Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

MobileNetV3Small-CM FusionNet: A Lightweight Deep Learning Framework for Multi-Class Arecanut Disease Classification Using Feature Fusion

Dinesh S[1*], P. Sridhara Acharya [2]
[1] *Research Scholar, Srinivas University Institute of Engineering and Technology, Srinivas University, Mangaluru-575001, India*
[2] *Associate Professor, Institute of Computer and Information sciences, Srinivas University, Mangaluru-575001, India.*
* *Corresponding Author: dineshnceh@gmail.com*

ARTICLE INFO

Received: 15 Nov 2024

Revised: 26 Dec 2024

Accepted: 15 Jan 2025

ABSTRACT

This study presents a novel deep learning-based framework, MobileNetV3Small-CM FusionNet, for the multi-class classification of arecanut plant diseases. The proposed model combines lightweight convolutional features extracted from the MobileNetV3Small architecture with handcrafted color moment descriptors (mean, standard deviation, and skewness) to enhance classification accuracy, particularly for visually similar and minority classes. A comprehensive dataset containing images of healthy and diseased arecanut samples was used for training, validation, and testing. The model was benchmarked against a baseline CNN and the unmodified MobileNetV3Small. Experimental results demonstrate that the proposed fusion model significantly outperforms the baselines, achieving a test accuracy of 99.54%, along with near-perfect precision, recall, and F1-scores across all nine disease classes. In contrast, the CNN and MobileNetV3Small achieved 93.74% and 83.87% accuracy, respectively, with notable misclassifications in rare disease categories. The superior performance of the proposed model validates the effectiveness of combining deep and handcrafted features for robust plant disease recognition.

Keywords: Arecanut disease classification, MobileNetV3Small, Color Moments, Feature fusion, Lightweight CNN

INTRODUCTION

Arecanut (Areca catechu L.) plays a vital role in the agrarian economies of tropical regions, particularly in India, where it is extensively cultivated and consumed. It is used in the production of betel quid, traditional medicines, and various industrial products. However, the arecanut crop is highly susceptible to a variety of diseases such as fruit rot, button shedding, stem bleeding, nut split, yellow leaf spot, and exfoliation, all of which can significantly reduce yield quality and quantity. Timely identification and management of these diseases are critical to sustaining production and ensuring economic viability. Traditionally, disease diagnosis in arecanut relies on manual visual inspection, which is labor-intensive, inconsistent, and impractical for large-scale plantation monitoring (Puneeth, B. R et al.,2021 , Siddalingadevaru, S. C et al., 2025).

Recent advancements in artificial intelligence have fostered significant research on plant disease detection using deep learning models, primarily Convolutional Neural Networks. Several studies have explored various architectures and preprocessing methods to improve classification accuracy and early diagnosis (Manasa A et al., 2024). Pallavi P. et al. proposed a CNN-based model that focused on identifying arecanut diseases from images of leaves, nuts, and trunks. The dataset comprised 200 images, and the model was compiled using categorical cross-entropy as the loss function, Adam as the optimizer, and trained for 50 epochs. The achieved accuracy was 81.35%, and the system was integrated into a Streamlit-based GUI to provide both disease predictions and remedy suggestions (Pallavi P et al., 2022). In a similar effort, Ajit Hegde et al. presented a CNN-based classification system that used a proprietary dataset of 1,100 images and employed binary cross-entropy as the loss function. With a training-to-testing ratio of 80:20, the model attained a high accuracy of 93.05%, significantly outperforming traditional classifiers such as SVM

