RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Issue 3: 436-442 (March 2026)

Vit-CNN Fusion for Robust Mango Quality Evaluation based on Classification Across Multiple Public Datasets

Anuja A. Gharpure[1], Neha Jain[1], Vaibhav E. Narawade[2]

10.18805/IJARe.A-6493

ABSTRACT

Background: Mango quality is crucial for economic viability, waste reduction and public health, yet traditional assessment methods often suffer from subjectivity, inefficiency and can be destructive. Existing deep learning approaches, specifically convolutional neural networks (CNNs), achieve higher accuracies but are unable to generalise across diverse datasets because they cannot capture contextual features.

Methods: To address these limitations, this research explores the application of machine learning algorithms, particularly deep learning techniques such as convolutional neural networks (CNNs) and Vision Transformers (ViTs), for an objective and efficient mango classification task. The research work concentrated on two publicly available datasets *viz*. Mango Ripeness Dataset, Mango Classification Dataset based on Mango. A comparative study of these datasets is performed to analyse their performance when fed to variants of CNNs, such as ResNet, MobileNet and ShuffleNet. By leveraging image analysis and feature extraction, this study aims to study the behaviour of different variants of the convolutional neural network on different datasets related to mangoes.

Result: Comparative experiments with CNN-ViT fusion achieve superior accuracy (98.19% on the Mango Ripeness Dataset and 100% on the Mango Classification Dataset), consistently. Additional ablation studies and cross-dataset validation confirm the robustness and scalability of the approach. This work establishes a reproducible framework for automated mango quality evaluation, paving the way for practical deployment in agricultural quality control and supply chain management.

Key words: Agriculture AI, Convolutional neural networks, Deep learning, Machine learning, Mango quality assessment, Vision transformers.

INTRODUCTION

Fruit quality is a critical factor throughout the supply chain, impacting waste reduction, economic viability and public health. Recognizing the limitations of traditional, often subjective, human-based quality assessments, this research proposeps the application of machine learning algorithms for a more objective and efficient solution. The use of machine learning and image processing will be used to find the various solutions in the food industry such as sorting, grading, *etc*. (Begum and Hazarika, 2022). We will leverage a variety of datasets, capturing the nuances of variations of mangoes and diseases found in mangoes, to rigorously evaluate the capacity of various machine learning models in accurately predicting key quality parameters. This investigation seeks to establish a robust and scalable framework for enhancing quality control practices within the agricultural and food sectors.

The advancements in the research focus on the classification, fruit identification, variety identification, maturity grading and defect detection. A key trend involves the increasing use of deep learning techniques, particularly Convolutional Neural Networks (CNNs), which have demonstrated remarkable accuracy in learning complex visual features directly from mango images (Li and Chen, 2022 and Patel and Gupta, 2023). Transfer learning, where models pre-trained on large image datasets are fine-tuned for mango-specific tasks, has also gained prominence due to its ability to achieve high performance with smaller datasets (Reddy and Rao, 2020; Joseph and Abraham, 2021).

[1]Pacific Academy of Higher Education and Research University, Udaipur-313 024, Rajasthan, India.
[2]Ramrao Adik Institute of Technology, D.Y. Patil Deemed to be University, Nerul, Navi Mumbai-400 706, Maharashtra, India.

Corresponding Author: Anuja A. Gharpure, Pacific Academy of Higher Education and Research University, Udaipur-313 024, Rajasthan, India. Email: aagharpure@gmail.com
ORCID: 0009-0006-6993-118X

How to cite this article: Gharpure, A.A., Jain, N. and Narawade, V.E. (2026). Vit-Cnn Fusion for Robust Mango Quality Evaluation based on Classification Across Multiple Public Datasets. *Indian Journal of Agricultural Research*. **60(3):** 436-442. doi: 10.18805/IJARe.A-6493.

