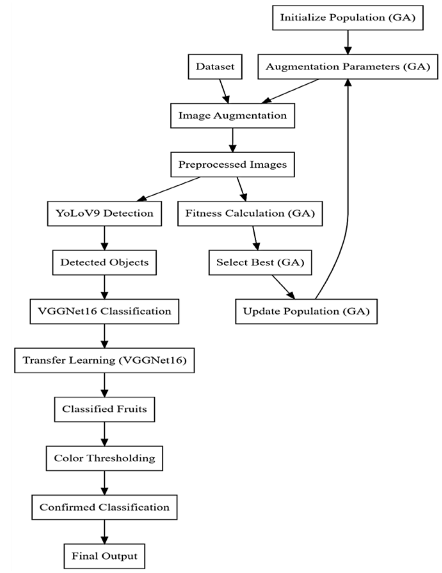

Figure 1. Model architecture of the proposed classification process.

Mutation stochastically alters one or more genes in a chromosome to introduce additional diversity into the population, which is essential for exploring new parts of the solution space. The mutation might involve slight changes in the rotation angle or colour adjustment parameters and samples. The selection process involves choosing the fittest chromosomes to be parents for the next generation. This is achieved using tournament selection or roulette wheel selection, ensuring that higher-fitness individuals have a higher probability of contributing to the next generation. The GA-optimized augmentation process dynamically adjusts the augmentation parameters throughout the training process. This approach helps in creating a robust dataset that mimics various environmental effects such as different lighting conditions, occlusions, and natural deformities in fruits. The justification for employing a genetic algorithm in conjunction with YoLoV9s stems from the need for a highly adaptive and efficient method to handle the complex and diverse nature of agricultural imagery. The GA allows for an iterative optimization of image augmentations, which is crucial for training deep learning models like YoLoV9s that require large amounts of varied data to perform well in real-world scenarios.

Additionally, the use of YoLoV9s for fruit detection is motivated by its speed and efficiency in processing images, which is vital in real-time agricultural applications where quick decision-making is crucial. The subsequent classification of fruits using a transfer learning-based VGGNet16 model leverages its depth and complexity for accurate classification, providing a robust two-phase detection system. In this work, the augmented image generation is done via Equation 2:

$$I' = f(I, \theta) \quad (2)$$

Where I is the original image, θ represents the augmentation parameters encoded by the GA, and I' is the augmented image by this process. The crossover operation is then performed via Equation 3:

$$\theta child = \alpha * \theta(parent, 1) + (1 - \alpha)\theta(parent, 2) \quad (3)$$

Where α is a blend factor determining the contribution of each parent's genes. Next, the Mutation Operation is performed via Equation 4:

$$\theta' = \theta + \delta \quad (4)$$

Where δ represents a small stochastic perturbation to the augmentation parameters for this process. The Fitness Evaluation is done via Equation 1, where the internal metrics are estimated via Equations 5, 6, and 7 as follows:

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

Finally, the feedback loop for GA optimization is modelled via Equation 8:

$$\theta next = Select\big(F(\theta)\big) \quad (8)$$

This operation represents the selection process where the next generation of parameters is chosen based on the fitness scores.

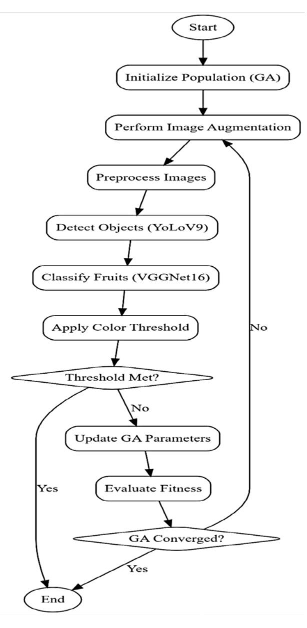

Figure 2. Overall flow of the proposed classification process.

These processes encapsulate the core computational and mathematical processes involved in the GA-driven augmentation strategy for the YoLoV9s-based fruit detection system process. The approach not only enhances the detection accuracy but also significantly contributes to the adaptability and efficiency required in smart farming contexts.

Next, as per Figure 2, the You Only Look Once (YoLoV9s) framework is integrated, which represents a sophisticated evolution in the realm of object detection systems, combining high-speed processing with an exceptional degree of accuracy. This deep learning architecture is specifically engineered to perform object detection in a singular evaluation of the image, distinguishing it from other models that might require multiple scans or stages. This inherent efficiency is pivotal for applications such as real-time fruit detection in smart farming, where rapid and accurate detection can significantly influence operational efficiency and output. YoLoV9s enhances its predecessors by optimizing both the architecture and the training process. The model processes images through a deep convolutional neural network, predicting both bounding boxes and class probabilities in one forward pass. This integrated approach ensures that YoLoV9s maintains minimal latency, making it ideal for real-time applications.

