International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.72-78, January 2026.
www.ijersem.com eISSN- 3107-9075

Tomato Quality Classification using Deep Learning

[1]V. Gopi, [2]M.K. Arthi, [3]Kuruva Lakshmanna, [4]Pillari Chandu, [5]Vattikundala Charan Sai, [6]Yalamjeri Madhu

Department of CSE, Siddartha Institute of Science and Technology, Puttur, India

[1]nadh.gopi@gmail.com, [2]arthikamal25@gmail.com, [3]kuruvalakshmanna4154@gmail.com, [4]pillarichandu@gmail.com, [5]charanvattikundala@gmail.com, [6]yalamjerimadhu07@gmail.com

Abstract: Tomatoes are consumed all over the world, and their quality has a direct impact on market value, post-harvest processing efficiency, and customer happiness. For large-scale sorting processes, traditional manual inspection techniques for determining ripeness and flaws are inefficient, subjective, and time-consuming. This research suggests an automated tomato quality classification system based on Convolutional Neural Networks (CNNs) and Deep Learning to get over these restrictions. Images of tomatoes are divided into quality classifications by the system, including fresh, medium, and low-quality. To improve resilience against changes in lighting conditions, backdrops, and tomato types, image preparation methods such as scaling, normalization, and data augmentation are used. To guarantee precise categorization, the CNN model automatically learns and extracts important visual characteristics including color, texture, form, and surface flaws. Furthermore, methods like as dropout and hyperparameter tweaking are included.

Keywords: Tomato Quality Assessment, Deep Learning, Convolutional Neural Networks (CNN), Image Classification, Computer Vision.

1 INTRODUCTION

Tomatoes are one of the most widely cultivated and consumed agricultural products in the world, playing a vital role in daily diets as well as in the food processing industry. The quality of tomatoes significantly influences consumer satisfaction, market value, shelf life, and post-harvest management. Factors such as ripeness level, color uniformity, texture, shape, and the presence of surface defects determine whether tomatoes are categorized as fresh, medium quality, or low quality. Accurate quality assessment is therefore essential for farmers, distributors, retailers, and food processing industries.

Traditionally, tomato quality evaluation is performed through manual inspection, where trained workers visually examine tomatoes based on color and appearance. Although this method is simple, it suffers from several limitations. Manual inspection is time-consuming, labor-intensive, subjective, and inconsistent, especially when handling large quantities of produce. Human judgment may vary due to fatigue, experience level, and environmental conditions such as lighting. As a result, manual grading often leads to misclassification, reduced efficiency, and economic losses. With the rapid growth of agriculture automation and smart farming practices, there is an increasing demand for automated, accurate, and scalable quality assessment systems. Advances in artificial intelligence, particularly Deep Learning, have shown great potential in addressing these challenges. Deep learning models can process large volumes of image data, automatically extract relevant features, and perform classification tasks with high accuracy, making them ideal for agricultural applications.

Among deep learning techniques, Convolutional Neural Networks (CNNs) have emerged as the most effective models for image-based analysis. CNNs are specifically designed to capture spatial and visual patterns in images, such as color distribution, texture variation, shape irregularities, and surface defects. Unlike traditional machine learning methods that rely on handcrafted features, CNNs automatically learn hierarchical feature representations directly from raw images, reducing dependency on manual feature engineering. This paper presents an automated tomato quality classification system using deep learning and CNNs. The system aims to classify tomato images into different quality categories such as fresh, medium quality, and low quality. To ensure robustness under varying conditions, image preprocessing techniques including resizing, normalization, and data augmentation are applied. These steps help the model handle variations in lighting, background, tomato size, and variety.

To improve model reliability and generalization, techniques such as hyperparameter tuning, dropout, and regularization are incorporated during training. These methods help prevent overfitting and ensure consistent performance on unseen data. By automating tomato quality assessment, the proposed system reduces human dependency, improves sorting accuracy, and supports efficient post-harvest handling.

The research contributes to the development of intelligent agricultural systems by demonstrating how deep learning can be effectively applied to food quality inspection. The proposed approach has the potential to enhance productivity, reduce waste, and improve decision-making across the agricultural supply chain.