Submitted: 19-12-2025 **Accepted:** 28-01-2026 **Online:** 04-02-2026

Studies have explored a diverse range of supervised learning algorithms beyond deep learning. For instance, Random Forests (RF), Support Vector Machines (SVM), K-Nearest Neighbors (KNN) and Decision Trees have been employed for classification based on features extracted through traditional image processing techniques or even non-image data like spectral information from Near-Infrared Spectroscopy (NIR) (Singh and Kumar, 2021; Das and Ghosh, 2022; Suresh and Mohan, 2024). These methods often provide interpretable models and can be effective depending on the specific classification task and the nature of the available data. Ensemble methods, which

combine the predictions of multiple learning algorithms, have also been investigated to improve the robustness and accuracy of mango fruit classification systems (Devi and Jain, 2024).

Further more, the focus of research has expanded to address real-world challenges such as variations in lighting, background complexity and the presence of noise in images. Preprocessing techniques like image enhancement, noise reduction and data augmentation are commonly applied to improve the quality of input data and the generalization ability of the classification models (Bhat and Mir, 2023). The development of mobile-based applications for on-site mango classification indicates a growing interest in deploying these technologies for practical use in agriculture and supply chain management (Joseph and Abraham, 2021). These advancements collectively highlight the significant potential of supervised learning to automate and enhance the efficiency and accuracy of mango fruit classification in various agricultural and commercial contexts.

The convolutional neural network (CNN) shows learning complex features, capturing local patterns and generalises in varying lighting conditions, backgrounds and mango conditions. The Vision Transformers (ViTs) work on large datasets and excels at modelling global dependencies.

The research combines the potentials of both the algorithms CNN and ViT to check whether fusion of CNN and ViT improves mango quality assessment accuracy or not and also to check the generalization across multiple public datasets representing mango ripeness levels and mango variants? The focus of the research work is to observe the classification problem on the various datasets related to mango fruits using supervised learning. The datasets are related to mangoes, variations observed in mangoes. The various machine learning and deep learning algorithms are tested on the datasets. The primary algorithms applied to these datasets include convolutional neural networks (CNNs) and their derivatives. Further more, the Vision Transformer architecture is utilised for classification tasks. Feature extraction is accomplished through the implementation of ResNet, MobileNet and ShuffleNet algorithms. The performance metrics obtained from these algorithms are then compared to determine the most accurate predictor for each dataset. This study aims to focus on a robust, scalable, reproducible framework for mango quality assessment that can be implemented for self-repositories created to address the agricultural problems.

Related work

The task of classifying various fruits and variants of fruits, fruit diseases, is automated using computer vision and machine learning algorithms based on visual features. A significant body of research focuses on developing and comparing different image processing techniques and feature extraction methods for fruit recognition, often tested across publicly available and custom-built fruit image datasets of varying complexity. Comparative analyses across different fruit image datasets, considering factors such as image resolution, lighting conditions, fruit variety and background clutter, provide valuable insights into the robustness and generalizability of different classification approaches.

The transfer learning using Convolutional neural network applied on FIDS 30 dataset, which consists of mixed fruit images, gives 94.8% accuracy for classification and fruit detection (Geerthik *et al.*, 2024). On the same dataset, AlexNet algorithm gives 75% accuracy (Geerthik *et al.,* 2024). On the same dataset, FIDS 30, when a Recurrent Neural Network was applied, the obtained result was 98.47% (Dhiman and Hu, 2021).

A public dataset available on the Kaggle website was used by many researchers to test the model. The Fruit 360 is a dataset available on the Kaggle website. This dataset consists of more than 40,000 images of various fruits organised in training and testing folders (Oltean, 2025). A Convolutional Neural Network was applied on the photos of the fruits apple, lemon and mango from the Fruit-360 dataset gave 95% accuracy (Bobde *et al.*, 2021).