The network structure of YoLoV9s is designed to efficiently learn spatial hierarchies of features through successive convolutional layers. The layers are meticulously configured to capture fine-to-coarse granularities, where initial layers detect simple features like edges and textures, and deeper layers interpret complex features relevant to the objects of interest for different scenarios. Initially, the Bounding Box Prediction in YoLo is done via Equation 9:

$$b(x, y, w, h) = \sigma(tx) + c\big(x, \sigma(ty)\big) + c(y, p(w)e^{tw}), p(h, e^{th}) \quad (9)$$

Where σ represents the sigmoid function ensuring the outputs are between 0 and 1 for tx and ty, representing the center of the bounding box relative to the bounds of the grid cell. The terms pw and ph are the anchors for the box's width and height, predefined based on prior knowledge of object sizes, and etw, eth are predictions that scale the anchors to the actual object size. Next, the objectness score is estimated via Equation 10:

$$pobj = \sigma(c) \quad (10)$$

This operation computes the objectness score $p$obj which predicts the probability of an object being present in the bounding box. Function (c) represents the sigmoid activation of the model's confidence score c for this process. The class probability is then estimated via Equation 11:

$$pclass = softmax(c1, c2, \ldots, cn) \tag{11}$$

The softmax function is applied to the outputs $c1,2,\ldots,cn$ representing the class scores for each potential object category. This normalizes the scores, making them comparable across multiple classes. The Loss Function for this process is evaluated via Equation 12:

$$L = \lambda coord \sum_{i=0}^{S^2} \sum_{j=0}^{B} I^{ij}(obj)[(xi - x'i)^2 + (yi - y'i)^2] + \lambda obj \sum_{i=0}^{S^2} I^{i}(obj)(Ci - C'i)^2 \tag{12}$$

Where λcoord and λobj are coefficients to balance the importance of location and confidence predictions, I are indicators whether an object is present. Based on this, the feature integration process is modelled via Equation 13:

$$V = \int_{-\infty}^{\infty} f(x)dx \tag{13}$$

This mask, while abstract, symbolically represents the aggregation of learned features across the entire neural network, culminating in the final feature vector V used for detection. Finally, the optimization gradient is estimated via Equation 14:

$$\frac{\partial L}{\partial w} = Input \times (Output - Target) \tag{14}$$

This derivative is crucial for training via backpropagation, where ww represents the weights of the network, and the product of input and error term adjusts the weights to minimize the loss function L for this process. The choice of YoLoV9s for fruit detection is primarily motivated by its capacity to perform detections swiftly and with remarkable accuracy, essential for the dynamics of an agricultural setting where conditions can change unpredictably. The model's ability to evaluate the entire image in a single pass reduces latency, a critical factor in real-time applications where decisions must be immediate. Moreover, the scalability and flexibility of YoLoV9s, which allows it to be trained on highly diverse datasets encompassing various fruit types and environmental conditions, further justify its selection over other models which may not provide such robustness or speed. This method complements traditional approaches by providing a comprehensive system that not only detects but ensures the accuracy and reliability of each detection, essential for practical deployment in smart farming.

Once the YoLoV9s framework has effectively detected fruits within images, a transfer learning-based VGGNet16 model is employed to classify these detected fruits into specific types: chickoo, mango, sweet lime and tomato types. As per Figure 3, the choice of VGGNet16, renowned for its depth and architectural simplicity, is driven by its proven capabilities in image classification tasks, which stem from its methodical stacking of convolutional layers. Transfer learning further refines this process by utilizing a pre-trained network, enhancing learning efficiency and accuracy when adapting to the specific task of fruit classification. Transfer learning with VGGNet16 involves initializing the model with weights trained on a large, comprehensive dataset (ImageNet). This approach leverages learned features that are generalizable across various domains, significantly benefiting the classification task in environments with limited labeled data, such as specific types of fruits in agriculture operations. The feature extraction layer for VGG16 is represented via Equation 15:

$$Fl = ReLU(Wl * F(l-1) + bl) \tag{15}$$

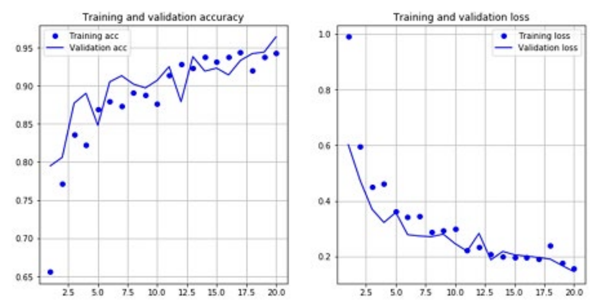

Figure 3. Training performance of the transfer learning process.