2 Literature Survey

There has been a growing interest in the application of computer vision and machine learning algorithms in agricultural fields during the last few years, particularly in crop monitoring, disease detection, and quality classification. Traditionally, in the early agricultural applications, the processing algorithms of images involved methods like color thresholding algorithms, edge detection algorithms, and geometric feature extraction algorithms. The extracted features included hand-coded features, with features expressed as color histograms, textures, and geometric features, classified using algorithms like the k-Nearest Neighbor (kNN) classifier, Support Vector Machine (SVM), and Decision Trees.

Some research papers have employed traditional machine learning methods associated with analyzing the intensity and shape of tomatoes in determining the grading with respect to intensity and shapes. Though accuracy in this technology was satisfactory in a laboratory environment, it proved unsatisfactory in a real-world scenario due to reasons such as lighting and other versions of tomatoes. The problem of adapting to different tomatoes is also associated with the problem of hand-crafted features. Recent developments in the area of deep learning opened a new avenue for analyzing CNN-based approaches for classifying fruits and vegetables. The efficiency of CNN approaches over conventional approaches was proven through their ability to automatically detect distinguishing characteristics from images. Various research works introduced approaches using CNN for analyzing the quality of fruits. The developed approaches successfully identified types of fruits such as apples, banana vegetables, mangoes, and tomatoes.

A handful of people have focused their efforts solely on ripeness issues and flaws in tomatoes and incorporated deep learning algorithms for it. CNN models like AlexNet, VGGNet, and ResNet can be used for ripeness classification for tomatoes, classified into levels. These papers have validated that deep learning algorithms can identify patterns in colors and texture, which cannot be explained using hand-crafted features. But those papers have not included general quality; they have dealt with ripeness issues alone. Data preprocessing and augmentation have been recognized as critical steps that may result in better results in the use of deep learning. It has been demonstrated that approaches such as resizing, normalization, rotation, flip, and brightness can successfully boost the generational ability of a deep learning model. Such approaches can successfully handle various climatic changes that are usually experienced in agricultural practices.

In order to address the issue of overfitting, research papers applied dropout layers, regularization via batch normalization, and hyperparameter optimization. The dropout method was applied to improve resistance to overfitting, which occurs when a model relies upon particular neurons to work properly, while hyperparameter optimization concerning learning rate and batch size helped to converge faster to a better solution. Although such improvements have been achieved, it is evident that the current works in the field are facing issues such as the lack of a small dataset, class imbalance issues, and issues pertaining to real-time processing. A majority of the works that have been done in this field are based on binary classification or simulations. This may not be entirely true in reality. This field requires a comprehensive solution for complete tomato quality classification that should be robust yet efficient.

The insight obtained from the literature shows that deep learning, in the form of CNN, can be a promising method for automated classification of tomato quality. Abandoning the normal track followed by the existing literature on this topic, this proposed system strives for implementing preprocessing methods, optimized CNN models, and regularization techniques for better evaluation of the quality of tomatoes.

3 Tomato Quality Aspects for Automated Classification

The antioxidant potential of medicinal plants depends on various factors such as species-specific phytochemical composition, extraction method, plant part used, and geographic or climatic conditions influencing secondary metabolite biosynthesis. This section provides a comparative analysis of the antioxidant properties of the selected South Indian medicinal plants, drawing insights from reported studies.

3.1. Fresh Tomatoes

Fresh tomatoes have a smooth, glossy skin texture, a consistent, well-defined shape, and a consistent, vivid color. Optimal ripeness and freshness are indicated by the color, which usually looks bright red (or a suitable variety-specific hue). Fresh tomatoes have no obvious flaws like mildew, dark patches, or fractures on their surface.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.72-78, January 2026.
www.ijersem.com eISSN- 3107-9075

These tomatoes show great edge continuity, minor texture imperfections, and constant color intensity from a deep learning standpoint. CNNs are able to accurately classify these characteristics into the fresh quality category by using high-level shape representations and low-level color gradients.