Fruits fresh and rotten fruit images are a publicly available dataset. The machine learning algorithm VGG 16 was used to extract features of specific fruits like apples, oranges and bananas. Apart from various classifiers tested on the dataset, the Support Vector Machine gave 99% accuracy (Mehta *et al.,* 2021). Another work on the fresh and rotten fruit images using VGG 16 and YOLOv5 was carried out on fruits like apples, oranges, bananas, grapes, tomatoes, onions, chilli and capsicum. The VGG 16 gave approximately 91% accuracy (Akshi *et al.*, 2024). The dataset of fresh and stale images of fruits and vegetables available publicly was tested with CNN, BiLSTM, CNN LSTM and CNN BiLSTM machine learning algorithms on selective fruits: apple, banana, tomato, bitter guard, capsicum and orange. The result obtained using CNN BiLSTM was better than the other three algorithms. The result of CNN BiLSTM is 97.76% (Yuan and Alhudhaif, 2024).

Banana ripeness was tested on the banana-ripening-process dataset using a Deep CNN approach. The dataset was publicly available, containing more than 18,000 images. The variants YOLOv8n to YOLOv8x were tested on this dataset. The result obtained was 94.60% to 96.30% accuracy, respectively (Aishwarya and Vinesh, 2023). Like on a banana, deep learning methods are also applied on grape leaf to detect diseases. Among the algorithms like DenseNet121, VGG19, VGG 16, InceptionV3 and ResNet50V2, DenseNet121 achieved highest accuracy 99.86% (Patil and More, 2025).

O.O. Abayomi-Alli, R. Damaševièius, S. Misra and A. Abayomi-Alli created a new dataset known as FruitQ. The dataset consists of images of 11 fruits, consisting of three

freshness classes such as fresh, mildly rotten and fully rotten. A dataset contains more than nine thousand images. The various deep learning algorithms, such as ShuffleNet, SqueezeNet, EfficientNet, ResNet18 and MobileNet-V2, were tested on this dataset. On this dataset, the ResNet 18 gave better performance up to 99.80% accuracy (Abayomi-Alli *et al.*, 2024).

It was being observed that the ripeness level in mango fruit was tested mostly using self-repositories. Images of variants of Mangoes, such as Harumanis and Sala, observed in Malaysia, were collected to check the ripeness level. The classifier, Support Vector Machine and odour sensor were used to classify the images into ripe and unripe categories (Huang *et al.*, 2023). To improve the mango grading accuracy on ripeness level, various classifiers such as Random Forest, Gradient Boosting, Support Vector Machine (SVM), K Nearest Neighbourhood (KNN) and Gaussian Naïve Bayes were applied on the Mango dataset. This dataset was prepared with Himsagor mangoes found in Bangladesh. For the feature extraction, the CNN and VGG16 were applied. The test accuracy obtained is given in Table 1.

A dataset based on Alphanso mangoes from Mysore, Karnataka, was prepared and tested using various machine learning classifiers such as Threshold-based, Naïve Bayes, LDA, SVM, KNN and PNN. The hierarchical classification gave approximately 83% accuracy and single-shot multiclass classification gave 82% accuracy (Raghavendra *et al.,* 2020).

Crop yield prediction is quantified by deep learning algorithms where convetional methods are applied and results in the lower prediction error and a strong correlation between predicted and actual values (Kalmani *et al.,* 2025).

A study on mango plant disease detection using ConvNext and Vision Transformer (ViT) on the Mango Fruit DDS dataset achieved 98.40% accuracy and on the Mango Leaf BD dataset gave 99.87% accuracy. These datasets are publicly available and cover various diseases observed on mango fruit and mango leaves (Alamri *et al.*, 2025).

MATERIALS AND METHODS

The work is carried out at the Department of Computer Science, Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India and Ramrao Adik Institute of Technology, D.Y. Patil Deemed to be University, Nerul, Navi Mumbai, India during January 2025 to August 2025.

The entire process of the implementation of the work starts with preprocessing and data augmentation. The preprocessed data is then fed to the fusion of CNN-ViT, where feature extraction is tested with variants of CNN, ShuffleNet, MobileNet and ResNet. These features are then fed to ViT for classification. The entire architecture is visualised through Fig 1.