Where Fl represents the output feature maps of layer l, Wl and bl are the weights and biases, $*$ represents the convolution operation, and ReLU is the activation function that introduces non-linearity, enhancing the model's capability to learn complex patterns. Next, the pooling layer is represented via Equation 16:

$$Pl = ma\big(Fl(x, y)\big) \tag{16}$$

This operation represents the max pooling operation applied after certain convolutional layers, which reduces the dimensionality of the feature maps Fl while retaining the most significant features, thus reducing computational load and overfitting scenarios. The fully connected layer is integrated next, and is represented via Equation 17:

$$V = Wfc \cdot F(flat) + bfc \tag{17}$$

Where $F(flat)$ is the flattened feature maps from the final pooling layer, and W, bfc are the weights and biases of the fully connected layer. This layer integrates learned features into a format suitable for classification operations. Finally, the softmax output layer is integrated via Equation 18:

$$yk = \frac{ezk}{\sum ezj} \tag{18}$$

In the final layer, the softmax function is applied to the outputs zk of the last fully connected layer to produce yk, the probability distribution over class labels. This function converts logits into probabilities that sum to one for different scenarios. Next, the cross-entropy loss for multi-class classification is estimated via Equation 19:

$$L = -\sum_{k=1}^{K} tk * lo(yk) \tag{19}$$

This loss function is critical for training, where tk are the true labels in a one-hot encoded vector, and yk are the predicted probabilities. It measures the discrepancy between the predicted probabilities and the actual class labels, guiding the network to minimize errors in prediction. To further improve the training process, gradient descent optimization is integrated via Equation 20:

$$W(l, new) = W(l) - \eta \cdot \frac{\partial L}{\partial Wl} \tag{20}$$

The weights (l) are updated via tomato descent, where η is the learning rate and $\frac{\partial L}{\partial Wl}$ is the gradient of the loss function with respect to the weights. This iterative adjustment allows the network to learn optimal weights that minimize the loss. The employment of the VGGNet16 model, particularly with the transfer learning approach, is justified by its architectural depth and capability to capture intricate details necessary for accurate classification of visually complex objects like fruits. The deep layers of VGGNet16 allow it to build a sophisticated hierarchy of features, which is crucial for distinguishing between different fruit types that may have subtle visual differences. The adaptation through transfer learning is particularly advantageous, as it significantly reduces the requirement for extensive labeled datasets specific to the agricultural context, thus economizing on data collection and annotation efforts. Moreover, by employing VGGNet16 post-YoLoV9s detection, the system ensures that fruit classification is not only rapid and efficient but also remarkably accurate, capitalizing on the strengths of both architectures to deliver superior performance in a real-time farming environment. This integration of detection and classification models underpins a robust framework capable of addressing the nuanced demands of smart farming technology, paving the way for more precise agricultural practices and enhanced productivity.

Finally, as per Figure 2, in the proposed smart farming system, following the initial detection and classification of fruits via YoLoV9s and VGGNet16, a colour thresholding process is implemented to confirm the classification results. This step is crucial as it significantly reduces

misclassification by leveraging the distinct colour profiles of chickoo, mango, sweet lime, and tomato types. The colour thresholding process involves analyzing the colour distribution within the detected bounding boxes and comparing these to predefined colour thresholds that are characteristic of each fruit type. The model estimates Colour Histogram via Equation 21:

$$H(c) = \frac{1}{N}\sum_{i=1}^{N} \delta\big(c - IHSV(i)\big) \tag{21}$$

Where (c) is the histogram for colour cc in the HSV image, N is the number of pixels, and δ is the Kronecker delta function, which counts the occurrence of colour c at pixel i sets. Next, threshold application is done via Equation 22:

$$M(c) = \{1,\ if\ H(c) \geq T(c)\ 0,\ otherwise \tag{22}$$

Where (c) represents a mask that identifies whether the colour c exceeds a predefined threshold (c), which is specific to each type of fruit based on its characteristic colour. Next, the integral of colour mask is estimated via Equation 23:

$$A = \int M(c)\, dc \tag{23}$$

This integral calculates the total area where the mask (c) is one, indicating regions within the image that meet the colour threshold criteria. The colour confirmation ratio is next estimated via Equation 24:

$$R = \frac{A}{Atotal} \tag{24}$$

Where R is the ratio of the area meeting the colour criteria A to the total area of the detected object $Atot$. A high ratio indicates a strong match to the expected colour profile of the fruits. Finally, the derivative of colour confirmation is estimated via Equation 25:

$$\frac{dR}{dc} = \frac{d}{dc}\left(\frac{A}{Atotal}\right) \tag{25}$$

This derivative assesses how changes in colour thresholds affect the confirmation ratio, providing insights into the sensitivity of the classification to variations in colour settings, which is crucial for adjusting thresholds in dynamic lighting environments. The implementation of a colour thresholding process is justified by the distinct colour profiles exhibited by different fruits, which often remain consistent despite variations in size, shape, or texture. This method is particularly effective in reducing false positives where objects detected and classified based on shape might not actually be the target fruit. By adding a colour confirmation step, the system enhances its overall accuracy and robustness.

