INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Check for updates

OPEN ACCESS

Received: 17/09/2026

Accepted: 08/10/2026

Published: 25/10/2026

Citation: Suneetha , Rashmi BS, Nandini HM (2026) A Hybrid Deep–Texture Feature Fusion Framework for Coconut Maturity Classification Using MobileNetV2 and Local Binary Patterns. Indian Journal of Science and Technology 19(38): 2639-2648. https://doi.org/10.17485/IJST/v19i38.1293

* **Corresponding author**.

rashbsh@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2026 Suneetha et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

A Hybrid Deep–Texture Feature Fusion Framework for Coconut Maturity Classification Using MobileNetV2 and Local Binary Patterns

Suneetha[1], **B S Rashmi**[2]*, **H M Nandini**[2]

1 Research Scholar, Department of Studies and Research in Information Technology, Karnataka State Open University, Mysuru, India
2 Assistant Professor, Department of Studies and Research in Information Technology, Karnataka State Open University, Mysuru, India

Abstract

Objectives: The main objective is to explore and devise a hybrid deep-handcrafted texture fusion approach for automated coconut image-based maturity classification using MobileNetV2 and Local Binary Patterns. This approach aims to enhance discrimination of three coconut maturity stages: tender, mature and overmature. **Method:** A newly created coconut dataset of 3050 coconuts annotated as tender, mature and overmature were used. The proposed framework integrates the lightweight MobileNetV2 CNN Model with handcrafted texture descriptor via feature fusion analogy. ImageNet is employed to pretrain MobileNetV2 model to extract deep features, while uniform LBP captures local texture information from coconut images. The feature vectors thus obtained are fused and prediction process is performed using Softmax classifier. The efficacy of the model was evaluated using three independent runs and standard classification metrics like accuracy, precision, recall and F1-score. **Findings:** The proposed MobileNetV2–LBP framework achieved 88.82% $\pm$ 1.67% mean validation accuracy across three runs, compared with 73.01% $\pm$ 0.19% for the MobileNetV2 baseline. The highest individual-run accuracy was 91.04%. A paired t-test across the three runs indicated a statistically significant improvement ($t(2) = 14.75$, $p = 0.0046$). The final run achieved 91.04% accuracy, with recalls of 97%, 86% and 86% and F1-scores of 0.95, 0.87 and 0.89 for tender, mature and overmature coconuts respectively. The results demonstrate that LBP provides complementary local texture information to MobileNetV2's deep semantic characterization. The empirical results indicate effective deep–handcrafted feature fusion for three-stage coconut maturity classification when compared with CNN baseline. **Novelty:** The study evaluates an application-oriented hybrid model by fusing MobileNetV2 deep representations with LBP texture features which substantially improves three-stage coconut maturity discrimination by providing complementary semantic and local texture information.

Keywords: Coconut maturity classification; MobileNetV2; Local Binary Pattern; Feature fusion; Smart agriculture

1 Introduction

Assessment of coconut (Cocos nucifera L.) maturity is essential in indicating time for using and processing. Visual clues of varied maturity levels are difficult by the fact that similar colour, shapes and textures can be present in coconut images. By the advancement of computer vision and deep-learning methods, it has become feasible to determine and categorize agricultural products automatically. Review of related studies in food processing and fruit image analysis have exhibited the use of Machine Learning (ML), Convolutional Neural Networks (CNN), Transfer Learning (TL)[(1)]. This study illustrates the problems related to computational overhead, lack of datasets, visual cues etc.,[(2)].

Recent works on coconut maturity have been focused on object detection and categorization by CNN. Faster R-CNN with ResNet-50 has been employed as backbone architecture to detect tender and mature coconut in complex environments[(3)]. In another work YOLOv5s, YOLOv4-Tiny and Mobile Net-SSD approach has been adapted for maturity determination work[(4)]. Both works have indicated the usage of raw and ripened coconut images. Similar to this, CNN and YOLOv8 architectures have been used to detect young and matured coconuts[(5)]. These works have proved the efficacy of deep learning models in coconut maturity detection, but their approaches lack in determining closely related maturity stages.

A latest study has adapted EfficientNet, MobileNetV2 and VGG16 models for classification work on 3000 images of coconuts representing mature, pre-mature and young stages. The results reported varied considerably with the type of architecture used[(6)]. In contrast, acoustic modality of features was employed with a deep learning approach for maturity detection purpose[(7)]. Another work emphasized on using wavelet scattering transform with MobileNetV2 model for maturity prediction[(8)]. Although these studies have worked on acoustic data, the visual patterns are not exploited directly from coconut images.