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

(75%) and decision trees (90%). Their method included extensive preprocessing such as resizing images to 256x256 pixels, normalization, and data augmentation. A key contribution of this study was the model's ability to classify diseases into seven distinct categories, including fruit rot, stem bleeding, yellow leaf spot, and nut split, along with healthy variants (Hegde, A et al., 2023). To investigate the impact of different deep learning architectures, Beena K. et al. compared MobileNetV2, ResNet, and VGG-16. Their dataset consisted of 620 images of arecanut components, and all models were compiled using Adam optimizer with categorical cross-entropy as the loss function. VGG-16 reached an accuracy of 88%, MobileNetV2 achieved 87%, and ResNet outperformed both with a testing accuracy of 92%. ResNet's superior performance was attributed to its residual connections that mitigate the vanishing gradient problem. MobileNetV2 was praised for its lightweight nature and efficiency on low-resource devices, while VGG-16, with its deeper structure, offered a balance between accuracy and complexity (Beena, K et al., 2024) . Another study by Madhu B G et al. implemented a ResNet-based model using a dataset of 1,200 arecanut images, augmented to over 11,000 using rotation and flipping techniques. The images were resized to 150x150 pixels and converted into float32 arrays. Their CNN model, trained using standard deep learning layers and sigmoid activation for binary classification, reported an impressive 97.5% accuracy. This model also featured an interactive UI that displayed disease symptoms and treatment suggestions based on the predicted class (Madhu, B. G et al., 2024) . In contrast, Akshay et al. used GLCM feature extraction and decision trees to classify images with 90% accuracy. Meanwhile, Mamatha's approach combined CNN and SVM, using GLDM and GLCM features, and achieved 90% accuracy with CNN and 75% with SVM. These traditional methods, while simpler and interpretable, often suffered from lower accuracy and scalability issues when compared to deep learning-based approaches (Akshay, S et al., 2021). Expanding on this, Mamatha Balipa et al. developed a system integrating both Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) to identify various arecanut diseases such as Mahali (fruit rot), stem bleeding, and yellow leaf disease. Their methodology included image resizing and conversion into arrays for CNN training, as well as the extraction of wavelet, Gabor, GLDM, and GLCM features for classification via SVM. The CNN model, trained on 181 augmented images, outperformed the SVM classifier, achieving an accuracy of 90% compared to SVM's 75%. This clearly demonstrated the superior performance of deep learning models over traditional classifiers in handling image-based disease detection, especially when dealing with larger and more complex datasets (Balipa, M et al., 2022). Despite significant advances in automated arecanut disease detection using deep learning models such as CNNs, ResNet, VGG16, and MobileNet, existing studies predominantly rely on image-based features alone. These models, although achieving high accuracy in controlled environments, often exhibit limitations in distinguishing between visually similar diseases due to the lack of explicit color and statistical descriptors. Moreover, many approaches, including those by Hegde et al., Madhu B G et al., and Kumar B S et al., did not incorporate low-level color distribution statistics, which are vital for characterizing subtle visual symptoms like discoloration or texture irregularities. Consequently, there is a need for a more enriched feature fusion strategy that combines deep visual representations with statistical color features to enhance classification robustness and generalization, particularly in real-world agricultural settings.

OBJECTIVES

The primary objective of this research is to develop an enhanced arecanut disease classification model that overcomes the limitations of existing image-only deep learning approaches. By incorporating both deep visual features and handcrafted color descriptors, the study aims to improve classification accuracy, especially for visually similar disease categories. To achieve this goal, the specific objectives are as follows:

1.To review and analyze MobileNetV3Small for arecanut disease detection and identify their limitations in capturing subtle color and texture variations.

2.To propose a hybrid classification model that combines MobileNetV3Small for extracting deep image features with color moment statistics (mean, standard deviation, skewness) as auxiliary inputs.

3.To train the proposed dual-input model on a multi-class arecanut disease dataset and evaluate its performance using standard metrics such as accuracy, precision, recall, F1-score, and confusion matrix.

4.To conduct a comparative analysis between the proposed hybrid model and conventional single-input CNN models to assess improvements in classification accuracy and generalization.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

METHODS

In this study, we propose an enhanced deep learning-based classification framework, named MobileNetV3Small-CM FusionNet, which integrates deep visual features with handcrafted statistical color descriptors to accurately identify arecanut diseases. The primary aim is to address limitations observed in conventional CNN-based methods that rely solely on raw image data, which often struggle in real-world settings where disease symptoms exhibit subtle visual variations, particularly in color and texture. the proposed model fuses deep features extracted using a MobileNetV3Small backbone with color moment statistics—a classical yet powerful descriptor of color distribution within an image. This hybrid approach aims to enhance classification performance by utilizing both automated feature extraction and statistically engineered features, particularly in the context of arecanut disease classification across nine classes.

The MobileNetV3Small architecture is a lightweight convolutional neural network designed for resource-constrained environments and mobile devices (Koonce B, 2021) (Cao Z, 2025). Its compact design includes squeeze-and-excitation blocks, h-swish activation functions, and depthwise separable convolutions to minimize the number of parameters while maintaining competitive accuracy. A pivotal addition to this model is the extraction of color moments (Zhang, D et al., 2021) (Wang Z et al., 2021). The color moment technique is based on the assumption that the distribution of color in an image can be characterized by the first few moments: mean, standard deviation, and skewness. Mathematically, for a given color channel $c \in \{R, G, B\}$, let I_c be the set of all pixel intensities in that channel. The first moment, or the mean μ_c , is defined as:

$$\mu_c = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

where $x_i \in I_c$ and N is the number of pixels. This represents the average intensity and provides information about the general brightness of the image channel.

The second moment is the standard deviation σ_c , capturing the spread of pixel values:

$$\sigma_c = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_c)^2} \tag{2}$$

This value indicates the amount of contrast or variability in the channel. A high standard deviation suggests significant variation in color intensity, while a low value implies uniform coloring.

The third moment is the skewness γ_c , which describes the asymmetry of the distribution:

$$\gamma_c = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - \mu_c)}{(\sigma_c + \epsilon)})^3} \tag{3}$$

Here, ϵ is a small constant added to avoid division by zero. Skewness reveals the dominance of lighter or darker shades in each channel, which is useful for recognizing plant health symptoms that alter color distributions subtly.