Datasets used in the work

Formation of the dataset, pre-processing and augmentation are crucial steps when the algorithm is trained and tested. The two datasets based on the mango classification are mainly focused on in the current work. The datasets used are focused on variants of mango and the mango ripeness level. Both datasets are publicly available in online mode.

For the sake of ease, throughout the paper, the datasets are referred to as the Mango Ripeness Dataset, which contains images related to mango ripeness levels and the Mango Classification Dataset, which contains images of various mangoes.

The datasets are restructured in training and testing folders as per the recommendation given in the research work prescribed by Mehta *et al.* (2021). Both folders consist of subfolders containing images in various groups. The images are organized in the form of the problem. The Mango Ripeness Dataset consists of classes such as Unripe, Early Ripe, Partial Ripe and Ripe (Prabhu, 2024). The Mango Classification Dataset consists of variants of mangoes found in Pakistan (Shahane, 2022). The varieties like Anwar Ratool, Chaunsa (Black, Summer Bahisht, White), Dosehri, Fajri, Langra and Sindhri are included in the dataset. The details of the datasets are given in Table 2.

Data augmentation

To achieve better results, machine learning and deep learning algorithms require large, high-quality datasets. The manual annotation and balancing of the dataset

Table 1: Result obtained on the mango dataset (Sikder *et al.*, 2025).

	CNN	VGG16
Gradient boosting	96.28%	93.72%
Random Forest	95.9%	93.59%
SVM	95.77%	93.97%
KNN	95.13%	93.97%
Gaussian Naïve Bays	86.92%	42.13%

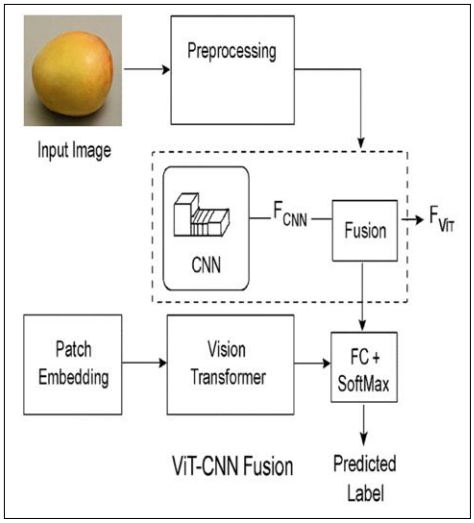


Fig 1: The workflow of CNN-ViT.

consume time. The data augmentation process is shown to be efficient (Mumuni and Mumuni, 2022). Multiple techniques are employed to enhance and balance the dataset. Geometric transformation, colour transformation are common data augmentation is common techniques applied to the dataset (Mumuni and Mumuni, 2022). In the proposed work, images from the dataset are converted to a tensor with resized images of 224 × 224. The geometric transformation rotation with degrees 15, flipping horizontally, is applied to the images randomly. The images are cropped to a size of 224 randomly with a padding size of 4. All the images are normalised with mean values of 0.485, 0.456, 0.406 and standard deviation values of 0.229, 0.224, 0.225 for each colour channel. All four datasets used in this work are augmented with these steps before the algorithm is applied.

Tested algorithms

The datasets (Mango Ripeness Dataset, Mango Classification Dataset) are inputted to traditional Convolutional Neural Network, MobileNet, VGG16, ResNet 50. All these models are initially judged without any data augmentation steps. The Convolutional Neural Network uses 2 convolutional layers, 2 Max Pooling layers, 2 Dense layers with SoftMax activation. To enhance the CNN's performance, data augmentation is applied and the number of layers is increased.

All three algorithms (MobileNet, VGG16, ResNet) are initially implemented with the default architecture. No augmentation is applied to the dataset before the algorithm is implemented. Adam optimiser is used and activation on the dense layer is applied using ReLU and SoftMax. To avoid overfitting, early stopping is used.

In the implementation of the proposed ViT transform, the model building was tried with pre-trained ResNet, MobileNet and ShuffleNet. Each Dataset is checked against the ViT transform with one of the pretrained models. Each dataset is initially augmented. Cross-entropy loss is used to calculate the loss in each step. Adam optimizer is used. The used architecture is given in the Fig 2.