An attempt has been made by implementing a hybrid approach by integrating fuzzy-deep learning approaches using Mask R-CNN to boost assessment of coconut maturity[(9)]. A two-step framework was introduced combining EfficientNetB0 and DenseNet121 for maturity detection and MobileNetV2 for harvest time prediction[(10)]. Although hybrid approaches were introduced, their combined efforts were lacking in exploring the contribution of lightweight CNN and local texture descriptor for three stage coconut maturity classification.

Most recent work focused on adapting YOLOv8, YOLOv9 YOLOv10 in conjunction with DETR, EfficientNet and Vision transformer-based methodology on a field recorded image dataset with varying degrees of occlusion[(11)]. Such studies determine continued efforts in the detection-oriented maturity analysis. But as per observation, they do not directly reveal the local texture information that can enhance the discrimination capability among similar maturity stages.

Earlier approaches have emphasized the criticality of varied dataset and their feature representation. A hybrid machine learning framework incorporating pre-processing and estimation approaches using heterogeneous agricultural data has been devised[(12)]. To enhance feature representation of agricultural images, a modified CNN approach with Jacobi polynomial and Laplacian based methods particularly for plant disease recognition was introduced[(13)]. These works have potential in revealing complementary feature representations rather than single learned representations.

For efficient texture analysis, Local Binary Pattern (LBP) is a well-established handcrafted descriptor for representing local spatial patterns[(14)]. In literature, a lot of variants have been proposed for enhancing the discriminative ability like Mid-hinge Local Binary Pattern for abrupt shot-change detection[(15)]. A very recent variant has been demonstrated combining LBP information with deep contour features for texture classification[(16)]. These studies present their intrinsic complementary information to learned representations across reported applications.

As per the literature study, the unresolved research gap is not the absence of deep-learning based coconut maturity recognition, but the limited investigation of the deep-handcrafted local texture feature representation. However, the reviewed literature provides limited evidence on fusing MobileNetV2 features with LBP local texture features for distinguishing coconut maturity stages. This motivates the feature fusion strategy to extract contextual features and subtle texture patterns related to varied coconut maturity stages. The main contributions of this work are:

- Creation of CocoNucifera Custom Dataset (CCD), comprising 3,050 annotated coconut images.
- Application and evaluation of MobileNetV2-LBP feature fusion framework for three state image-based coconut maturity classification.
- Comparison of the classification performance of the proposed MobileNetV2-LBP framework with the MobileNetV2 baseline model.

2 Methodology

The proposed work develops a hybrid deep learning framework for coconut image-based maturity classification as shown in Figure 1. The potentiality of the work is evaluated using the newly created coconut dataset. A two-stage feature extraction work is executed to achieve robust classification of coconut maturity stages. Initially, MobileNetV2 CNN model is employed to extract

deep features followed by LBP descriptor to extract handcrafted texture features. Subsequently, a concatenated feature vector is formed and propagated through Softmax classifier to categorize coconut maturity stages such as tender, mature or overmature. All the coconut images in the dataset undergo a pre-processing stage and is carried forward to the feature extraction stage.

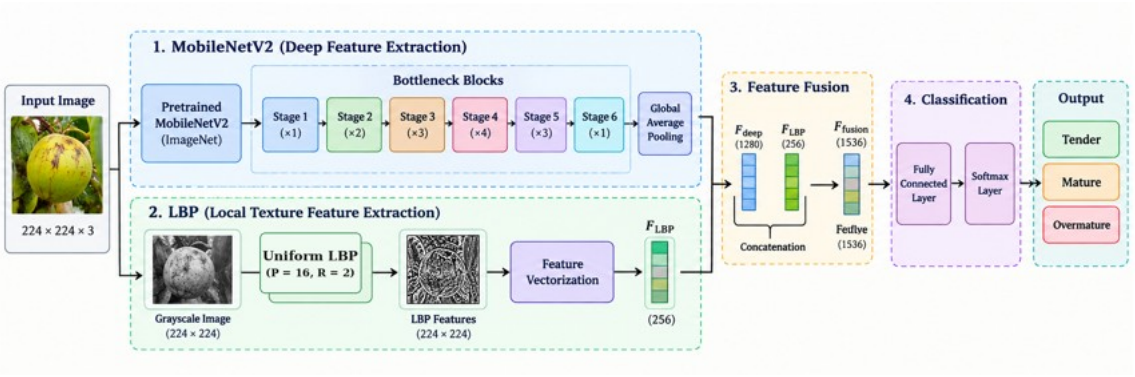


Fig 1. Architecture of the proposed Hybrid Deep-Texture Fusion Framework

2.1 MobileNetV2 Deep Feature Extraction

MobileNetV2 is a widely used lightweight CNN deep learning model to extract semantic features from coconut images as illustrated in Figure 1. Every input image I is set to 224×224×3 resolution and provided to pre-train the MobileNetV2 network model. Here, inverted residual blocks and depth wise separable convolutions are used as an aid to reduce the number of parameters and computational complexity. This operation is expressed as:

$$Y_{i,j,k} = \sum_{m,n} K_{m,n,k} X_{i+m,j+n,k} \tag{1}$$

The inverted residual block is derived as:

$$F(X) = P_2(D(P_1(X))) \tag{2}$$

This pipeline facilitates MobileNetV2 to learn discriminative high-level representations useful and aids in reducing the overhead cost.

In order to produce contextual information, the above convolutional feature maps are converted into a compact feature vector using Global Average Pooling (GAP). The k[th] feature channel is expressed as:

$$f_k = \frac{1}{HW} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{i,j,k} \tag{3}$$

The GAP operation formulates with feature vector that is denoted by:

$$F_{deep} = \text{GAP}(F_{MobileNetV2}(I)) \tag{4}$$

F_{deep} is the deep feature extractor that represents contextual characteristics of the coconut image and is further used for feature fusion.

2.2 LBP Texture Feature Extraction

In order to represent the roughness and fibre patterns in the coconut image, LBP value is computed for every pixel in the coconut image which is defined as:

$$LBP_{(P,R)} = sum_{(p=0)}^{(P-1)} s(g_p - g_c) 2^p \tag{5}$$

Where P = 16, R = 2 and $s(g_p - g_c) = \{ 1, g_p >= g_c ; 0, g_p < g_c \}$.

Further, in order to depict the rotation texture representation, the uniform LBP formulation is employed based on the number of transitions in circular fashion and is expressed as:

$$U(LBP_{(P,R)}) = |s(g_{(P-1)} - g_c) - s(g_0 - g_c)| + sum_{(p=1)}^{(P-1)} |s(g_p - g_c) - s(g_{(p-1)} - g_c)| \quad (6)$$

A texture histogram for each coconut image is produced using the frequency of each LBP pattern as:

$$H_{(i)} = sum_{(x,y)} I\left(LBP_{(P,R)}(x,y) = i\right) \quad (7)$$

Then, the histogram is normalized as:

$$H_{norm(i)} = H(i)/(sum_{(j=0)}^{(N-1)} H(j) + epsilon) \quad (8)$$

Finally, the LBP feature vector is produced using the normalized histograms that represent the spatial intensity variations of all the consecutive coconut images denoted as:

$$F_{LBP} = \left[H_{norm(0)}, H_{norm(1)}, ..., H_{norm(N-1)}\right] \quad (9)$$

2.3 Feature Fusion

The integrated hybrid representation is constituted by deep features and hand-crafted texture features to help in detecting coconut maturity stages which is formulated as:

$$F_{fusion} = \left[F_{deep} || F_{LBP}\right] \quad (10)$$

Where F_{deep} and F_{LBP} illustrates the MobileNetV2 deep feature vector and the normalized LBP feature vector respectively. For a 224 × 224 input image, the MobileNetV2 backbone yields a 1280-dimensional global representation after Global Average Pooling. A fully connected layer reduces the representation dimension to 256. The uniform LBP descriptor P=16 and R=2 produces a 58-dimensional histogram in the implemented configuration, which is normalized by its total histogram count and projected to 256-dimensions with a fully connected layer. The obtained deep and texture representations are concatenated to form a fused feature vector. No further normalization is applied to the projected MobileNetV2 feature before the fusion.

2.4 Classification Layer

In order to determine the three maturity classes: Tender, Mature and Overmature the fused feature representation F_{fusion} is propagated through fully connected layers to compute class scores such as $F_1 = ReLU\left(W_1 F_{fusion} + b_1\right)$ and $F_2 = ReLU\left(W_2 F_1 + b_2\right)$ respectively. Subsequently, Softmax layer function transforms the scores into probabilities for each maturity class as follows:

$$P(y = k|F_2) = exp(z_k)/sum_{(j=1)}^{3} exp(z_j) \quad (11)$$

The final predicted maturity stage belongs to the class possessing the highest probability.

2.5 Performance Evaluation

The evaluation criteria used for the maturity predicted results are accuracy, precision, recall and F1-score. Accuracy measures the proportion of correctly classified samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

Precision represents the actual correctness of the positive predictions for a specific class and recall measures the model's ability to recognize the actual instances of that class.