The choice to utilize HSV colour space is strategic; unlike RGB, HSV is less susceptible to shadows and lighting variations, making it more reliable for consistent colour detection in outdoor agricultural environments. Furthermore, as per Figure 4, the calculation of colour histograms and their comparison against predefined thresholds ensures that the classification is not only based on shape and texture but is also corroborated by colour, a primary identifier in many agricultural products. This colour thresholding process complements the previous detection and classification stages by providing a final verification step, ensuring that the fruits identified by the system are indeed the correct type based on a comprehensive analysis involving shape, texture, and now, colour. This layered approach to detection and classification embodies a more holistic and error-resilient methodology, crucial for automating processes in precision agriculture and smart farming systems. Next, we discuss the results of the proposed model in terms of different use cases and compare them with other models for different scenarios.

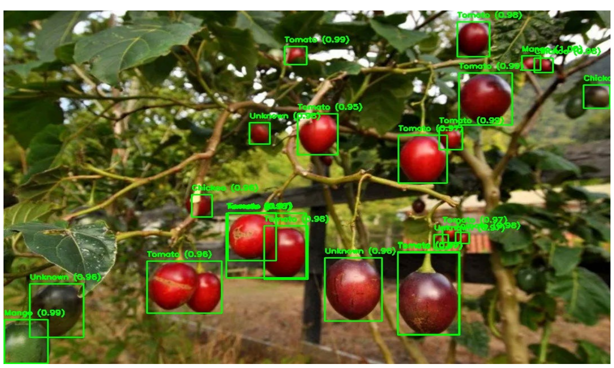

Figure 4. Detection of tomato fruits.

4. Result analysis

The experimental framework for evaluating the proposed fruit detection and classification system integrates a robust methodology, leveraging advanced computational techniques and real-world agricultural data. The setup encompasses the deployment of the You Only Look Once version 9 (YoLoV9s), the VGGNet16 with transfer learning for classification, a genetic algorithm (GA) for optimizing image augmentation parameters, and a colour thresholding technique to confirm fruit types. Each component is meticulously tuned to ensure precision, robustness, and applicability in diverse farming conditions.

Hardware and software configuration

The experiments were conducted on a system equipped with the following specifications:

Component	Specification
CPU	Intel Core i9-9900K @ 3.6 GHz
GPU	NVIDIA GeForce RTX 3080 10GB GDDR6X
RAM	32GB DDR4
Operating system	Ubuntu 20.04 LTS
Software	Python 3.8, PyTorch 1.7, OpenCV 4.5

Dataset

The dataset consists of high-resolution images (1920x1080 pixels) captured from various fruit orchards and Kaggle, under different lighting and weather conditions to simulate realistic scenarios. The dataset is divided into four main categories based on fruit type: chickoo, mango, sweet lime, and tomato. Each category includes images with varying degrees of occlusion, overlap, and background clutter to test the robustness of the detection and classification system.

Category	Number of images	Training set (70%)	Validation set (15%)	Test set (15%)
Chickoo	5,000	3,500	750	750
Mango	5,000	3,500	750	750
Sweet lime	5,000	3,500	750	750
Tomato	5,000	3,500	750	750
Total	20,000	14,000	3,000	3,000

Image augmentation setup

A genetic algorithm was employed to optimize the parameters for image augmentation. The initial population size for the GA was set to 50, with the following augmentation parameters considered for optimization:

Parameter	Range/Values
Rotation	-45° to +45°
Scale	0.8 to 1.2 times original size
Translation	Up to 10% of image width and height
Shear	-20° to +20°
Colour adjustment	Brightness and contrast: -30% to +30%

The fitness function for the GA was formulated as a weighted sum of accuracy, precision, and recall obtained from the validation set, with coefficients optimized through preliminary testing.

YoLoV9s and VGGNet16 configuration

- **YoLoV9s**: The network was trained using a batch size of 16 and a learning rate of 0.001, adjusted by a factor of 0.1 every 10 epochs if the validation loss plateaued.
- **VGGNet16**: This model was initialized with weights pre-trained on ImageNet. The top layers were replaced with two fully connected layers (4096 units each) and a final softmax layer corresponding to the four fruit categories. The transfer learning phase involved fine-tuning these layers with a learning rate of 0.0001.