RESULTS AND DISCUSSION

Initially, the Mango Ripeness Dataset, Mango Classification Dataset were tested with CNN, MobileNet, VGG16 and ResNet 16. These algorithms are implemented using the default architecture the result is much lower than expected. Results are discussed in Table 3. The research mentioned that the CNN sometimes fail to recognize the fruit due to various factors such as illumination, occlusion and surface deformation (Picon *et al.*, 2020; Huang *et al.*, 2023). Since the MobileNet can capture features easily and can avoid overfitting on smaller datasets makes it more effective (Tan *et al.*, 2022). Since the ViT has superior capacity in long range modelling and capturing context in features of the image hence to enhance the accuracy of the algorithm ViT transform is used (Dosovitskiy *et al.*, 2021; Islam *et al.*, 2021; Huang *et al.*, 2023).

The comparative analysis given in Table 3 states that the result obtained from the implementation of the

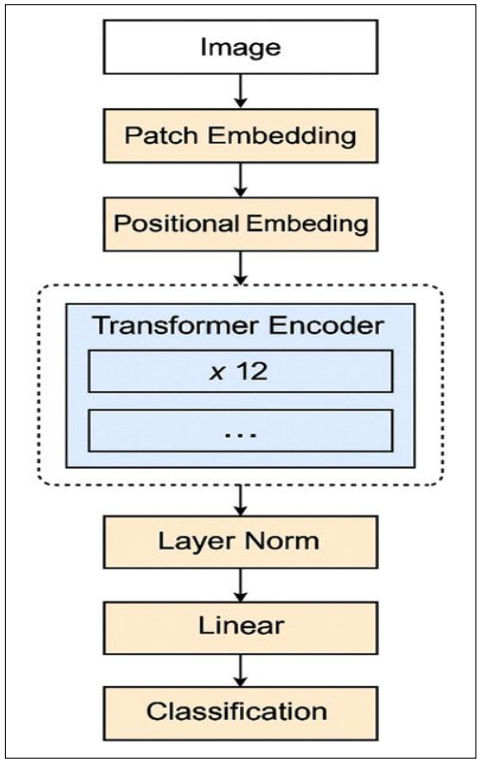


Fig 2: The proposed ViT transform.

Table 2: Details of the dataset.

Dataset	No of classes	No. of images in training folder	No. of images in test folder	Total images
Mango ripeness dataset	4	786	166	952
Mango classification dataset	8	1360	240	1600

Table 3: Comparison of results of datasets with the default architectures of CNN, MobileNet, VGG16 and ResNet16.

Name of the dataset	CNN	MobileNet	VGG16	ResNet 16
Mango ripeness dataset	84.15	84.15	73.17	46.95
Mango classification dataset	88.74	94.16	89.99	25.41

Table 4: Result of proposed ViT transformer with pretrained Shufflenet, Mobilenet and ResNet model.

Epoch	Mango ripeness dataset			Mango classification datatset		
	ShuffleNet	MobileNet	ResNet	ShuffleNet	MobileNet	ResNet
10	96.39	98.19	99.40	99.17	100	99.50
20	98.19	98.80	98.80	100	100	100
30	97.59	98.19	99.40	99.58	100	100
40	98.80	97.59	98.19	100	100	100
50	98.19	98.19	98.80	100	100	100
60	96.99	98.19	98.19	100	100	100
70	98.80	98.19	98.80	100	100	100
80	97.59	98.19	98.19	100	100	100
90	98.19	98.19	99.40	100	100	100
100	98.19	98.19	97.59	100	100	100

MobileNet algorithm is much higher than the other three algorithms. To obtain more accuracy, the proposed work is concentrated on the ViT transform with a pretrained model.