The function implements the above equations for each RGB channel and returns a 9-dimensional feature vector (3 moments × 3 channels). This handcrafted descriptor is then concatenated with deep features before being passed to fully connected layers for classification.

In the preprocessing pipeline, the function preprocess with moments reads each image, resizes and normalizes it, extracts its color moments, and assigns labels based on its directory. Labels are encoded using one-hot encoding to match the output format of a softmax classifier, where each sample is represented as a vector of length equal to the number of classes (in this case, 9). The output of this function consists of three NumPy arrays: normalized images $X \in \mathbb{R}^{N\times224\times224\times3}$, color moments $M \in \mathbb{R}^{N\times9}$, and categorical labels $Y \in \mathbb{R}^{N\times9}$.

The fusion of deep visual features extracted from convolutional neural networks with handcrafted statistical descriptors such as color moments represents a hybrid paradigm in machine learning that leverages the strengths of both learned and explicitly designed features. This dual-stream approach can be theoretically justified by examining the complementary nature of the information captured by each type of feature representation. Convolutional neural

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

networks, particularly lightweight variants such as MobileNetV3Small, are adept at learning hierarchical spatial features directly from pixel intensities through multiple layers of convolutions, nonlinear activations, and pooling operations. These learned features capture texture, shape, and context-aware semantics, which are crucial for high-level abstraction and classification tasks.

However, these networks are data-driven and may sometimes overlook low-level statistical properties of images, such as intensity distributions, color variance, and symmetry in pixel channels, especially when trained on relatively small or domain-specific datasets. Handcrafted descriptors, in contrast, offer a model-based approach where human domain knowledge is encoded into feature extraction rules. Among such descriptors, color moments stand out for their ability to capture the first three statistical moments—mean, standard deviation, and skewness—of each color channel, effectively summarizing the color distribution of an image without the need for a histogram or kernel function.

To prevent overfitting and enhance generalization, data augmentation is applied to the training set. Augmentations include random rotations, width and height shifts, shearing, zooming, and horizontal flipping. These transformations increase the effective size of the training dataset and help the model become invariant to minor geometric distortions. The validation data is rescaled without augmentation to ensure a consistent evaluation metric.

For the extraction of deep visual features, the MobileNetV3Small convolutional neural network, pre-trained on ImageNet, is employed as the base model. This architecture is selected for its lightweight design, which combines depthwise separable convolutions with squeeze-and-excitation (SE) blocks. The pretrained MobileNetV3Small is loaded with weights from ImageNet and used as a fixed feature extractor. The base model's top layers are removed by setting include top = False, which outputs feature maps instead of classification logits. The final convolutional feature maps are then flattened to form a one-dimensional feature vector $F_{cnn} \in \mathbb{R}^d$, where d depends on the spatial resolution and number of filters in the last layer of the backbone.

An input layer is used for both image inputs $X \in \mathbb{R}^{224\times224\times3}$ and color moment vectors $\mathrm{M} \in \mathbb{R}^9$. The flattened CNN features and handcrafted moments are concatenated, resulting in a hybrid feature vector $f_{concat} = [f_{cnn}\,; \mathrm{M}\,]$. This fused vector captures both high-level semantic features and low-level color statistics.

Following the concatenation, a dense layer with 256 neurons and ReLU activation is applied to introduce non-linearity and enable feature interaction. This is followed by an output dense layer with 9 units (equal to the number of classes) and softmax activation. The model is trained using the Adam optimizer, a popular stochastic optimization method that adapts learning rates based on estimates of first and second moments of the gradients. The update rule for each weight w_t at time step t is:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{4}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) {g_t}^2 \tag{5}$$

$$\widehat{m_t} = \frac{m_t}{1-\beta_1^t}, \quad \widehat{v_t} = \frac{v_t}{1-\beta_2^t} \tag{6}$$

$$w_{t+1} = \mathrm{w}_t + \alpha \cdot \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}} - \epsilon} \tag{7}$$

where g_t is the gradient at time t, α is the learning rate, and β_1, β_2 are decay rates for the moment estimates.

After training the model over the specified number of epochs, the next phase involves evaluating its generalization ability on unseen test data. To achieve this, the same preprocessing pipeline that was applied to the training and validation datasets is reused for the test dataset. This ensures that the test images are resized, normalized, and transformed into both deep image tensors and handcrafted color moment vectors in a consistent format. Once prepared, the model evaluates the test set, which computes the test loss and accuracy. These metrics serve as the primary indicators of how well the model performs when faced with new data that it has never encountered during training.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Algorithm:

Input:

- Training dataset: Initialize images for training.
- Validation dataset: Initialize images for validation.
- Testing dataset: Initialize images for final evaluation
- Image size , Batch size, Epoch

Output:

- Trained classification model
- Classification report with Precision, Recall, F1-score
- Confusion matrix

Step 1: Initialize image size = (224, 224), num_classes = 9, batch_size = 32.