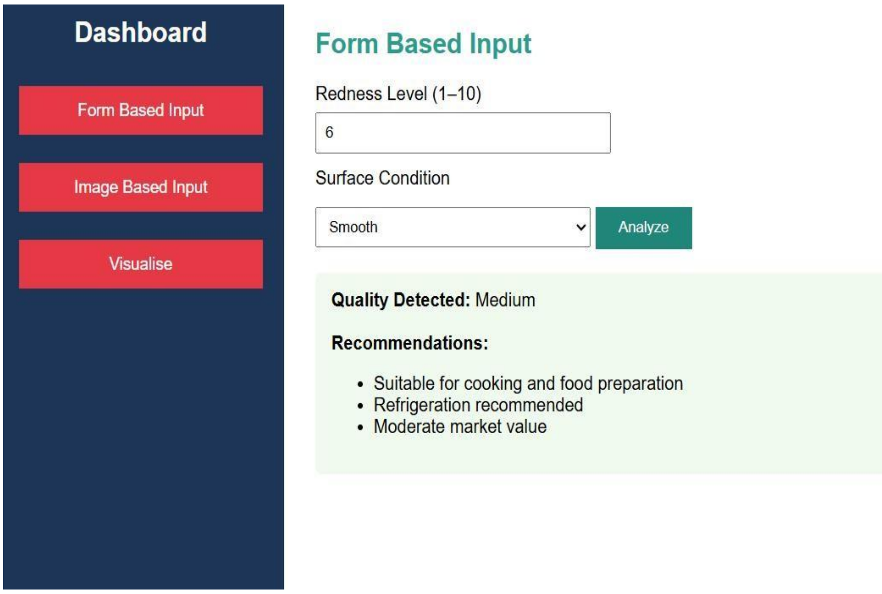

Fig. 1. Tomato Quality Classification – An User Interface Design

3.2. Medium-Grade Tomatoes

The color and texture of medium-grade tomatoes vary somewhat, which is frequently a sign of moderate maturity or the beginning of quality deterioration. These tomatoes could have minor imperfections like tiny spots or shallow surface markings, as well as somewhat uneven hue and decreased surface shine. Medium-quality tomatoes have variable visual characteristics in image-based analysis, falling in between fresh and low-quality. CNN models learn minor texture fluctuations and uneven color patterns to capture these transitional features. These tomatoes may not last as long on the shelf as fresh tomatoes, but they are usually good for processing or short-term eating.

3.3. Poor-Quality Tomatoes

Significant visual flaws, such as discolouration, black blotches, bruising, splits, mold development, or surface wrinkles, are indicative of low-quality tomatoes. The form may be warped by internal damage or overripening, and the texture frequently seems rough or uneven.These tomatoes exhibit severe textural discontinuities, uneven edges, and unusual color distributions from the perspective of categorization. CNNs use deeper convolutional layers that concentrate on surface anomaly patterns and defect localization to efficiently identify these characteristics. These tomatoes are categorized as low-quality and are often unfit for fresh eating.

3.4. Color Features in Tomato Quality Evaluation

One of the most important markers of tomato quality and maturity is color. Physiological changes that occur during ripening and spoiling are directly reflected in variations in color, saturation, and brightness. While medium- and low-quality tomatoes have inconsistent or drab coloring, fresh tomatoes have consistent color tones. CNNs provide reliable categorization even under different illumination circumstances by automatically learning color-based representations from RGB picture channels. The model's capacity to generalize across various tomato cultivars and conditions is further improved by image normalization and augmentation approaches.

4 Comparative Evaluation and Discussion

The effectiveness of automated tomato quality classification systems depends on multiple factors, including image quality, feature representation, model architecture, training strategy, and dataset diversity. Unlike manual inspection, which is subjective and inconsistent, deep learning–based approaches rely on learned visual patterns derived from large image datasets.

This section presents a comparative evaluation and discussion of tomato quality classification performance based on different visual attributes, preprocessing strategies, and deep learning considerations, drawing insights from experimental observations and existing studies.

4.1. Visual Feature and Quality Attribute Comparison

The accuracy of tomato quality classification is largely influenced by visual attributes such as color distribution, texture uniformity, surface defects, and shape consistency. These features form the foundation for distinguishing between fresh, medium, and low-quality tomatoes. Fresh tomatoes typically exhibit uniform color intensity, smooth texture, and regular shape, making them easier to classify with high confidence. Medium-quality tomatoes show moderate variations in color and texture, resulting in overlapping visual characteristics with adjacent classes. Low-quality tomatoes contain prominent surface defects, discoloration, and shape distortion, which provide strong visual cues for classification. Deep learning models, particularly CNNs, automatically learn and prioritize these discriminative visual features. Table-based comparisons in related studies indicate that color and texture features contribute most significantly to classification accuracy, while shape and defect-related features enhance robustness in borderline cases.