All the pretrained models, such as Shufflnet, MobileNet and ResNet, are executed and implemented with the ViT transform. The gradual increase by 10 in the epochs is used to observe the algorithm's behaviour against each dataset. The threshold values of the epoch where the algorithm gave a consistent result, or the increase/ decrease in the result, are observed.

The behaviour of the Shufflenet, MobileNet and ResNet algorithm against each dataset is noted in Table 4. The epoch wise validation accuracy is observed for each dataset. The epoch wise evaluation will give a strong insight for stability, convergence and overfitting behaviour (Kamilaris and Prenafeta-Boldú, 2020). It is being mentioned that after 90 epochs, the Mango Ripeness Dataset gave a consistent behaviour. The Mango Classification Dataset gave 100% accuracy after 40th epochs. For MobileNet, the Mango Ripeness Dataset gave 98.19% accuracy after 50th epochs. The Mango classification dataset achieved 100% accuracy after the first 10 epochs. The pretrained ResNet model was applied to both datasets. The Mango Ripeness Dataset gave approximately 98% accuracy. The Mango Classification Dataset gave 100% accuracy.

It is observed that both datasets gave the same result as ShuffleNet and MobileNet. The Mango Classification Dataset had given 100% accuracy once they reached to their threshold value in ResNet, MobileNet and ShuffleNet. The Mango Ripeness Dataset gave 98.19% accuracy with Shufflenet and MobileNet but 97.59% accuracy in the application of the ResNet algorithm.

A one-way ANOVA test was performed to compare ShuffleNet, MobileNet and ResNet across training epochs for both datasets. For the Mango Ripeness dataset, the ANOVA result ($p = 0.0213$) indicated significant differences. Wilcoxon tests showed no significant difference between ShuffleNet and MobileNet ($p=0.445$), while ResNet performed significantly worse than both.

For the Mango Classification dataset, ANOVA also revealed significant differences ($p = 0.0006$). Wilcoxon tests again confirmed that ShuffleNet and MobileNet performed similarly, whereas ResNet showed significantly lower performance.

It is being observed that among the three pertained models, ResNet, MobileNet and ShuffleNet, the ResNet dataset gave the lowest performance as compared with ShuffleNet and MobileNet in the Mango Ripeness Dataset. The pretrained models ShuffleNet and MobileNet gave the same result for both datasets. Rather, the threshold value for the consistent result is different in both algorithms. Since more images are present in the Mango Classification Dataset, the performance obtained for this dataset is 100%. This suggest that lightweight CNN backbone combined with ViT gave higher accuracy reducing computational cost which can be used in real time quality monitoring system (Tan *et al.*, 2022; Huang *et al.*, 2023).

CONCLUSION

The limitations of existing mango classification methods are tested to identify those that lead to a robust, dataset-independent solution. To leverage the solution, a fusion of CNN and ViT was tested across diverse datasets, reporting not only accuracy but also robustness. The consistent accuracy (98.19% and 100% respectively) for the Mango ripeness dataset and Mango Classification Dataset is observed in pre-trained variants of CNN such as ShuffleNet, MobileNet and ResNet. The statistical evaluation strongly support the proposed model ViT- ShuffleNet and ViT-MobileNet are more robust and reliable for classification problem. The further study can be applied to self-repositories and real-world agricultural environments. Moreover, lightweight versions for mobile phones can be used to check the on-site mango quality assessment for farmers. Overall, this work establishes a reproducible and scalable foundation for automated fruit quality evaluation.

ACKNOWLEDGEMENT

The present study was supported by the guidance and encouragement of Dr. Vaibhav E. Narawade (Professor,

RAIT, Mumbai) and Dr. Neha Jain (Assistant Professor, PAHER University). The authors also acknowledge the facilities and cooperation provided by the Department of Computer Science, PAHER University, Udaipur.

Disclaimers

The views and conclusions expressed in this article are solely those of the authors and do not necessarily represent the views of their affiliated institutions. The authors are responsible for the accuracy and completeness of the information provided, but do not accept any liability for any direct or indirect losses resulting from the use of this content.