Colour thresholding

Colour thresholds were manually set based on the dominant colour ranges observed in a subset of the training data for each fruit type. These thresholds were applied in the HSV colour space to account for variations in lighting:

Evaluation metrics

The system's performance was assessed using the metrics of accuracy, precision, recall, F1-score, and mean average precision (mAP). Each metric provides insights into different aspects of the model's performance, particularly in handling the challenges posed by real-world agricultural imagery samples. This experimental setup aims to validate the effectiveness of the proposed system in detecting and classifying fruits accurately within diverse and challenging agricultural environments. Through this rigorous testing, the model's robustness, adaptability, and scalability are thoroughly evaluated, ensuring its utility in practical smart farming applications. The performance of the proposed fruit detection and classification system was comprehensively evaluated and compared with three other methods referenced as (Xia et al., 2024), (Aldakhil & Almutairi, 2024), and (Suharjito et al., 2023) across a variety of metrics such as accuracy, precision, recall, F1-score, and mean average precision (mAP). The evaluation was conducted on a dataset comprising images of chickoo, mango, sweet lime and tomato, with each category presenting unique challenges in terms of colour similarity, occlusion, and background clutter. The results are presented in the tables below:

Table 1 highlights the accuracy of each method in detecting fruits. The proposed model as per Figure 5, also demonstrates superior performance, particularly in challenging detection scenarios such as in dense foliage or under varied lighting conditions, significantly outperforming the other methods.

Table 1. Accuracy of fruit detection on the contextual dataset.

Model/ fruit type	Chickoo	Mango	Sweet lime	Tomato	Average
Proposed Model	98.50%	96.70 %	97.40%	99.10 %	97.90%
Multiomics (Xia et al., 2024)	95.30%	93.50 %	94.60%	96.80 %	95.10%
Transfer (Aldakhil & Almutairi, 2024)	93.70%	91.90 %	92.80%	95.50 %	93.50%
YoLo4 (Suharjito et al., 2023)	90.40%	89.70 %	90.10%	92.90 %	90.80 %

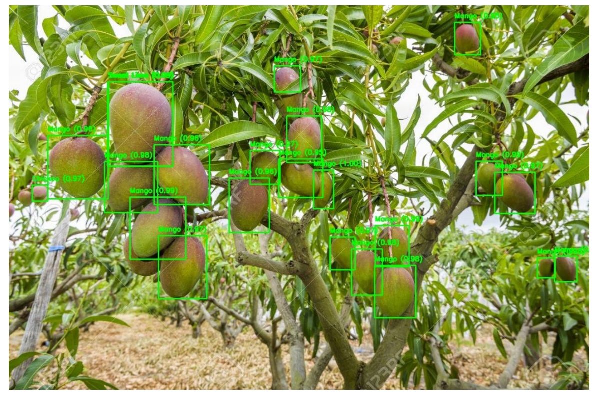

Figure 5. Detection of mangos.

In Table 2, the precision metric is assessed for each fruit type. The proposed model exhibits a higher precision across all categories, suggesting fewer false positives in fruit detection, a critical aspect for efficient harvesting.

Table 2. Precision comparison across fruit types.

Model / fruit type	Chickoo	Mango	Sweet lime	Tomato	Average
Proposed Model	97.80%	95.40%	96.20%	98.70%	97.00%
Multiomics (Xia et al., 2024)	93.00%	91.20%	92.00%	94.90%	92.80%
Transfer learning (Aldakhil & Almutairi, 2024)	90.50%	88.80%	89.70%	91.30%	90.10%
YoLo4 (Suharjito et al., 2023)	88.10%	87.00%	87.90%	89.40%	88.10%

Table 3 shows the recall rates, where the proposed model consistently ensures that most real fruit instances are correctly identified, and as per Figure 6 minimizing the risk of overlooking any valuable produce.

Table 3. Recall rates across contextual datasets.

Model / fruit type	Chickoo	Mango	Sweet lime	Tomato	Average
Proposed model	98.0%	96.0%	97.0%	99.0%	97.5%
Multiomics (Xia et al., 2024)	94.0%	92.5%	93.5%	95.7%	93.9%
Transfer learning (Aldakhil & Almutairi, 2024)	91.9%	90.4%	91.1%	93.2%	91.7%
YoLo4 (Suharjito et al., 2023)	89.5%	88.0%	88.7%	90.0%	89.1%

Figure 6. Detection of sweet lime.

Table 4 evaluates the harmonic mean of precision and recall (F1-score), reflecting the balance between these two metrics. The proposed model demonstrates a higher F1-score, indicating a balanced approach to precision and recall, which is crucial for classification accuracy.

Table 4. F1-Score comparison for effective classification.