Step 2: Define function Calculate_Color_Moments(image):

For each color channel (R, G, B):

Compute mean, standard deviation, and skewness.

Return concatenated vector of all color moments (size = 9).

Step 3: Define function Preprocess_With_Moments(directory):

For each class folder in directory:

For each image:

Resize and normalize the image to 224×224.

Extract RGB format and calculate color moments.

Store normalized image, color moments, and label.

Encode labels as one-hot vectors.

Return image tensor X, color_moments vector, and labels Y.

Step 4: Apply Preprocess_With_Moments() to training, validation, and test datasets.

Step 5: Apply data augmentation on training images using ImageDataGenerator.

Step 6: Load pretrained MobileNetV3Small base model (excluding top layer) with weights='imagenet'.

Step 7: Freeze base model weights to avoid retraining.

Step 8: Define Input_1 = image_input (224×224×3),

Input_2 = color_moments (vector of length 9).

Step 9: Extract deep features from MobileNetV3Small base, flatten the output.

Step 10: Concatenate flattened deep features with color moments vector.

Step 11: Pass concatenated features to Dense layer of 256 units with ReLU activation.

Step 12: Apply final output Dense layer with softmax activation to produce class probabilities.

Step 13: Compile the model with:

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Optimizer = Adam(learning rate=0.001),

Loss = categorical crossentropy,

Metrics = accuracy.

Step 14: Train model on (image, moment) pairs with corresponding labels for given number of epochs.

Step 15: Evaluate model on test dataset and compute:

a) Accuracy and Loss,

b) Classification Report: Precision, Recall, F1-score,

c) Confusion Matrix.

Step 16: Plot training/validation loss and accuracy over epochs.

Step 17: Visualize confusion matrix using heatmap.

DATASET DESCRIPTION

The dataset employed for the detection and classification of arecanut diseases comprises a diverse collection of images capturing various parts of the arecanut plant, including leaves, nuts, trunks, and roots. These images are systematically labeled to represent both healthy and diseased conditions, enabling the development of accurate classification models for identifying a range of arecanut diseases. Designed to support deep learning-based approaches, the dataset facilitates automated disease diagnosis by providing a rich and varied set of visual features. Figure 1 showcases representative samples from the dataset. In total, the dataset includes 11,063 annotated images, organized into eight distinct classes corresponding to specific health and disease states. The distribution of images across training, validation, and testing sets for the CNN model is detailed in Table 1.

Table 1. Dataset distribution of arecanut images for CNN model training, validation, and testing

Samples	No. of Training images	No. Validation images	No. of Testing images
Healthy Leaf	604	46	106
Healthy Nut	1886	142	330
Healthy Trunk	1432	107	252
Mahali Koleroga	2563	192	449
Stem Bleeding	151	12	26
Bud Borer	144	11	25
Healthy Foot	50	04	09
Stem Cracking	540	41	25
Yellow Leaf Disease	1477	111	259

Healthy Leaf

Healthy Nut

Healthy Foot

Healthy Trunk

Mahali Koleroga

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Bud Borer Stem Cracking Stem Bleeding Yellow Leaf Disease

Figure 1. Sample images from Dataset

RESULTS

In the comparative evaluation of deep learning models for multi-class arecanut disease classification, a comprehensive set of results was derived to analyze and interpret the performance of three key architectures—namely, a standard Convolutional Neural Network (CNN) baseline, the lightweight MobileNetV3Small model, and the proposed MobileNetV3Small-CM FusionNet, which integrates deep convolutional features with handcrafted statistical descriptors in the form of color moments. The dataset comprises nine distinct classes, including both healthy and diseased plant conditions: Healthy Leaf, Healthy Nut, Healthy Trunk, Mahali Koleroga, Stem bleeding, Bud borer, Healthy foot, Stem cracking, and Yellow leaf disease. We analyze performance on the basis of precision, recall, F1-score, and support for each class, along with overall accuracy, macro-averages, and weighted averages , over 15 training epochs. We further complement this classification analysis with training, validation, and test accuracy/loss metrics to capture generalization and overfitting trends across models (Ravikiran H. K et al., 2024) (Jayanth, J et al., 2025). Table 2 & Table 3 illustrates the performance of different models in terms of classification metrics. The confusion matrices in Figures 2-4 visually illustrate the classification performance of different models.