Informed consent

All animal procedures for experiments were approved by the Committee of Experimental Animal Care and Handling Techniques were approved by the University of Animal Care Committee.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Abayomi-Alli, O.O., Damaševièius, R., Misra, S. and Abayomi-Alli, A. (2024). FruitQ: A new dataset of multiple fruit images for freshness evaluation. *Multimedia Tools and Applications.* **83(4):** 11433-11460. https://doi.org/10.1007/s11042-023-16058-6.

Aishwarya, N. and Vinesh Kumar, R. (2023). Banana Ripeness Classification with Deep CNN on NVIDIA Jetson Xavier AGX. Proceedings of the 7th International Conference on I-SMAC. *Institute of Electrical and Electronics Engineers*. https://doi.org/10.1109/I-SMAC58438.2023.10290326.

Akshi, Varshney, P., Avasthi, S. and Agarwal, K. (2024). Fruit and vegetable classification and freshness detection using machine learning. Proceedings of the MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon 2024). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/MITADTSoCiCon60330.2024.10575107.

Alamri, F.S., Sadad, T., Almasoud, A.S., Aurangzeb, A. and Khan, A. (2025). Mango disease detection using fused vision transformer with ConvNeXt architecture. *Computers, Materials and Continua*. https://doi.org/10.32604/cmc.2025.061890

Ali, S., Ibrahim, M., Ahmed, S.I., Nadim, M., Rahman, M.R., Shejunti, M.M. and Jabid, T. (2022). *MangoLeafBD dataset* (Version 1). Mendeley Data. https://doi.org/10.17632/hxsnvwty3r.1.

Begum, N. and Hazarika K.M. (2022). Deep learning based image processing solutions in food engineering: A Review. *Agricultural Reviews.* **43(3):** 267-277. doi: 10.18805/ag.R-2182.

Bhat, A., Khan, F. and Mir, I.A. (2023). Impact of image preprocessing techniques on the performance of deep learning models for mango fruit classification. *Information Processing in Agriculture.* **10(3):** 450-462.

Bobde, S., Jaiswal, S., Kulkarni, P., Patil, O., Khode, P. and Jha, R. (2021). Fruit quality recognition using deep learning algorithm. Proceedings of the international conference on smart generation computing, communication and networking (SMART GENCON 2021). *Institute of Electrical and Electronics Engineers*. https://doi.org/10.1109/SMARTGENCON51891.2021.9645793.

Das, J., Chakraborty, S. and Ghosh, D. (2022). Non-destructive mango quality assessment using support vector machine on near-infrared spectral data. *Postharvest Biology and Technology.* **185:** 111780.

Devi, P., Meena, S. and Jain, R. (2024). Enhanced mango fruit classification using ensemble of multiple deep learning models. *Artificial Intelligence in Agriculture.* **8:** 100120.

Dhiman, B., Kumar, Y. and Hu, Y.C. (2021). A general-purpose multi-fruit system for assessing the quality of fruits using recurrent neural networks. *Soft Computing.* **25(14):** 9255-9272. https://doi.org/10.1007/s00500-021-05867-2.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T. and Houlsby, N. (2021). An image is worth 16×16 words: Transformers for Image Recognition at Scale. ICLR.

Geerthik, S., Senthil, G.A., Oliviya, K.J. and Keerthana, R. (2024). A system and method for fruit ripeness prediction using transfer learning and CNN. Proceedings of the International Conference on Communication, Computing and Internet of Things (IC3IoT 2024). *Institute of Electrical and Electronics Engineers*. https://doi.org/10.1109/IC3IoT60841.2024.10550209.

Huang, Z., Zhang, C., Wang, L. and Wang, H. (2023). Fruit disease detection and classification using CNN and transformer hybrid models. *Expert Systems with Applications*. **213:** 119070.

Islam, M.Z., Hossain, M.S. and Andersson, K. (2021). Vision transformer-based fruit disease recognition using attention-guided feature extraction. *IEEE Access*. **9:** 165404-165414.

