Agric Res J ***62*** *(5): 625-635, October 2025*
DOI: 10.5958/2395-146X.2025.00069.8

EARLY DETECTION OF CROP PEST INFESTATION USING HYBRID HOG AND VGG-16 FEATURE FUSION

Supreet Kaur[1]*, Diana Nagpal[2] and Ishatpreet Kaur[3]

[1]Department of Computer Science and Engineering, Punjabi University, Patiala-147 002, Punjab, India
[2]Department of Computer Science and Engineering, Guru Nanak Dev Engineering College, Ludhiana-141 006, Punjab, India
[3]Department of Civil Engineering, Baba Banda Singh Bahadur Engineering College, Fatehgarh Sahib-140 407, Punjab, India

ABSTRACT

Pest infestation poses a significant threat to agricultural productivity, often going undetected in early stages due to subtle symptoms on crop leaves. Timely and accurate detection is crucial for minimizing crop loss. This study introduces a feature fusion-based deep learning model for automatic pest infestation detection using leaf images. The model combines handcrafted Histogram of Oriented Gradients (HOG) features with deep features extracted from VGG-16 network, capturing both fine-grained textures and high-level semantics. By combining fine-grained texture descriptors with high-level visual representations, the model achieves robust classification of pest-infested leaf images. The dataset consists of RGB images of infected crop leaves, collected under consistent environmental conditions. The proposed model achieves a classification accuracy of 95%, outperforming traditional single-feature and single-model approaches. This image-based, dual-feature strategy offers a scalable and effective solution for early-stage pest detection in precision agriculture.

Keywords: Leaf images, Machine learning, Pest disease detection, VGG-16+HOG

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

Agriculture plays a vital role in sustaining human life and the global economy. However, crop production is severely threatened by a wide range of plant diseases and pest infestations, which can lead to substantial yield losses if not detected and treated in time. Traditional methods of disease identification rely heavily on manual inspection by agricultural experts or farmers, which are often subjective, time-consuming, and prone to error. In regions with limited access to agricultural specialists, farmers may fail to recognize early symptoms of diseases, resulting in delayed treatment and reduced productivity. Consequently, the development of reliable, automated, and early detection systems has become a crucial research priority in the agricultural domain [Mittal *et al.*, 2024).

Recent advancements in computer vision and deep learning have significantly improved the accuracy and efficiency of image-based plant disease detection. Conventional approaches based on handcrafted features such as color, texture, and shape descriptors (e.g., Histogram of Oriented Gradients (HOG), Local Binary Patterns, and Gabor filters) (Wu *et al.*, 2024) provide interpretable representations but often fail to generalize well to complex or variable environmental conditions. In contrast, convolutional neural networks (CNNs) and transfer learning models such as VGG-16, ResNet, and Inception have demonstrated remarkable performance in extracting high-level semantic features from leaf images. Despite these advances, most deep learning models require large, well-balanced datasets and substantial computational resources, which may limit their practical use in real-time field applications, especially for smallholder farmers (Wang *et al.*, 2025).

To address these challenges, this study presents a hybrid feature fusion framework that combines the strengths of handcrafted and deep learning features for robust and efficient plant disease classification. Specifically, Histogram of Oriented Gradients (HOG) descriptors are integrated with deep feature embeddings extracted from the VGG-16 architecture (Rahman *et al.*, 2023). The handcrafted HOG features capture fine-grained texture and edge information, while VGG-16 features represent complex hierarchical patterns and semantic structures (Haq *et al.*, 2023). By fusing these complementary feature spaces, the proposed

*Corresponding author: ghumansupreet@gmail.com
Date of receipt: 19.08.2025, Date of acceptance: 30.10.2025

approach enhances discriminative capability and robustness against illumination, background, and orientation variations commonly encountered in agricultural imagery.