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \quad (13)$$

To evaluate false and missed classified classes, precision and recall measures are combined to depict F1-score which is expressed as:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (14)$$

2.6 Dataset Description

CCD[17] is developed by capturing coconut images in agricultural land and from local vendors located in Mysuru region of Karnataka, India. During the acquisition process, 3,050 coconut images were captured using high-resolution digital camera under varied illumination conditions and with varied angles. The developed dataset is annotated with assistance of farming personnel into three classes 'tender', 'mature' and 'overmature' as depicted in Table 1.

Labeling was performed with the help of five farming personnel experienced in coconut planting and maturity determination. Each coconut's maturity was defined based on its visual parameters such as color, overall appearance, size, shape and husk characteristics. The annotation stage did not include any measurements of the coconut's interior and was not laboratory-based but instead relied on the visual inspection of the coconut's external features. The dataset was split at the image level and thus images of the same specimen were not necessarily grouped together. This should be validated with a specimen-level split, grouping all images of a single coconut in one partition.

Table 1. Class-wise distribution of the coconut maturity dataset

Class	Total images	Training	Validation
Tender	1337	936	401
Mature	1056	740	316
Overmature	657	460	197
Total	3050	2136	914

The dataset exhibited moderate class imbalance, with tender coconuts representing the largest class and overmature coconuts the smallest. The distribution of the total number of datasets was fixed between experiments, yet the baseline and MobileNetV2–LBP experiments had a different procedure of splitting into partitions. The baseline utilized the Keras' validation split=0.30, while the MobileNetV2–LBP model implemented a stratified train_test_split with random state corresponding to each experimental seed.

2.7 Experimental Settings

The experimental setup was implemented using Python and TensorFlow/Keras to promote three effective coconut multistage maturity classification such as tender, mature and overmature. The MobileNetV2 CNN model is initialized with ImageNet-trained weights and LBP was employed for encoded spatial information of coconut images. Also, during the training process data augmentation techniques like zoom, rotation and horizontal flipping were deployed. The experimental analysis examined the effectiveness over three independent runs using different random seeds with other specifications as tabulated in Table 2:

Table 2. Specifications of the Parameter

Parameter	Setting
Dataset	3,050 coconut images
Classes	Tender, Mature, Overmature
Train/Validation	70% / 30%
MobileNetV2	ImageNet pretrained
Input size	128 × 128 (baseline), 224 × 224 (proposed)
Optimizer	Adam, LR = 0.0001
Epochs	20
LBP	P = 16, R = 2, Uniform
Experimental runs	3 (Seeds: 1, 21, 42)
Evaluation	Accuracy, Precision, Recall, F1, ROC-AUC

The proposed MobileNetV2-LBP framework was trained on 224 × 224 sized RGB images with a batch size of 8. The MobileNetV2 architecture was initialized with ImageNet pretrained weights. All the MobileNetV2 layers preceding the fine-tuned layers were frozen during training while the last 30 layers were fine-tuned. The Deep Branch used Global Average Pooling followed by a fully connected layer with 256 units with ReLU activation and 0.5 dropout. The LBP branch regressed a 256-unit ReLU fully connected layer with 0.5 dropout. The feature vectors from both the LBP and Deep Branch Regressor were then concatenated and inputted into a fully connected layer with 512 units with ReLU activation and 0.6 dropout. This was followed

by another fully connected layer with 256 units with ReLU activation and a three-unit Softmax classifier for the final predictions. The algorithm was optimized using the Adam optimizer with a learning rate of 0.0001 and a categorical cross-entropy loss function. The proposed fusion model did not utilize any online image augmentations. The experiments were carried out three times independently with random seeds 1, 21 and 42. There was no explicit early-stopping or model selection policy based on a best-checkpoint criterion. The model was trained for the designated number of epochs and the weights realized at the end of the training period were used for evaluation.

3 Results and Discussion

3.1 Performance Comparison of Models

Table 3 compares the classification performance and computational cost of the two models, showing a substantial improvement in the proposed MobileNetV2–LBP framework with higher accuracy, precision, recall and F1-score at the expense of increased training time.

Table 3. Classification performance and computational efficiency of the models

Model	Mean Accuracy	Std. Deviation	Weighted Precision	Weighted Recall	Weighted F1-score	Mean Training Time(s)	Highest Individual-Run Accuracy (%)
MobileNetV2 (Baseline)	0.7301	0.0019	0.73	0.73	0.72	11036.32	73.19
MobileNetV2 +LBP (Proposed)	0.8882	0.0167	0.91	0.90	0.90	29541.28	91.04
Improvement	+15.81	-	-	-	-	+18,504.96	+17.85

Initially both EfficientNet and MobileNetV2 convolutional architectures were deployed for coconut maturity prediction work. EfficientNet results were reported with meagre 43.87% validation accuracy revealing with misclassification of mature or overmature samples. In contrast to this, MobileNetV2 experimental results were significantly raised to 73.19% validation accuracy. Accordingly, MobileNetV2 is facilitated as a baseline model to carry out feature fusion analysis. Experimental analysis was carried out with a combination of EfficientNet features with LBP features and found an increase in results from 43.87% to 56.28% validation accuracy.