Joseph, T., Mathew, G. and Abraham, S. (2021). Development of a mobile application for mango variety identification using fine-tuned convolutional neural networks. *Applied Engineering in Agriculture.* **37(6):** 987-995.

Kalmani, V.H., Dharwadkar, N.V. and Thapa, V. (2025). Crop yield prediction using deep learning algorithm based on CNN-LSTM with attention layer and skip connection. *Indian Journal of Agricultural Research*. **59(8):** 1303-1311. doi: 10.18805/IJARe.A-6300.

Kamilaris, A. and Prenafeta-Boldú, F.X. (2020). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*. **147:** 70-90.

Li, W., Zhang, Q., Wei, R. and Chen, S. (2022). High-accuracy mango variety classification using deep convolutional neural networks with attention mechanisms. *Computers and Electronics in Agriculture.* **198:** 107025.

Mehta, D., Sehgal, S., Choudhury, T. and Sarkar, T. (2021). Fruit quality analysis using modern computer vision methodologies. Proceedings of the IEEE Madras Section International Conference (MASCON 2021). *Institute of Electrical and Electronics Engineers*. https://doi.org/10.1109/MASCON51689.2021.9563427.

Mumuni, A. and Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array.* **16:** 100258. https://doi.org/10.1016/j.array.2022.100258.

Oltean, M. (2025). Fruit-360 Dataset [Dataset]. Kaggle.

Patel, R., Sharma, V. and Gupta, A. (2023). Real-time mango maturity grading based on lightweight CNN architectures for embedded systems. *Journal of Agricultural Engineering Research.* **34(2):** 112-125.

Patil R.G. and More A. (2025). A comparative study and optimization of deep learning models for grape leaf disease identification. *Indian Journal of Agricultural Research*. **59(4):** 654-663. doi: 10.18805/IJARe.A-6242.

Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J. and Johannes, A. (2020). Deep convolutional neural networks for mobile capture device-based crop disease classification. *Computers and Electronics in Agriculture*. **161:** 104892.

Prabhu, A. (2024). Alphonso Mango Ripening Stage Classification Dataset (Version 1). Mendeley Data. https://doi.org/10.17632/tyghd6gxw2.1

Raghavendra, A., Guru, D.S., Rao, M.K. and Sumithra, R. (2020). Hierarchical approach for ripeness grading of mangoes. *Artificial Intelligence in Agriculture.* **4:** 243-252. https://doi.org/10.1016/j.aiia.2020.10.003.

Reddy, L., Devi, K. and Rao, M. (2020). Transfer learning for efficient mango disease classification with limited data. *Plant Pathology Journal.* **36(4):** 380-387.

Shahane, S. (2022). Mango Varieties Classification and Grading [Dataset]. Kaggle.

Sikder, M.S., Islam, M.S., Islam, M. and Reza, M.S. (2025). Improving mango ripeness grading accuracy: A comprehensive analysis of deep learning, traditional machine learning and transfer learning techniques. *Machine Learning with Applications.* **19:** 100619. https://doi.org/10.1016/j.mlwa.2025.100619.

Singh, S., Verma, N. and Kumar, P. (2021). Mango defect detection using random forest classifier with optimized feature selection from digital images. *Precision Agriculture.* **22(5):** 1456-1478.

Suresh, K., Priya, R. and Mohan, S. (2024). Comparative analysis of machine learning algorithms for mango maturity classification using color and texture features. *Journal of Horticultural Science.* **79(1):** 55-68.

Tan, S., Zhang, Y., Li, W. and Xu, J. (2022). A lightweight deep learning model for fruit ripeness and disease classification based on MobileNet and attention mechanism. *Computers and Electronics in Agriculture*. **198:** 107054.

Yuan, Y., Chen, J., Polat, K. and Alhudhaif, A. (2024). Detecting fruit and vegetable freshness through integration of convolutional neural networks and bidirectional long short-term memory networks. *Current Research in Food Science.* **8:** 100723. https://doi.org/10.1016/j.crfs.2024.100723.