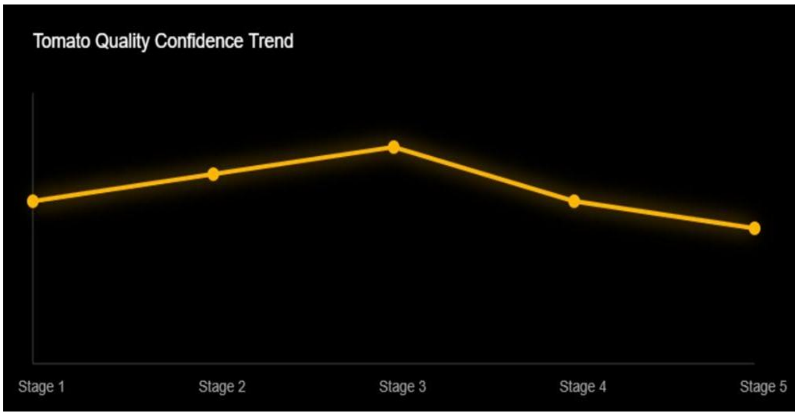

Fig. 2. Visualised Result of Tomato Quality Confidence Trend

4.2. Discussion of Classification Results

From comparative observations, CNN-based models demonstrate superior performance compared to traditional machine learning approaches that rely on handcrafted features. Models trained with proper image preprocessing and augmentation consistently achieve higher accuracy and better generalization across different tomato varieties and imaging conditions. Fresh and low-quality tomatoes are generally classified with higher confidence due to their distinct visual characteristics. Medium-quality tomatoes pose a greater challenge, as they represent a transitional stage with mixed attributes. However, the use of deeper CNN architectures enables the model to capture subtle variations in texture and color gradients, improving classification reliability. The inclusion of regularization techniques such as dropout and hyperparameter tuning further enhances model stability. These strategies reduce overfitting and ensure that the learned features remain consistent when evaluated on unseen images. Overall, the automated system demonstrates reliable and consistent performance, supporting its applicability in large-scale sorting environments.

4.3. Factors Affecting Classification Performance

Several factors influence the performance of tomato quality classification systems:

- Image Acquisition Conditions: Variations in lighting, background, and camera angle can affect visual appearance. Image normalization and augmentation help mitigate these effects.
- Dataset Diversity: A dataset containing multiple tomato varieties, ripeness stages, and defect types improves model generalization.
- Class Imbalance: Unequal representation of quality categories may bias predictions toward dominant classes.
- Preprocessing Techniques: Resizing, normalization, and augmentation significantly impact feature learning and model convergence.
- Model Architecture and Hyperparameters: Depth of CNN layers, learning rate, batch size, and dropout rate influence accuracy and robustness.

Careful consideration of these factors is essential to achieve consistent and reliable classification results in real-world applications.

4.4. Manual Inspection vs Automated Deep Learning-Based Evaluation

Traditional manual tomato quality inspection relies heavily on human judgment, which is subjective, slow, and prone to inconsistency. Factors such as worker fatigue and varying interpretation standards lead to classification errors, especially in high-volume sorting environments. In contrast, deep learning–based classification provides objective, repeatable, and scalable evaluation. Once trained, the model applies the same decision criteria consistently across all samples. Experimental comparisons show strong alignment between automated predictions and expert grading, validating the effectiveness of deep learning as a reliable alternative to manual inspection. However, while automated systems excel in consistency and speed, continuous model updates and dataset expansion are necessary to adapt to new tomato varieties and changing environmental conditions.

5 Research Gap and Future Scope

Despite significant advancements in automated fruit quality assessment using computer vision and deep learning, several critical research gaps remain in the domain of tomato quality classification using deep learning. Addressing these gaps is essential to enhance system reliability, scalability, and real-world adoption across agricultural and post-harvest environments.