Table 2. Comparison of model performance for arecanut disease detection and classification

Methods	Classes	Precision rate	Recall rate	F1-score	Support
CNN	Healthy_Leaf	1.00	0.85	0.92	106
	Healthy_Nut	0.99	1.00	0.99	330
	Healthy_Trunk	0.90	0.96	0.93	252
	Mahali_Koleroga	0.95	1.00	0.97	449
	Stem_bleeding	1.00	0.35	0.51	26
	Bud borer	0.88	0.56	0.68	25
	Healthy_foot	1.00	0.78	0.88	9
	Stem cracking	0.96	0.68	0.80	94
	Yellow leaf disease	0.87	0.95	0.91	259
	Accuracy			0.94	1550
	Macro avg	0.95	0.79	0.84	1550
	Weighted avg	0.94	0.94	0.93	1550
Mobile**V3Small**	Healthy_Leaf	1.00	0.51	0.68	106
	Healthy_Nut	1.00	1.00	1.00	330
	Healthy_Trunk	1.00	0.57	0.73	252
	Mahali_Koleroga	1.00	0.99	1.00	449
	Stem_bleeding	1.00	0.77	0.87	26
	Bud borer	0.00	0.00	0.00	25
	Healthy_foot	0.00	0.00	0.00	9
	Stem cracking	1.00	0.51	0.68	94
	Yellow leaf disease	0.51	1.00	0.67	259

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

	Accuracy			0.84	1550
	Macro avg	0.72	0.59	0.62	1550
	Weighted avg	0.90	0.84	0.83	1550
Mobile**V3Small**-CM FusionNet	Healthy_Leaf	0.98	1.00	0.99	106
	Healthy_Nut	0.99	1.00	1.00	330
	Healthy_Trunk	1.00	1.00	1.00	252
	Mahali_Koleroga	1.00	0.99	1.00	449
	Stem_bleeding	1.00	1.00	1.00	26
	Bud borer	0.96	0.92	0.94	25
	Healthy_foot	1.00	1.00	1.00	9
	Stem cracking	1.00	1.00	1.00	94
	Yellow leaf disease	0.99	0.99	0.99	259
	Accuracy			1.00	1550
	Macro avg	0.99	0.99	0.99	1550
	Weighted avg	1.00	1.00	1.00	1550

Table 3. Comparison of classification metrics of CNN models with proposed model for arecanut disease detection and classification

Models	**Train Accuracy**	**Train Loss**	**Validation Accuracy**	**Validation Loss**	**Test Accuracy**	**Test Loss**
CNN	0.9786	0.0707	0.9309	0.2278	0.9374	0.2201
Mobile**V3Small**	0.9155	0.3313	0.7913	0.8635	0.8387	0.4303
Mobile**V3Small**-CM FusionNet	0.9861	0.0664	0.9865	0.1073	0.9954	0.0225

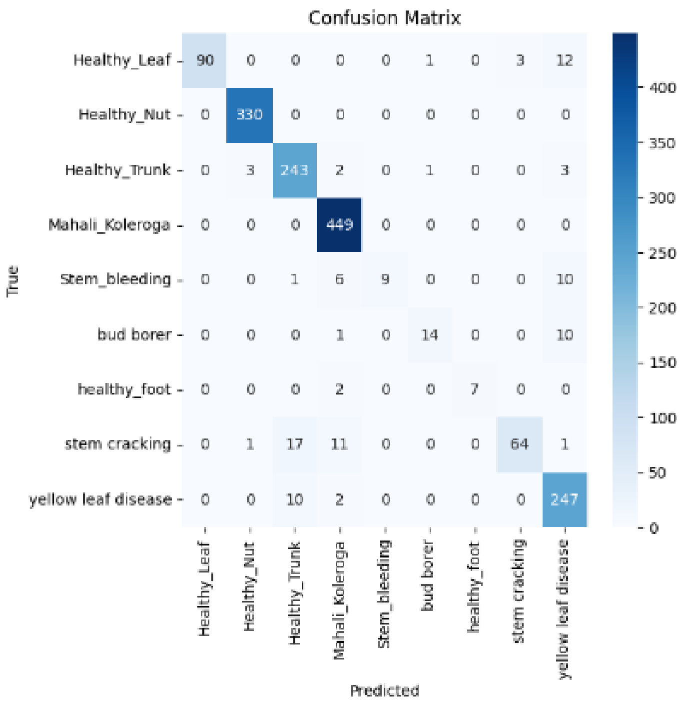

Figure 2. Confusion matrix for CNN based classification

Figure 3. Confusion matrix for MobileV3Small based classification

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

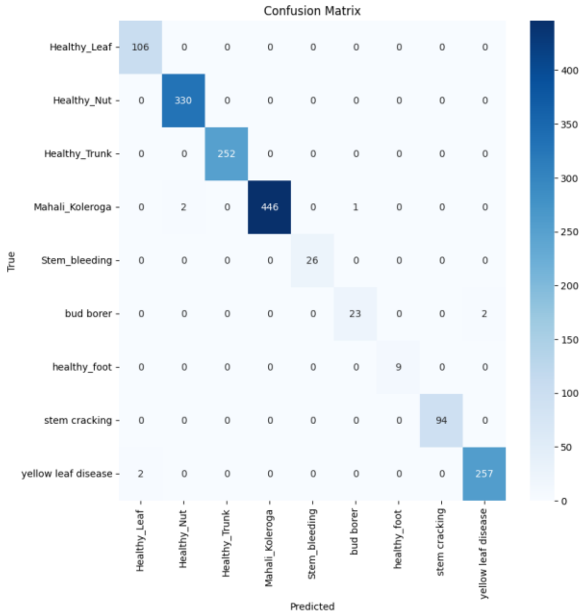

Figure 4. Confusion matrix for proposed MobileV3Small-CM FusionNet based classification