Further, integrating LBP features with the baseline MobileNetV2 resulted in a 15.81% improvement in mean accuracy. The proposed model yields 91.04% accuracy compared to the baseline model with 73.19% accuracy. The baseline MobileNetV2 used 128×128 images, whereas the proposed MobileNetV2–LBP framework used 224×224 images as a potential confounding factor to preserve spatial texture information while evaluating the effect of incorporating LBP. Notably, texture feature fusion findings demonstrate an increase in recall ratio from 0% to 21% and 64% respectively. Thus, the LBP descriptor acts as a supporting aid in expressing the surface roughness and micro variation of coconut image in combination with deep features.

It is also observed that the proposed model consumes more mean training time (29,541.28) than baseline model (11,036.32) indicating computational overhead. Timing variation is caused due to differentiation in batch size, input resolution and tuning approach of both the models. The proposed MobileNetV2–LBP framework required greater training time than the MobileNetV2 baseline. The mean training time increased from 11,036.32 s (3.07 h) to 29,541.28 s (8.21 h), corresponding to an additional 18,504.96 s (approximately 5.14 h) or approximately 2.68 times the baseline training time. This additional computational cost was accompanied by an increase in mean validation accuracy from $73.01\% \pm 0.19\%$ to $88.82\% \pm 1.67\%$, representing a 15.81 percentage-point improvement. Thus, the proposed framework involves a measurable trade-off between higher classification performance and increased training cost. Whether this trade-off is appropriate depends on the intended application and available computational resources. Because inference latency was not measured in the present experiments, practical deployment efficiency cannot be assessed from training time alone. Future evaluation should therefore report per-image inference time, including LBP feature-extraction time, together with memory usage and computational requirements.

Figures 2(a) and 2(b) depict the classification accuracy and loss curve achieved by both the models respectively.

3.2 Class-wise Performance Analysis

Table 4 is detailed with the significant results of the fusion model across all coconut maturity classes in terms of precision, recall and F1-score.

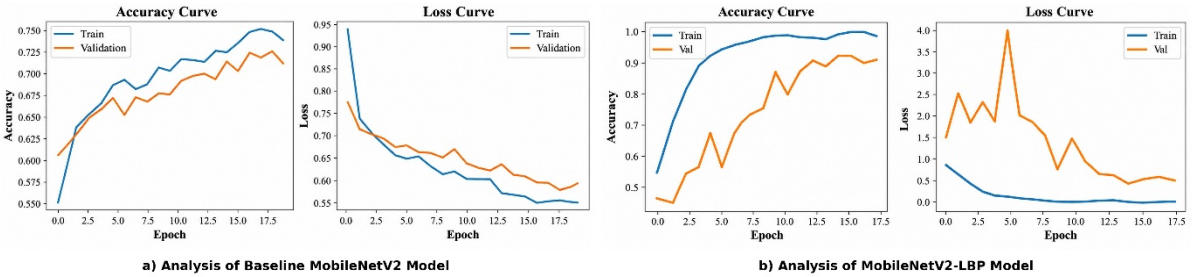


Fig 2. Classification Accuracy and Loss Curve

Table 4. Class-wise performance comparison

Class	Model	Precision	Recall	F1-score
Tender	MobileNetV2	0.71	0.89	0.79
	Fusion Model	**0.92**	**0.97**	**0.95**
Mature	MobileNetV2	0.70	0.49	0.58
	Fusion Model	**0.88**	**0.86**	**0.87**
Overmature	MobileNetV2	0.82	0.81	0.82
	Fusion Model	**0.92**	**0.86**	**0.89**

Compared with the MobileNetV2 baseline, the fusion model increased recall from 0.89 to 0.97 for tender coconuts, from 0.49 to 0.86 for mature coconuts and from 0.81 to 0.86 for overmature coconuts, corresponding to improvements of 8, 37 and 5 percentage points respectively.