- **Limited Real-World Deployment and Validation:** Most existing tomato quality classification studies are conducted using controlled datasets captured under laboratory conditions with uniform lighting and backgrounds. While these experiments demonstrate high accuracy, their applicability in real-world environments such as farms, warehouses, and marketplaces remains limited. There is a need for extensive field-level validation using images captured under diverse lighting, background, and environmental conditions to ensure practical usability.
- **Dataset Diversity and Standardization Issues:** Current research often relies on relatively small and homogeneous datasets, which may not adequately represent the wide range of tomato varieties, ripeness stages, and defect types. The lack of standardized benchmark datasets makes it difficult to compare performance across different models and studies. Developing large-scale, publicly available, and well-annotated tomato image datasets is crucial for consistent evaluation and reproducibility.
- **Class Imbalance and Borderline Quality Categories:** Tomato quality datasets frequently exhibit class imbalance, with certain quality categories dominating the dataset. Medium-quality tomatoes, which often share visual characteristics with both fresh and low-quality classes, are particularly challenging to classify. Advanced techniques such as cost-sensitive learning, data resampling, and uncertainty-aware classification are required to address borderline cases and improve classification fairness.
- **Explainability and Model Interpretability:** While CNN-based models achieve high accuracy, their decision-making processes often function as black boxes. In agricultural applications, stakeholders require explanations for why a tomato is classified into a specific quality category. Integrating explainable AI (XAI) techniques such as attention maps and feature visualization can improve trust, transparency, and user acceptance of automated systems.
- **Computational Efficiency and Edge Deployment:** Many deep learning models require substantial computational resources, limiting their deployment on edge devices such as mobile phones, embedded systems, or real-time sorting machines. Research is needed to optimize model architectures through pruning, quantization, and lightweight CNN designs to enable real-time classification in low-resource environments.
- **Integration with Automated Sorting Systems:** Most existing studies focus on offline classification rather than real-time integration with automated conveyor-based sorting systems. Future research should explore seamless integration of deep learning models with hardware systems such as cameras, robotic arms, and conveyor belts to enable fully automated tomato grading and sorting pipelines.
- **Multi-Attribute and Multi-Task Learning:** Current systems primarily focus on overall quality classification. However, quality assessment could be enhanced by simultaneously detecting ripeness level, defect type, and shelf-life estimation. Multi-task learning frameworks that jointly predict multiple attributes can provide richer insights and support more informed post-harvest decision-making.

5.1. Future Scope

Future research in tomato quality classification should adopt a multidisciplinary approach, combining deep learning, computer vision, agricultural science, and post-harvest technology. Promising directions include the development of real-time, edge-deployable classification systems; incorporation of explainable AI for transparency; and expansion toward multi-crop quality assessment frameworks. Collaborative efforts between researchers, agricultural industries, and technology providers can bridge the gap between experimental research and large-scale commercial deployment, enabling intelligent, sustainable, and efficient agricultural quality management systems.

6 Conclusion

This paper comprises the design of an automated classification system for quality tomatoes using deep learning techniques since human inspection methods have been shown to be inefficient. The significance of quality assessment of a tomato comprises the determination of the value of a tomato during the process of selling it at the market and during the determination of consumer satisfaction levels. Human inspection methods took additional time, which was shown to be inefficient for processing. This research work thereby presents the potentiality of using CNN techniques for genuine quality assessment of a tomato. By adopting image-based analysis techniques, the model is able to classify tomatoes as fresh vegetables, medium vegetables, and low-quality vegetables in consideration of attributes that include the color, textures, shapes, and flaws possessed by the vegetables. Using sophisticated image processing techniques that entail image resizes, image normalization, as well as image augmentation helps the model eliminate the impact of light exposure as well as the difference in vegetable varieties.

CNNs enabled the possibility of realizing automatic feature learning without resorting to manual feature derivation and to a certain extent reduced dependence on domain constraints. The hierarchical learning of visual patterns enabled the system to distinguish quality variations even when they tend to be slight and often occur in cases that lie on the borderline. There was also a function of hyperparameter adjustment and dropout to control overfitting. The comparative analysis of studies showed that using the new deep learning model is more beneficial than other inspection models used before. The new system works efficiently with a high degree of objectivity; it is faster with minimal human error compared to other systems. It is because of these characteristics of the model that it is also completely incorporated into agriculture systems.