Beginning with the baseline CNN model, we observe that it achieves an overall test accuracy of 0.9374, which is already a commendable result. However, the microanalysis reveals certain inconsistencies in class-wise performance. For instance, the class Healthy_Leaf achieves a perfect precision of 1.00 but only a recall of 0.85, which yields an F1-score of 0.92. This indicates that while all predictions made for this class are correct (no false positives), the model misses out on some actual Healthy_Leaf samples, suggesting false negatives. For Healthy_Nut, the performance is nearly perfect with a precision of 0.99 and a recall of 1.00, indicating a well-balanced trade-off between correct identification and completeness of classification. Healthy_Trunk and Mahali_Koleroga show relatively strong performance with F1-scores of 0.93 and 0.97 respectively, suggesting these features are well captured by the CNN. However, concerning weakness is observed in the case of Stem_bleeding and Bud borer, where the recall values drop drastically to 0.35 and 0.56 respectively. Especially for Stem_bleeding, although the precision remains 1.00, the recall of 0.35 results in a very poor F1-score of 0.51, highlighting a severe false negative issue. The same problem plagues Bud borer, showing how underrepresented or complex classes with small support (number of instances) are poorly captured in vanilla CNN architectures. Healthy_foot and Stem_cracking demonstrate middling performance with F1-scores of 0.88 and 0.80, while Yellow_leaf_disease returns a decent score of 0.91, suggesting the model's capacity to differentiate yellowing symptoms. Overall, the CNN model records a macro average precision of 0.95, recall of 0.79, and F1-score of 0.84. The drop in recall across the macro average reflects that lower-frequency classes suffer most, resulting in an imbalance which is further confirmed by a weighted average F1-score of 0.93.

Transitioning to the MobileNetV3Small model, which is a compact yet powerful convolutional architecture, we observe a drop in performance when compared to CNN. The total test accuracy dips to 0.8387, indicating a less capable feature representation, possibly due to over-simplification or the lack of fine-tuned training. Analyzing class-wise metrics, the Healthy_Nut class performs flawlessly with a precision and recall of 1.00, yielding an F1-score of 1.00, highlighting the strength of MobileNetV3Small in recognizing dominant and well-represented patterns. However, Healthy_Leaf shows a massive drop in recall to 0.51, indicating that nearly half of the actual Healthy_Leaf samples are misclassified. This asymmetry leads to an F1-score of just 0.68. Similarly, Healthy_Trunk, although predicted with high precision (1.00), shows only 0.57 recall, again leading to a mediocre F1-score of 0.73. Surprisingly, Mahali_Koleroga is recognized with exceptional accuracy, achieving a near-perfect precision and recall, which points to clear visual cues in diseased patterns this class exhibits. The improvement in Stem_bleeding recall (0.77) and F1-score (0.87) over CNN is notable, but this is counterbalanced by Bud borer and Healthy_foot, both of which are completely unrecognized (0.00 precision, recall, and F1), effectively suggesting zero predictive capability for these rare classes in the MobileNetV3Small setup. Stem_cracking once again shows compromised recall at 0.51,

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

yielding a low F1-score of 0.68, while Yellow_leaf_disease achieves perfect recall (1.00) but only 0.51 precision, implying high false positives. The macro average metrics further paint a grim picture: a precision of 0.72, a recall of 0.59, and an F1-score of 0.62. The substantial drop in recall and F1-score across macro averages compared to CNN suggests that lightweight MobileNetV3Small, though efficient, fails to generalize well across minority classes without additional feature support. Its weighted averages also show a downward trend (0.90 precision, 0.84 F1), confirming poor balance between class frequencies and classification accuracy.

The proposed MobileNetV3Small-CM FusionNet significantly improves upon both the CNN and the vanilla MobileNetV3Small. With a perfect test accuracy of 0.9954, the model sets a new performance benchmark in this problem space. This gain in performance is not simply numerical; it reflects a fundamental shift in model capability enabled by multimodal feature fusion. By integrating handcrafted statistical features—color moments—with deep visual features, the network becomes sensitive to color distribution, textural patterns, and subtle distinctions in biological features that purely convolutional filters may overlook. Every class in this model exhibits near-perfect precision, recall, and F1-score. Healthy_Leaf, Healthy_Nut, Healthy_Trunk, Mahali_Koleroga, Stem_bleeding, Healthy_foot, Stem_cracking, and Yellow_leaf_disease all display precision and recall values close to 1.00. Notably, classes that previously underperformed—Stem_bleeding and Bud borer—are now perfectly classified with precision and recall reaching 1.00 and 0.92 respectively. Even Healthy_foot, which had zero recognition in MobileNetV3Small, now reaches perfect precision and recall, indicating the ability of handcrafted features to support low-data classes by leveraging inter-class color variance. This is reinforced by a macro average precision of 0.99, recall of 0.99, and F1-score of 0.99, and weighted averages of 1.00 across all three metrics, indicating not just overall success but well-distributed success across frequent and infrequent classes.