The improvement witnessed in MobileNetV2–LBP might be explained by the fact that the information contained in local surface patterns is complementary to that captured by MobileNetV2. The coconut maturity can present different color, shape, texture, illumination and brightness[3],[4]. The local binary pattern describes local intensity relationship and texture, which complements the high-level semantic features extracted by MobileNetV2. This improvement can be linked to the increase of mature class recall from 0.49 to 0.86. However, fibre structure, shell roughness and other properties affecting the visual appearance of a coconut were not measured independently in this study and thus should be considered as potential sources of the improvement.

Further confusion matrix illustrates the coconut maturity classification in Figures 3(a) and 3(b) respectively. It can be noticed that mature coconuts have been misclassified as tender or overmature due to subtle differences in the visual features. It is observed that variant features of LBP integrated with contextual information retrieved from MobileNetV2 aids in the prominent decline of misclassification.

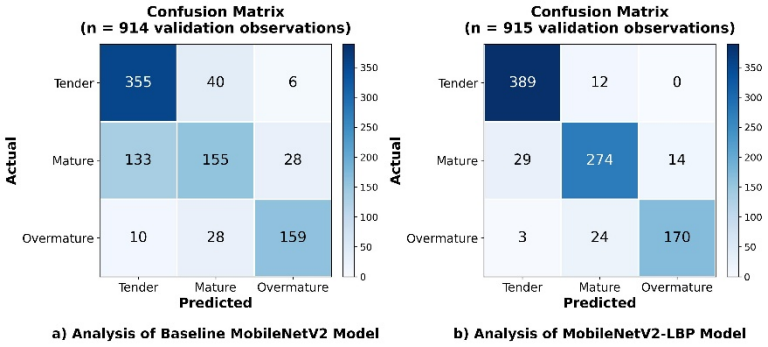


Fig 3. Confusion Matrix Analysis

The Receiver Operating Characteristic (ROC) curves have been illustrated in Figures 4(a) and 4(b) respectively. The fusion model accomplished greater Area Under the Curve (AUC) values for all the classified maturity stages than the baseline model.

The baseline MobileNetV2 achieved AUC values of 0.825, 0.960 and 0.899 for tender, mature, and overmature coconuts, respectively, whereas the MobileNetV2–LBP model achieved 0.983, 0.965 and 0.980 respectively.

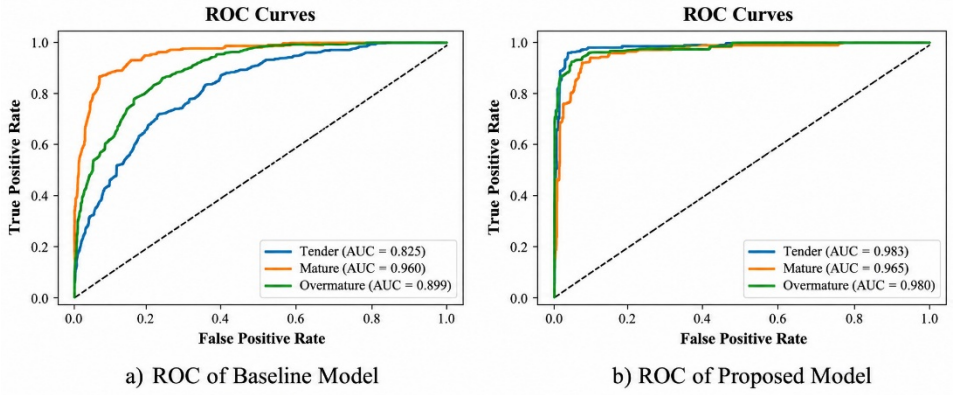


Fig 4. Receiver Operating Characteristic (ROC) curves

3.3 Contextual Comparison of Reported Coconut Maturity Assessment Studies

During empirical study, the results of the proposed coconut maturity assessment were carried out with the recent state-of-the-art approaches as depicted in Table 5. Usually, the existing studies implementation work varies with imaging technique, maturity distinctions, dimension of the dataset, evaluation protocols etc., and reported performances are not directly comparable with proposed results.

Table 5. Contextual comparison of reported coconut maturity assessment studies

Study	Modality	Dataset/sample size	Maturity classes	Method	Reported performance
	Image	3000 images	3 stages	EfficientNet	80.17% accuracy
De Luna et al. [6]	Image	3000 images	3 stages	MobileNetV2	65.33% accuracy
	Image	3000 images	3 stages	VGG16	96.17% accuracy
Sattar [8]	Acoustic	122 signals	3 stages	WST + MobileNetV2	73.12% accuracy
Rizki & Halim [7]	Acoustic	129 signals	3 stages	CNN + LSTM	98.36% accuracy
Perera & Perera [10]	Image	Dataset-specific	3 stages	EfficientNetB0 + DenseNet121	99% accuracy
Proposed study	image	3050 images	3 stages	MobileNetV2 + LBP	91.04% accuracy