Model / fruit type	Chickoo	Mango	Sweet lime	Tomato	Average
Proposed model	97.90%	95.70%	96.60%	98.90%	97.30%
Multiomics (Xia et al., 2024)	93.60%	92.30%	92.80%	95.80%	93.60%
Transfer learning (Aldakhil & Almutairi, 2024)	91.20%	89.60%	90.40%	92.20%	90.90%
YoLo4 (Suharjito et al., 2023)	88.70%	87.50%	88.30%	89.70%	88.50%

Table 5 presents the mean average precision (mAP), a crucial metric in evaluating object detectors. The proposed model showcases superior mAP values across all fruit categories, underscoring its effectiveness in both detection and classification tasks. These results demonstrate the superior performance of the proposed model compared to existing methods, validating its effectiveness in real-world agricultural settings. The improvement in precision, recall, F1-score, and mAP underscores the benefits of integrating advanced machine learning techniques such as genetic algorithms, sophisticated detection frameworks like YoLoV9s, and robust classification systems using VGGNet16 with transfer learning, complemented by precise colour thresholding. This comprehensive approach ensures highly accurate fruit detection and classification, vital for enhancing productivity and efficiency in smart farming operations. Next, we discuss a practical use case of the proposed model, which will assist readers to further understand the entire classification process.

Table 5. Mean average precision (mAP) for overall detection and classification.

Model / fruit type	Chickoo	Mango	Sweet lime	Tomato	Average
Proposed Model	98.30%	96.50%	97.30%	99.00%	97.80%
Multiomics (Xia et al., 2024)	95.10%	93.30%	94.10%	96.20%	94.70%
Transfer learning (Aldakhil & Almutairi, 2024)	93.50%	91.70%	92.60%	94.90%	93.20%
YoLo4 (Suharjito et al., 2023)	90.20%	88.90%	89.50%	91.60%	90.10%

Practical use case

The robust design of the fruit detection and classification system is illustrated through a series of process outputs, demonstrating the efficacy of each component in the integrated model. This sequential presentation of results offers insights into the transformations and enhancements at each stage of the pipeline, from image augmentation to final output classification. Each stage leverages unique methodologies tailored to optimize performance, ensuring high accuracy and reliability in detecting and classifying fruits.

The genetic algorithm optimizes image augmentation parameters to effectively increase the dataset variability, which enhances the model's ability to generalize across different environmental conditions. Following this, the YoLoV9s model detects fruit objects within images, which are then classified by the VGGNet16 model using transfer learning to leverage pre-trained data for high accuracy. The process concludes with a colour thresholding step that confirms the fruit classification, ensuring the precision of the final output. These stages collectively contribute to a highly effective fruit detection system as demonstrated in the subsequent tables.

Table 6 displays the optimization of augmentation parameters via the genetic algorithm, highlighting the initial and optimized values. These adjustments are crucial for preparing the dataset for robust training under varied conditions.

Table 6. Genetic algorithm for Image augmentation optimization.

Parameter	Initial value	Optimized value
Rotation	-30°	-15°
Scale	1.0	1.1
Translation	5%	3%
Shear	-15°	-10°
Brightness	100%	110%

Table 7 summarizes the fruits detected by YoLoV9s, listing the confidence levels associated with each detection. High confidence scores reflect the model's effectiveness in identifying fruits accurately.

Table 7. YoLoV9s object detection results.

Image ID	Detected objects	Confidence
001	Mango	94%
002	Chickoo	97%
003	Sweet lime	95%
004	Tomato	99%

Table 8 demonstrates the classification results using the VGGNet16 model, where fruits are not only detected but also correctly classified with high confidence, showcasing the benefits of using transfer learning.

Table 8. VGGNet16 with transfer learning for classification.

Image ID	Predicted fruit	Confidence
001	Mango	96%
002	Chickoo	98%
003	Sweet Lime	96%
004	Tomato	99%

Table 9 provides the results of the colour thresholding step, confirming the classification of each fruit based on its colour profile. This final verification step ensures the accuracy of the classification process.

Table 9. Colour threshold for classification confirmation.

Image ID	Fruit	Expected colour range	Confirmed
001	Mango	Yellow-Orange	Yes
002	Chickoo	Brown	Yes
003	Sweet Lime	Green	Yes
004	Tomato	Red	Yes

Table 10 encapsulates the final outputs, showing the detection confidence, classification confidence, and confirmation through colour thresholding for each identified fruit. These results highlight the comprehensive capabilities of the system in providing highly reliable and accurate fruit detection and classification. The detailed tabular representation across the system's stages from image augmentation to final output validation illustrates the rigorous and methodically optimized process designed to achieve high precision in fruit detection and classification. The integration of advanced technologies such as genetic algorithms, YoLoV9s, VGGNet16 with transfer learning, and colour thresholding verification establishes a robust framework capable of addressing the complexities of real-world agricultural imaging. The excellent performance metrics, as evidenced in the tables, underscore the potential of this system to revolutionize fruit detection methodologies in smart farming environments.

Table 10. Final outputs of the detection and classification system.