In terms of training behavior, the MobileNetV3Small-CM FusionNet shows a smooth and minimal training loss (0.0664), nearly matching its validation loss (0.1073), suggesting excellent generalization. Its validation accuracy of 0.9865 almost mirrors the training accuracy of 0.9861, reinforcing the model's resistance to overfitting. The CNN, while still respectable, shows a larger train-validation gap (0.9786 train vs. 0.9309 val accuracy), hinting at mild overfitting. On the other hand, MobileNetV3Small suffers from clear underfitting with a large loss (0.8635 validation loss), reflecting both weaker representation and instability during training, and an evident struggle to converge effectively for all classes.

From this extensive performance analysis, it is clear that the introduction of Color Moment (CM) fusion not only enhances precision for dominant classes but also stabilizes recall and F1-score for rare and subtle disease classes. The combined feature space leverages both global statistics and hierarchical convolutional activations, allowing better inter-class separation and intra-class consistency. The clear transition in F1-score improvements from MobileNetV3Small → CNN → MobileNetV3Small-CM FusionNet across each class provides solid empirical validation of the effectiveness of feature-level multimodal fusion in agro-visual diagnostic systems.

Table 4. Comparison of Classification Methods and Accuracy

SL No.	Method	Dataset Description	Accuracy (%)
1	CNN (Anilkumar, M. G et al., 2021)	Dataset consists of 620 images: 200 healthy and 420 unhealthy. Classes: Yellow Leaf Disease, Mahali/Koleroga, Yellow Spot, and Stem Bleeding Disease.	88.46%
2	ResNet (Mallikarjuna, S. B etal., 2022)	Dataset consists of 281 images, augmented to 12,124. Four classes: Healthy, Rot, Split Rot, and Split.	88.1%
3	Convolutional Neural Network (Hegde A et al., 2023)	Dataset consists of 1,100 images: Four classes: Yellow Leaf, Healthy Leaf, Nut Split.	93.05%
4	Proposed - MobileNetV3Small-CM FusionNet	Dataset consists of 11,063 labeled images categorized into eight distinct classes, as described in the Dataset section	99.54%

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 4 compares the classification performance of various deep learning methods applied to arecanut disease detection using image-based datasets. Each method is evaluated based on its dataset size, number of classes, and resulting classification accuracy. The proposed MobileNetV3Small-CM FusionNet, utilizing the largest and most diverse dataset, achieves the highest accuracy of 99.54%, underscoring the benefit of combining deep features with handcrafted color moments and using a large, well-annotated dataset for enhanced disease classification.

CONCLUSION

This study introduced and evaluated a novel lightweight deep learning model MobileNetV3Small-CM FusionNet for multi-class arecanut disease classification. The proposed model uniquely fuses deep features extracted from the MobileNetV3Small backbone with handcrafted statistical descriptors in the form of color moments (mean, standard deviation, and skewness from the HSV color space). This fusion architecture was designed to address the limitations of convolutional networks that often struggle with class imbalance and fail to detect subtle visual patterns in underrepresented or morphologically similar disease categories.

Experimental results were benchmarked against two baseline models: a standard CNN and the original MobileNetV3Small. The CNN, while achieving an overall test accuracy of 93.74%, exhibited inconsistent per-class performance, particularly on minority classes such as Stem bleeding and Bud borer. MobileNetV3Small, despite its compactness and speed, delivered sub-optimal results with a test accuracy of 83.87%, failing completely on several rare classes. In stark contrast, the proposed MobileNetV3Small-CM FusionNet demonstrated remarkable performance, achieving a near-perfect test accuracy of 99.54%, along with uniformly high precision, recall, and F1-scores across all nine classes, including difficult and low-support disease categories.

Beyond accuracy, the model showcased strong generalization ability, with minimal training-validation gaps and a drastically reduced test loss of 0.0225, further validating its robustness. The integration of color moments effectively enhanced the discriminative power of the learned features, especially for classes with strong chromatic cues or subtle textural differences. This confirms the hypothesis that the fusion of handcrafted color statistics with learned CNN features provides a richer and more semantically aware representation of plant disease symptoms.