De Luna et al. adapted EfficientNet, MobileNetV2 and VGG16 architectures to carry out the coconut maturity classification work [6]. Experimental analysis was carried out using 3000 coconut images representing mature, pre-mature and young categories and their results reported with accuracy of 80.17%, 65.33% and 96.17% respectively. The proposed framework achieved a mean validation accuracy of 88.82% $\pm$ 1.67%, while the highest individual-run validation accuracy of 91.04% was observed across the three runs. Among image-based studies, this result exceeds results of MobileNetV2 and EfficientNet reported by De Luna et al. but lower than their VGG16 result of 96.17%. However, this interpretation cannot justify that one architecture is better than the other, since there are no identical evaluation procedures and protocols.

In relation to acoustic sensing of coconut maturity, Sattar adapted WST+MobileNetV2 CNN architecture and reported 73.12% accuracy [8]. Similar to this sensing modality Rizki and Halim employed CNN-LSTM model for the same three maturity stages and reported 98.36% testing accuracy and F1-score of 99% [7]. Both studies predicted results on three stage coconut maturity classification similar to the proposed model by considering acoustic data rather than RGB image data.

Perera and Perera adapted a dual EfficientNetB0–DenseNet121 framework in conjunction with MobileNetV2 and reported classification accuracy above 99% [10]. It is noticed that two deep CNN architectures aid with significant results. In particular, the proposed approach does not combine two deep CNN backbones for classification, but combines one lightweight CNN representation with a handcrafted local texture representation.

The proposed MobileNetV2–LBP framework achieved a mean validation accuracy of 88.82% $\pm$ 1.67% across three independent runs, with a highest individual-run accuracy of 91.04%. These results are presented as a contextual comparison with previous coconut maturity studies rather than as evidence of superiority, because the reported studies employ different datasets, maturity definitions, sensing modalities, acquisition conditions and evaluation protocols.

4 Conclusion

In this work, a hybrid model was developed by fusing MobileNetV2 deep features with LBP texture descriptors for three-stage coconut maturity classification: tender, mature and overmature. The main objective was to integrate the high-level semantic features extracted by MobileNetV2 with the local texture characteristics represented by LBP. The experimental evaluation was conducted on a newly created dataset containing 3,050 coconut images. The MobileNetV2–LBP framework achieved a mean validation accuracy of 88.82% $\pm$ 1.67% across three independent runs, compared with 73.01% $\pm$ 0.19% for the MobileNetV2 baseline. The highest individual accuracy was 91.04%, while the recalls for tender, mature and overmature coconuts were 97%, 86% and 86% respectively.

The main contribution of this study is the application and evaluation of a deep-handcrafted texture fusion strategy, in which MobileNetV2 provides semantic representations and LBP captures local texture characteristics. Across the three experimental runs, the MobileNetV2–LBP configuration achieved higher validation accuracy than the baseline, with improvements from 72.76% to 86.99%, 73.09% to 88.42%, and 73.19% to 91.04% for seeds 1, 21 and 42 respectively. The statistically significant distinction between the two configurations was obtained using a paired t-test ($t(2) = 14.75$, $p = 0.0046$). However, this result should be interpreted as evidence of a difference between the evaluated experimental configurations rather than as an independent estimate of the effect of LBP, since the baseline and hybrid models also differed in certain training settings.

The evaluation was performed on a single dataset of 3,050 images and therefore does not establish generalization across different coconut varieties, geographical locations, cameras, backgrounds, illumination conditions or field environments. In addition, the comparison was based on an internal validation set rather than an independent external dataset. The hybrid framework also required more training time than the MobileNetV2 baseline, indicating additional computational cost associated with the fusion architecture and training configuration. These factors should be considered for practical deployment.

Future studies should include cross-dataset validation using coconut images from different geographical locations, cultivars, camera devices and acquisition environments. Evaluation under controlled variations in illumination, occlusion, background complexity, viewing angle and image quality would further assess robustness under real-world conditions. A controlled ablation study using identical training settings should also be conducted to isolate the contribution of LBP from other architectural and training differences. Further research may investigate attention-based feature fusion to learn the relative importance of deep and local texture representations. In addition, multimodal fusion of visual, acoustic, spectral and other non-destructive sensing information could be explored for more comprehensive coconut maturity and quality assessment.

Overall, the study provides experimental evidence that combining MobileNetV2 deep representations with handcrafted LBP texture information is a promising approach for three-stage image-based coconut maturity classification. The work demonstrates the potential of combining complementary semantic and local-texture representations within a deep-handcrafted fusion framework, while further cross-dataset and real-world validation is required before practical deployment.