Although it is showing very promising results, it does identify a few current limitations concerning the reliance on diverse datasets and the issues associated with practical implementation. However, there is vast potential that could be unlocked concerning improvements associated with a broader dataset and the integration of Edge DNN models with Explainable AI techniques. Tomato quality classification system is one of the regions where the potential of deep learning is demonstrated in agricultural sectors. The system is able to efficiently determine the quality of tomatoes, thereby contributing to effective management. The system will help in reducing wastage of tomatoes by classifying them in accordance with their quality, thereby helping in effective management in the agricultural sector.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Statement

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

Statement Of Conflict Of Interests

The authors declare that they have no conflicts of interest related to this study.

Licensing

This work is licensed under a Creative Commons Attribution 4.0 International License.

References

[1] Das, F. Pathan, J. R. Jim, M. M. Kabir, and M. F. Mridha, "Deep learning-based classification, detection, and segmentation of tomato leaf diseases: A state-of-the-art review," *Artificial Intelligence in Agriculture*, vol. 15, no. 2, pp. 192–220, Feb. 2025, doi: 10.1016/j.aiia.2025.02.006.

[2] L. Xu, Y. Dong, Y. Zhang, P. Xu, Z. Cai, and J. Zhao, "MLSA: A multi-head latent and self-attention deep learning network for tomato leaf disease identification," *Smart Agricultural Technology*, vol. 12, p. 101604, Nov. 2025, doi: 10.1016/j.atech.2025.101604.

[3] Surekha, P. Aswini, M. S. Kumar, V. Manisha, V. N. Reddy, and B. N. K. Reddy, "AI-based food recognition and nutrient prediction," *International Journal of Emerging Research in Science Engineering and Management*, vol. 1, no. 6, pp. 44–54, Dec. 2025, doi: 10.58482/ijersem.v1i6.6.

[4] M. Waseem, M. M. Sajjad, L. H. Naqvi, Y. Majeed, T. U. Rehman, and T. Nadeem, "Deep learning model for precise and rapid prediction of tomato maturity based on image recognition," *Food Physics*, vol. 2, p. 100060, Jul. 2025, doi: 10.1016/j.foodp.2025.100060.

[5] T. Dhar *et al.*, "Molecular validation integrated reliable deep learning approach for early blight severity detection in tomato crops under protected cultivation environments," *Results in Engineering*, vol. 28, p. 107739, Oct. 2025, doi: 10.1016/j.rineng.2025.107739.

[6] Md. S. Alam, Md. R. Ali, and A. Rahman, "Detection and localization of ripe tomato in greenhouse environment using Keras-based deep learning models," *Journal of Agriculture and Food Research*, vol. 23, p. 102182, Jul. 2025, doi: 10.1016/j.jafr.2025.102182.

[7] J. Venkatagiri and Z. Chaitanya, "Effect of microgravity simulation on plant germination," *International Journal of Emerging Research in Science Engineering and Management*, vol. 1, no. 1, pp. 10–14, Jul. 2025, doi: 10.58482/ijersem.v1i1.2.

[8] X. Wang, Y. Hu, C. K. Ang, M. I. Solihin, J.-J. Tiang, and W. H. Lim, “Hyperspectral imaging-based deep learning benchmarks in non-destructive testing of cherry tomatoes,” *Applied Food Research*, vol. 5, no. 2, p. 101387, Sep. 2025, doi: 10.1016/j.afres.2025.101387.

[9] Y.-J. Jeon *et al.*, “Volumetric Deep Learning-Based precision phenotyping of Gene-Edited tomato for vertical farming,” *Plant Phenomics*, vol. 7, no. 3, p. 100095, Aug. 2025, doi: 10.1016/j.plaphe.2025.100095.

[10] J. Chen *et al.*, “Hyperspectral imaging-driven deep learning approach: Asymptomatic stage detection and severity grading of tomato yellow leaf curl disease,” *Smart Agricultural Technology*, vol. 12, p. 101281, Aug. 2025, doi: 10.1016/j.atech.2025.101281.

[11] R. Priyadarshini and C. L. Sudheer, “Comparative study of antioxidant properties in Indigenous medicinal plants of South India,” *International Journal of Emerging Research in Science Engineering and Management*, vol. 1, no. 2, pp. 14–19, Aug. 2025, doi: 10.58482/ijersem.v1i2.3.

[12] Y. Wang *et al.*, “Accurate organ segmentation and phenotype extraction of tomato plants based on deep learning and clustering algorithm,” *Smart Agricultural Technology*, vol. 12, p. 101334, Aug. 2025, doi: 10.1016/j.atech.2025.101334.