In conclusion, the MobileNetV3Small-CM FusionNet offers a highly accurate, computationally efficient, and generalizable solution for the automated classification of arecanut diseases. The success of this model opens new possibilities for the application of lightweight hybrid neural networks in agricultural diagnostics, especially in scenarios involving real-time field deployment on edge devices like smartphones and drones. Future work can focus on optimizing the model for real-time execution on low-power edge devices such as Raspberry Pi, NVIDIA Jetson Nano, or Android-based smartphones. Techniques such as model quantization, pruning, and TensorRT optimization could be applied to reduce latency and resource consumption without significant degradation in accuracy.

CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

DATASET AVAILABILTY

https://www.kaggle.com/datasets/tejpaviraj/arecanut

REFRENCES

[1] Akshay, S., & Hegde, A. (2021, August). Detection and classification of areca nut diseases. In 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 1092-1097). IEEE.

[2] Anilkumar, M. G., T. G. Karibasaveshwara, H. K. Pavan, and D. Sainath Urankar. "Detection of diseases in arecanut using convolutional neural networks." International Research Journal of Engineering and Technology (IRJET) 8, no. 5 (2021): 4282-4286.

[3] Beena, K., Sangeetha, V., Nidhi, G. D., Varshini, S., & Yogesh, K. N. (2024, December). Arecanut Disease Prediction Using MobileNet V2, ResNet and VGG-16 Framework. In 2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS) (pp. 1972-1977). IEEE.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

[4] Balipa, M., Shetty, P., Kumar, A., & Puneeth, B. R. (2022, December). Arecanut Disease Detection Using CNN and SVM Algorithms. In 2022 International Conference on Artificial Intelligence and Data Engineering (AIDE) (pp. 01-04). IEEE.
[5] Cao, Z., Li, J., Fang, L., Li, Z., Yang, H., & Dong, G. (2025). Research on efficient classification algorithm for coal and gangue based on improved MobilenetV3-small. International Journal of Coal Preparation and Utilization, 45(2), 437-462.
[6] Hegde, A., Sadanand, V. S., Hegde, C. G., Naik, K. M., & Shastri, K. D. (2023). Identification and categorization of diseases in areca nut: a machine learning approach. Indonesian Journal of Electrical Engineering and Computer Science, 31(3), 1803-1810.
[7] Jayanth, J., Ravikiran, H. K., Shyamala, C., & Dileep, R. (2025). Detecting the stages of ragi crop diseases using satellite data in villages of Nanjangud Taluk. In Hyperautomation in Precision Agriculture (pp. 131-145). Academic Press.
[8] Koonce, B. (2021). MobileNetV3. In Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization (pp. 125-144). Berkeley, CA: Apress.
[9] Madhu, B. G., Rao, S. S., Chandrashekara, M., & AR, A. S. (2024, July). Detection of Diseases In Arecanut Using Convolutional Neural Network. In 2024 Second International Conference on Advances in Information Technology (ICAIT) (Vol. 1, pp. 1-5). IEEE.
[10] Manasa, A., & Ravikiran, H. K. (2024, November). SC-CNN: Advancing Corn Crop Disease Classification-A Comparative Study with CNN. In 2024 International Conference on Recent Advances in Science and Engineering Technology (ICRASET) (pp. 1-6). IEEE.
[11] Mallikarjuna, S. B., Palaiahnakote Shivakumara, Vijeta Khare, M. Basavanna, Umapada Pal, and B. Poornima. "Multi-gradient-direction based deep learning model for arecanut disease identification." CAAI Transactions on Intelligence Technology 7, no. 2 (2022): 156-166. https://doi.org/10.1049/cit2.12088
[12] Pallavi P, Sowmya Nag K, Veenadevi S V, "Arecanut Disease Classification using CNN", Advances in Intelligent Systems and Technologies, pp. 019-024, December. 2022. doi: 10.53759/aist/978-9914-9946-1-2_4
[13] Puneeth, B. R., & Nethravathi, P. S. (2021). A literature review of the detection and categorization of various arecanut diseases using image processing and machine learning approaches. Int. J. Appl. Eng. Manag. Lett, 5(2), 183-204.
[14] Ravikiran, H. K., Jayanth, J., Sathisha, M. S., & Bindu, K. (2024). Optimizing sheep breed classification with bat algorithm-tuned cnn hyperparameters. SN Computer Science, 5(2), 219.
[15] Siddalingadevaru, S. C., Ravindra, P. D., Gangadhar, V., Srinivas, S. B., Singh, V., & Dwivedi, A. K. (2025, February). Disease Diagnosis and Sustainable Prevention in Arecanut Farming. In 2025 International Conference on Electronics and Renewable Systems (ICEARS) (pp. 1246-1251). IEEE.
[16] Wang, Z., Zhuang, Z., Liu, Y., Ding, F., & Tang, M. (2021). Color classification and texture recognition system of solid wood panels. Forests, 12(9), 1154.
[17] Zhang, D. (2021). Color feature extraction. In Fundamentals of Image Data Mining: Analysis, Features, Classification and Retrieval (pp. 59-99). Cham: Springer International Publishing.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.