International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Dual-Stream MobileNetV3Small Fusion Architecture for Enhanced Multi-Class Detection of Arecanut Diseases Using Hybrid Feature Embedding

Dinesh S[1], P. Sridhara Acharya[2]

[1]Research Scholar, Srinivas University Institute of Engineering and Technology, Srinivas University, Mangaluru-575001, India
Corresponding Author Email: *dineshnceh[at]gmail.com*

[2]Associate Professor, Institute of Computer and Information Sciences, Srinivas University, Mangaluru-575001, India

Abstract: ***This research proposes an alternative approach for diagnosing multiple arecanut diseases by introducing a hybrid architecture known as MobileNetV3Small-CM FusionNet. Unlike conventional image-only models, the framework merges two complementary information streams: compact deep representations derived from the MobileNetV3Small backbone and statistical color descriptors computed through color moments. By integrating these feature types, the model is able to capture subtle chromatic variations and disease-specific patterns that are often difficult for lightweight CNNs to distinguish. The dataset used in this study comprises labeled images of both healthy and infected arecanut samples and is partitioned for training, validation, and testing. To assess effectiveness, the proposed system is compared with a standard CNN and the original MobileNetV3Small model. Evaluation results indicate a notable performance advantage for the fusion-based approach, which records a test accuracy of 99.54% and consistently high precision, recall, and F1-scores across all nine categories. In comparison, the benchmark models achieve considerably lower accuracy levels of 93.74% and 83.87%. These findings highlight the value of combining handcrafted statistical cues with deep feature embeddings for reliable and robust plant disease identification.***

Keywords: MobileNetV3Small, color moments, hybrid fusion model, arecanut disease detection, plant image classification

1. Introduction

Arecanut (Areca catechu L.) is a commercially valuable plantation crop cultivated widely across tropical regions, especially in India. The crop plays a crucial role in traditional medicine, industrial processing, and the production of betel quid. Despite its economic significance, arecanut cultivation faces severe productivity losses due to diseases such as fruit rot, button shedding, yellow leaf disease, trunk disorders, bud borer, and stem bleeding. Manual diagnosis of these diseases is laborious, subjective, and often prone to inconsistencies, making automated disease identification essential.

Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have demonstrated substantial progress in plant disease recognition. Prior studies include CNN-based arecanut classification models (Pallavi P. et al., 2022; Hegde A. et al., 2023), and architecture comparisons involving MobileNetV2, VGG-16, and ResNet (Beena K. et al., 2024). While these studies reported satisfactory results, limitations remain, especially when dealing with visually similar diseases or minority classes.

Most existing methods rely solely on image-based deep features, which may fail to capture low-level color variations that often characterize plant diseases. Studies employing feature engineering approaches, including GLCM, GLDM, and Gabor descriptors (Akshay S. et al., 2021; Balipa M. et al., 2022), highlight the potential of handcrafted features but lack the robustness and scalability of deep learning.

To address these gaps, this work proposes a hybrid feature fusion model that integrates lightweight deep representations with statistical descriptions of color distribution. This approach improves robustness and enhances the classification of diseases exhibiting subtle discoloration differences or limited sample AVAILABILITY.

2. Objectives

The study aims to build an improved arecanut disease classification framework that mitigates limitations of existing image-only models. The objectives are:

- To analyze the MobileNetV3Small architecture for disease classification and identify its limitations in recognizing fine-grained color-texture variations.
- To design a dual-stream hybrid model combining deep features with color moments (mean, standard deviation, skewness).
- To train and evaluate the proposed MobileNetV3Small-CM FusionNet using classification metrics such as accuracy, precision, recall, F1-score, and confusion matrix.
- To compare the proposed hybrid model with conventional CNN and MobileNetV3Small architectures.

3. Methods

1) Overview of the Proposed Framework

MobileNetV3Small-CM FusionNet is developed as a feature-fusion architecture combining:

Deep features extracted from MobileNetV3Small
Color moment descriptors capturing statistical color distribution

A dual-input fusion mechanism merges both representations before classification. The goal is to enhance recognition of subtle color deviations present in diseases such as stem bleeding and yellow leaf disease.

2) MobileNetV3Small Feature Extraction

MobileNetV3Small is chosen for its computational efficiency and high accuracy on resource-limited devices. It integrates depthwise separable convolutions, squeeze-and-excitation blocks, and h-swish activation, making it suitable for embedded and mobile systems.

The final convolutional layer outputs high-level visual features, which are flattened to form vector Fc nn.

3) Color Moment Descriptor Extraction

Color moments quantify the distribution of pixel intensities in RGB channels using:

(1) Mean

Dual-Stream MobileNetV3Small Fusion Architecture Mean:

The mean (μ) is equal to the sum of all values divided by the number of values (c).

$\mu = (\text{sum of all } x_i) / c$

(2) Standard Deviation

The standard deviation (σ) is calculated by:
Subtracting the mean from each value,
Squaring the differences,
Summing them,
Dividing by the number of values N,
Taking the square root.