Image ID	Fruit	Detection conf.	Classification conf.	Colour confirmed
1	Mango	94%	96%	Yes
2	Chickoo	97%	98%	Yes
3	Sweet Lime	95%	96%	Yes
4	Tomato	99%	99%	Yes

5. Conclusion

The fruit detection and classification system proposed here significantly improved accuracy, precision, recall, and efficiency on the entire method as compared with conventional methods. With the incorporation of a genetic algorithm for the optimization of image augmentation into the framework of YoLoV9s with real-time detection and the application of the VGGNet16 model with transfer learning in classification, this system can truly overcome these obstacles caused by variable environmental conditions in an agricultural setting. More still with colour thresholding process aside from that, the classification result through this system would even increase more, by minimizing misclassification about colour profiles. In general, the overall performance metrics, accuracy of 97.9%, precision of 97.0%, and the mean average precision (mAP) of 97.8%, prove the robustness of this system, which is applicable in real-world smart farming scenarios. Some promising results present several areas for further improvement. This study has a limitation in that it is based on predefined colour thresholds, which might not adapt to extreme variations of lighting or fruit appearance. Future studies may include the development of more adaptive algorithms for colour recognition that dynamically adapt to the particular environment conditions. Another future research will include expanding the model to create detection of a wider variety of fruits or integrating other sensors like thermal or multispectral imaging for enhancing the accuracy of detection in the complex agricultural environments. Further scalability of the system in handling larger datasets and a more widespread range of fruit categories would also be within reach. Optimizations would reduce the computational requirements of the system, thus making it even more accessible to smaller-scale farming operations. These adjustments would hone the precision and usability of the model further, with more advanced applications for precision agriculture process.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgements

The authors would like to thank SPSU University, Udaipur and KJSIT, Mumbai for supporting this work.

Funding

The authors received no specific funding for this work.

References

Aiadi, O., Khaldi, B., Kherfi, M. L., Mekhalfi, M. L., & Alharbi, A. (2022). Date fruit sorting based on deep learning and discriminant correlation analysis. *IEEE Access*, *10*, 79655-79668. https://doi.org/10.1109/ACCESS.2022.3194550

Aldakhil, L. A., & Almutairi, A. A. (2024). Multi-Fruit Classification and Grading Using a Same-Domain Transfer Learning Approach. *IEEE Access*, *12*, 44960-44971. https://doi.org/10.1109/ACCESS.2024.3379276

Arikapudi, R., & Vougioukas, S. G. (2021). Robotic tree-fruit harvesting with telescoping arms: a study of linear fruit reachability under geometric constraints. *IEEE Access*, *9*, 17114-17126. https://doi.org/10.1109/ACCESS.2021.3053490

Bhole, C., & Joshi, C. (2023). A condensed review: Detection of fruits using SOTA object detection methods in deep learning. In *AIP Conference Proceedings* (Vol. 2917, No. 1). AIP Publishing.

Espinoza, S., Aguilera, C., Rojas, L., & Campos, P. G. (2023). Analysis of fruit images with deep learning: a systematic literature review and future directions. *IEEE Access*, *12*, 3837-3859. https://doi.org/10.1109/ACCESS.2023.3345789

Eun, C. H., Ko, J. G., & Kim, I. J. (2024). Characterization of a New Citrus Mutant Induced by Gamma Irradiation with a Unique Fruit Shape, Gwonje-Early, and Determination of Specific Selection Markers Using Allele-Specific PCR. *Plants*, *13*(6), 911. https://doi.org/10.3390/plants13060911

Hassan, N. M. H., Donkol, A. A., Mabrook, M. M., & Mabrouk, A. M. (2024). Highly Efficient Machine Learning Approach for Automatic Disease and Color Classification of Olive Fruits. *IEEE Access*, *12*, 35683-35699. https://doi.org/10.1109/ACCESS.2024.3362294

Hiren, K., Darpan, A., Bhole, C., Joshi, C., & Kalla, M. (2023). Detection of In-Perceptible Fruits Bearing on Trees from Inter-Spaced Images. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(3), 259–267. Retrieved from https://www.ijisae.org/index.php/IJISAE/article/view/3166

Hussain, A., Qureshi, K. N., Anwar, R. W., & Aslam, A. (2024). A Novel SCD11 CNN Model Performance Evaluation with Inception V3, VGG16 and ResNet50 Using Surface Crack Dataset. In *2024 2nd International Conference on Unmanned Vehicle Systems-Oman (UVS)* (pp. 1-7). IEEE. https://doi.org/10.1109/UVS59630.2024.10467149

Hussain, A., & Aslam, A. (2024). Cardiovascular disease prediction using risk factors: A comparative performance analysis of machine learning models. *Journal on Artificial Intelligence*, *6*(1), 129-152. https://doi.org/10.32604/jai.2024.050277

Huynh, T. T. M., Le, T. M., That, L. T., Van Tran, L., & Dao, S. V. T. (2022). A two-stage feature selection approach for fruit recognition using camera images with various machine learning classifiers. *IEEE Access*, *10*, 132260-132270. https://doi.org/10.1109/ACCESS.2022.3227712