5 Acknowledgements

NIL

Author details: Suneetha is a Research Scholar; B S Rashmi, H M Nandini are Assistant Professors

References

1) Zhu L, Spachos P, Pensini E, Plataniotis KN. Deep learning and machine vision for food processing: a survey. *Current Research in Food Science*. 2021;4:233–249. https://doi.org/10.1016/j.crfs.2021.03.009.
2) Espinoza S, Aguilera C, Rojas L, Campos PG. Analysis of fruit images with deep learning: a systematic literature review and future directions. *IEEE Access*. 2024;12:3837–3859. https://doi.org/10.1109/ACCESS.2023.3345789.
3) Parvathi S, Selvi ST. Detection of maturity stages of coconuts in complex background using Faster R-CNN model. *Biosystems Engineering*. 2021;202:119–132. https://doi.org/10.1016/j.biosystemseng.2020.12.002.
4) Mandava RK, Mittal H, Hemalatha N. Identifying the maturity level of coconuts using deep learning algorithms. *Materials Today: Proceedings*. 2024;103:410–414. https://doi.org/10.1016/j.matpr.2023.09.071.
5) Luthfie AA, Alamsyah. Identifying coconut maturity levels using CNN and YOLOv8 deep learning algorithms. *Journal of Information Systems Exploration and Research*. 2025;3(2):53–62. https://doi.org/10.52465/joiser.v3i2.595.

6) de Luna RG, Alegre JCM, Martinez RAP, Ramos MRN, Sadian HMA, Villanueva OGG. Coconut maturity level classification using image preprocessing and deep transfer learning. *2025 RIVF International Conference on Computing and Communication Technologies (RIVF)*. 2025;p. 1140–1144. https://doi.org/10.1109/RIVF68649.2025.11365215.
7) Rizki FM, Halim A. Improving the accuracy of deep-learning models for coconut maturity classification using acoustic signals. *Jurnal Pekommas*. 2025;10(2):45–52. https://doi.org/10.56873/jpkm.v10i2.6004.
8) Sattar F. Predicting maturity of coconut fruit from acoustic signal with applications of deep learning. *Biology and Life Sciences Forum*. 2024;30(1):16. https://doi.org/10.3390/IOCAG2023-16880.
9) Megalingam RK, Manoharan SK, Maruthababu RB. Integrated fuzzy and deep learning model for identification of coconut maturity without human intervention. *Neural Computing and Applications*. 2024;36(11):6133–6145. https://doi.org/10.1007/s00521-023-09402-2.
10) Perera PA, Perera PM. Dual-stage deep learning framework for hybrid coconut maturity classification and harvest timeline prediction. *Asian Journal of Research in Computer Science*. 2025;18(9):76–90. https://doi.org/10.9734/ajrcos/2025/v18i9755.
11) Tejaswi K, Madhushree B, Bharathi RK, Tejaswi S, Vinanthi S. Tender coconut maturity stage detection using YOLOv8-v10: comparative analysis with DETR, EfficientDet and vision transformers. *Procedia Computer Science*. 2026;282:135–145. https://doi.org/10.1016/j.procs.2026.05.062.
12) Leena KV, Chitra K. Hybrid machine learning framework for agricultural data collection and preprocessing for crop yield prediction. *Indian Journal of Science and Technology*. 2026;19(12):837–850. https://doi.org/10.17485/IJST/v19i12.236.
13) Janes PS, Chithra PL. A novel deep convolutional neural network approach using Jacobi polynomial and Laplacian function (JPLF) in recognition of plant leaf disease. *Indian Journal of Science and Technology*. 2024;17(14):1450–1463. https://doi.org/10.17485/IJST/v17i14.2651.
14) Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002;24(7):971–987. https://doi.org/10.1109/TPAMI.2002.1017623.
15) Nandini HM, Chethan HK, Rashmi BS. Abrupt shot change detection using Midhinge Local Binary Pattern. *2020 IEEE Hyderabad Section International Conference (IEEE-HYDCON)*. 2020;p. 1–5. https://doi.org/10.1109/HYDCON48903.2020.9242841.
16) Chang M, Ji L, Zhu J. Multi-scale LBP fusion with the contours from deep CellNNs for texture classification. *Expert Systems with Applications*. 2024;238:122100. https://doi.org/10.1016/j.eswa.2023.122100.
17) CocoNucifera Custom Dataset (CCD). *Google Drive*. 2026 Sep. Available from: https://drive.google.com/drive/folders/12OHxuj-Qboynbhfl1SEhFbEEYKv-8W1w?usp=sharing.