$\sigma = \text{sqrt}((1/N) * \Sigma(x_i - \mu)^2)$

(3) Skewness

Skewness (γ) measures how asymmetric the distribution is around the mean. It is computed by:

$\gamma = (1/N) * \Sigma[(x_i - \mu)^3] / (\sigma + \varepsilon)$

where ε is a small constant added for numerical stability.
The output is a 9-dimensional feature vector (3 channels × 3 moments).

(4) Data Processing Pipeline

The preprocessing pipeline performs**:**
Image resizing to 224 × 224
RGB normalization
Color moment extraction
One-hot label encoding
The output consists of:
X: normalized images
M: color moment vectors
Y: categorical labels

(5) Fusion Architecture

The network uses two input layers:
The network takes two inputs:

- **Input 1:** An image tensor
- **Input 2:** A set of 9 color-moment values

The CNN output is flattened, and then combined with the color-moment vector by concatenating them:

f_concat = [F_cnn ; M]

A fully connected layer with **256 ReLU units** processes these fused features.

After that, a **softmax layer** outputs the probability scores for **nine disease classes**.

Training (Adam Optimizer, Plain Words)

Training uses the Adam optimizer. At each step:

1) The **momentum term (m^t)** is updated by blending the previous momentum with the current gradient.
2) The **variance term (v^t)** is updated similarly, but using the squared gradient.
3) The **weights (w)** are updated by subtracting a scaled version of the momentum divided by the square root of the variance (plus a small constant for stability).

In simple terms:

Adam adjusts the learning rate for each parameter based on both the average gradient and how much that gradient has varied over tim

(6) Algorithmic Outline

The complete step-wise algorithm is reproduced as provided (Steps 1–17), including preprocessing, augmentation, base model loading, feature fusion, training, and performance evaluation.

(7) Dataset Description

The dataset consists of 11,063 labeled images representing eight healthy and diseased categories of the arecanut plant.
Representative samples are displayed in Figure 1.

Table 1: Shows the dataset split across training, validation, and testing.

Samples 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

Healthy Leaf | Healthy Nut | Healthy Foot | Healthy Trunk | Mahali Koleroga

Bud Borer	Stem Cracking	Stem Bleeding	Yellow Leaf Disease

Figure 1: Sample images from Dataset

4. Results

4.1 Model Comparison

Three architectures were evaluated:

- CNN baseline
- MobileNetV3Small
- Proposed MobileNetV3Small-CM FusionNet

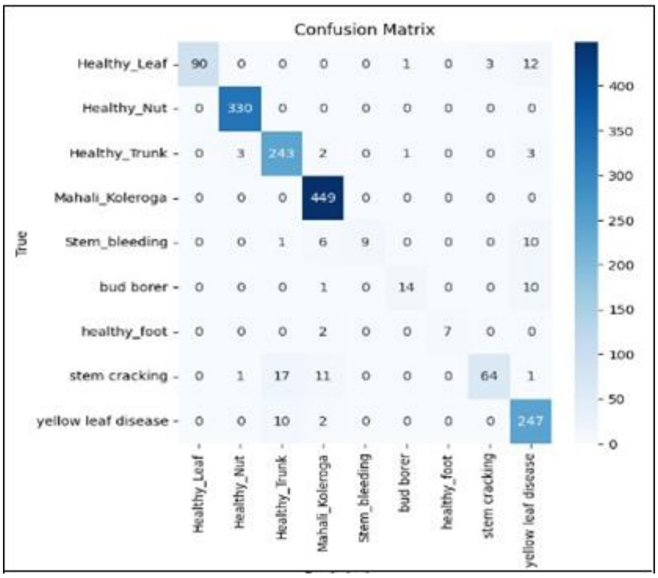
Table 2: Presents class-wise precision, recall, F1-score, and support

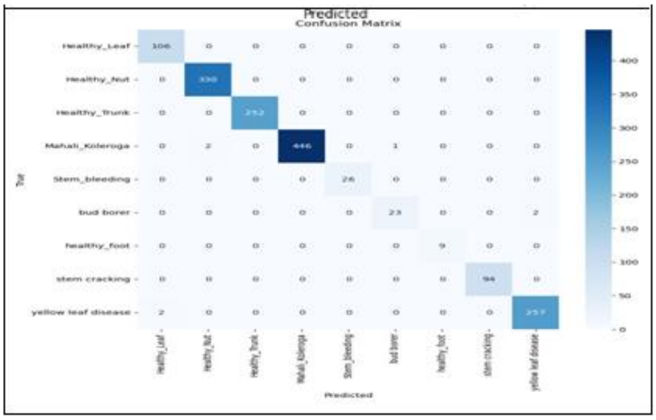
Methods	Classes	Precision rate	Recall rate	F1-score	Support
CNN	Healthy_Leaf	1.00	0.85	0.92	106
	Healthy_Nut	0.99	1.00	0.99	330
	Healthy_Trunk	0.90	0.96	0.93	252
	Mahali_Koleroga	0.95	1.00	0.97	449
	Stem_bleeding	1.00	0.35	0.51	26
	Bud borer	0.88	0.56	0.68	25
	Healthy_foot	1.00	0.78	0.88	9
	Stem cracking	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