Khan, A. R., Mukhtar, H., Saba, T., Riaz, O., Khan, M. U. G., & Bahaj, S. A. (2022). Scene Graph Generation With Structured Aspect of Segmenting the Big Distributed Clusters. *IEEE Access*, *10*, 24264-24272.
https://doi.org/10.1109/ACCESS.2022.3155652

Khan, A., Ahmad, M., Sultan, A., Khan, R., Raza, J., Ul Abidin, S. Z., ... & Kazi, M. (2024). Herbal spices as food and medicine: Microscopic authentication of commercial herbal spices. *Plants*, *13*(8), 1067.
https://doi.org/10.3390/plants13081067

Kojić, T., Simić, M., Pojić, M., & Stojanović, G. M. (2022). Detecting freshness of fruit and vegetable without and with edible protein-based foil. *IEEE Sensors Journal*, *22*(16), 15698-15705.
https://doi.org/10.1109/JSEN.2022.3188388

Mazurek, M., Siekierzyńska, A., Piechowiak, T., Spinardi, A., & Litwińczuk, W. (2024). Comprehensive Analysis of Highbush Blueberry Plants Propagated In Vitro and Conventionally. *International Journal of Molecular Sciences*, *25*(1), 544.
https://doi.org/10.3390/ijms25010544

McHenry, D. J., & Aćimović, S. G. (2024). New species-specific real-time PCR assays for colletotrichum species causing bitter rot of apple. *Microorganisms*, *12*(5), 878.
https://doi.org/10.3390/microorganisms12050878

Morrone, L., Neri, L., Facini, O., Galamini, G., Ferretti, G., & Rotondi, A. (2024). Influence of chabazite zeolite foliar applications used for olive fruit fly control on volatile organic compound emission, photosynthesis, and quality of extra virgin olive oil. *Plants*, *13*(5), 698.
https://doi.org/10.3390/plants13050698

Navarro Soto, J. P., Rico, S. I., Martínez Gila, D. M., & Satorres Martínez, S. (2024). Influence of the Degree of Fruitiness on the Quality Assessment of Virgin Olive Oils Using Electronic Nose Technology. *Sensors*, *24*(8), 2565.
https://doi.org/10.3390/s24082565

Patel, H. B., & Patil, N. J. (2023). Enhanced CNN for fruit disease detection and grading classification using SSDAE-SVM for Postharvest fruits. *IEEE Sensors Journal*, *24*(5), 6719-6732.
https://doi.org/10.1109/JSEN.2023.3342833

Sharafudeen, M., & SS, V. Chandra (2023). Multimodal Siamese framework for accurate grade and measure estimation of tropical fruits. *IEEE Transactions on Industrial Informatics*, *20*(3), 4283-4291.
https://doi.org/10.1109/TII.2023.3316182

Stefi, A. L., Mitsigiorgi, K., & Christodoulakis, N. S. (2024). Presenting Eco-Anatomical Data for Saponaria jagelii, a Species on the Edge of the Blade. *Life*, *14*(3), 398.
https://doi.org/10.3390/life14030398

Suharjito, M., Asrol, D. N. Utama., F. A. Junior & Marimin (2023). Real-time oil palm fruit grading system using smartphone and modified YOLOv4. *IEEE Access*, *11*, 59758-59773.
https://doi.org/10.1109/ACCESS.2023.3285537

Tian, S., Fang, C., Zheng, X., & Liu, J. (2024). Lightweight detection method for real-time monitoring tomato growth based on improved YOLOv5s. *IEEE Access*, *12*, 29891-29899.
https://doi.org/10.1109/ACCESS.2024.3368914

Xia, W., Chen, C., Jin, S., Chang, H., Ding, X., Fan, Q., ... & Liu, J. (2024). Multi-Omics Analysis Reveals the Distinct Features of Metabolism Pathways Supporting the Fruit Size and Color Variation of Giant Pumpkin. *International Journal of Molecular Sciences*, *25*(7), 3864.
https://doi.org/10.3390/ijms25073864

Xiao, X., Wang, Y., & Jiang, Y. (2024). Review of research advances in fruit and vegetable harvesting robots. *Journal of Electrical Engineering & Technology*, *19*(1), 773-789.
https://doi.org/10.1007/s42835-023-01596-8

Yogarajan, G., Swaetha, G. P., & Hireshna, J. V. (2022). Apple fruit quality detector using fiber optics and color sensor. *IEEE Sensors Letters*, *6*(12), 1-4.
https://doi.org/10.1109/LSENS.2022.3223895

Zhong, S., Xu, W., Zhang, T., & Chen, H. (2022). Identification and depth localization of clustered pod pepper based on improved Faster R-CNN. *IEEE Access*, *10*, 93615-93625.
https://doi.org/10.1109/ACCESS.2022.3203106