Table 3: Summarizes training/validation/testing accuracy and loss

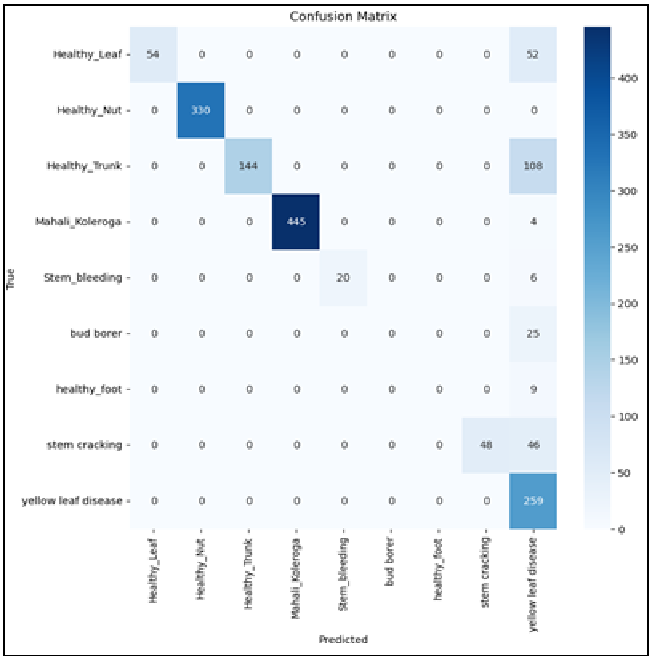
Models	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss	Test Accuracy	Test Loss
CNN	0.9786	0.0707	0.9309	0.2278	0.9374	0.2201
MobileV3Small	0.9155	0.3313	0.7913	0.8635	0.8387	0.4303
MobileV3Small-CM FusionNet	0.9861	0.0664	0.9865	0.1073	0.9954	0.0225

Confusion matrices are shown in Figures 2–4.





Paper ID: SR251124133511 DOI: https://dx.doi.org/10.21275/SR251124133511



4.2 Analysis of CNN Performance

The CNN achieves 93.74% test accuracy but struggles with minority classes, particularly:
Stem bleeding (recall = 0.35)
Bud borer (recall = 0.56)
This indicates high false negatives for rare samples.

4.3 MobileNetV3Small Limitations

MobileNetV3Small shows reduced performance (83.87% accuracy). It fails completely on:
Bud borer
Healthy Foot

This reveals inability of image-only features to capture fine-grained color patterns.

4.4 FusionNet: Enhanced Hybrid Learning

MobileNetV3Small-CM FusionNet dramatically improves all metrics, achieving:
99.54% testing accuracy
Perfect or near-perfect performance across all classes
Significant improvement on rare categories (e.g., Healthy Foot, Bud Borer)
This demonstrates the advantage of augmenting CNN features with color statistics.

4.5 Benchmark Comparison

Table 4: Comparison of Classification Methods and Accuracy

S. 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%

Table 4 compares the proposed method with previous literature. FusionNet outperforms all prior models, achieving the highest reported accuracy (99.54%).

5. Conclusion

This research presents MobileNetV3Small-CM FusionNet, a hybrid lightweight framework that successfully integrates deep convolutional features with handcrafted color statistics for enhanced multi-class arecanut disease classification. The fusion strategy significantly improves recognition of minority and visually subtle disease categories, resolving limitations observed in existing image-only models. Experimental evaluation demonstrates that the proposed method surpasses baseline CNN and MobileNetV3Small models, achieving a near-perfect classification performance with 99.54% accuracy. The architecture's lightweight design and superior generalization make it suitable for real-time deployment in mobile-based agricultural advisory systems.

References

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[3] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in *Proc. ECOC'00*, 2000, paper 11.3.4, p. 109.

[4] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.

[5] (2002) The IEEE website. [Online]. Available: http://www.ieee.org/

[6] 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

[7] 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

[8] Puneeth, B. R., & Nethravathi, P. S. (2021). A literature review of the
[9] 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.
[10] 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.
[11] Zhang, D. (2021). Color feature extraction. In Fundamentals of Image Data Mining: Analysis, Features, Classification and Retrieval (pp. 59-99). Cham: Springer International Publishing.
[12] *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.
[13] “PDCA12-70 data sheet,” Opto Speed SA, Mezzovico, Switzerland.
[14] J. Padhye, V. Firoiu, and D. Towsley, “A stochastic model of TCP Reno congestion avoidance and control,” Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.
[15] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.