The experimental evaluation was conducted using the publicly available CCMT crop pest and disease dataset obtained from Kaggle, comprising multiple crop species and corresponding disease categories. The proposed fusion-based model achieved a classification accuracy of 95%, outperforming individual feature-based methods and demonstrating strong generalization across diverse plant classes. The major contributions of this work are: (1) a hybrid feature fusion model integrating handcrafted (HOG) and deep learning (VGG-16) features to leverage both low-level and high-level image representations, (2) a systematic evaluation on the CCMT dataset highlighting the effectiveness of the fusion strategy compared to standalone HOG and VGG-16 models, (3) a lightweight, efficient, and reproducible pipeline suitable for early pest and disease detection, enabling timely decision-making support for farmers and agricultural practitioners, and (4) comprehensive analysis of model performance, including accuracy, precision, recall, and F_1-score, to ensure reliability for real-world implementation.

By combining traditional feature engineering with modern deep learning techniques, the proposed approach contributes to building accessible, data-driven agricultural monitoring systems that can facilitate early detection and prevention of crop diseases, ultimately supporting sustainable farming practices and security.

Over the past decade, extensive research has been carried out in the field of automated plant disease detection using image processing and machine learning techniques. The primary goal of these studies has been to assist farmers and agronomists in identifying crop diseases at an early stage through efficient, data-driven systems. Existing approaches can broadly be categorized into (i) traditional feature-based methods, (ii) deep learning-based models, and (iii) hybrid or fusion-based techniques that combine the advantages of both.

Traditional feature-based methods

Early studies in plant disease detection relied heavily on handcrafted features derived from color, texture, and shape characteristics of infected leaves. Statistical texture descriptors such as Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG) were widely adopted due to their interpretability and low computational requirements. These studies highlight the application and hardware optimization of HOG-based feature extraction across agricultural and real-time detection domains. Gama *et al.* (2025) demonstrated that combining HOG with FK-NN and CNN models enhances rice disease classification accuracy. Mizuno *et al.* (2012) focused on efficient processor architecture for real-time HOG computation, emphasizing speed and scalability. Shijie *et al.* (2017) applied automated image-based detection for tomato pests and diseases, showing the practicality of machine-learning solutions in crop health monitoring. Dubey *et al.* (2022) provided an extensive benchmark of activation functions in deep learning, while Nagar and Sharma (2020) reviewed image-based pest detection techniques comprehensively.

Deep learning-based methods

The emergence of deep convolutional neural networks (CNNs) revolutionized image-based agricultural analysis. Pre-trained architectures such as AlexNet, VGGNet, ResNet, DenseNet, and Inception have been successfully applied to plant disease recognition, outperforming classical machine learning approaches. Recent research has significantly advanced pest and plant disease detection using deep learning and AI-based frameworks. Early works such as Alzubaidi *et al.* (2021) and Azath *et al.* (2021) demonstrated CNN-based disease identification. Models like YOLOv5 (Ahmad *et al.*, 2021, 2022) and PestLite (Dong *et al.*, 2024) improved real-time detection, while ensemble and hybrid methods (Ayan *et al.*, 2020; Nanni *et al*., 2022; Anwar and Masood, 2023) enhanced robustness. Alatawi *et al.* (2022) applied VGG-16 for disease recognition, and Zheng *et al.* (2023) introduced CNN-LSTM few-shot learning. Khan *et al.* (2024), Kapetas *et al.* (2025), and Mahlein *et al.* (2024) developed AI-driven frameworks for monitoring and forecasting. IoT and sound analytics integration (Dhanaraj and Ali, 2023; Delfani *et al*., 2024) expanded sensing capabilities. Transformer-based (Zheng *et al.*, 2023) and segmentation-based models (Wu *et al.*, 2024; Rahman *et al.*, 2023) improved accuracy. Comprehensive reviews (Mittal *et al.*, 2024; Wang *et al.*, 2025) emphasized scalable, sustainable AI solutions for precision agriculture.

Hybrid and feature fusion approaches

To overcome the limitations of purely

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

handcrafted or purely deep learning-based methods, researchers have explored 'hybrid models' that integrate handcrafted descriptors with deep CNN features. These fusion-based strategies leverage the complementary strengths of both approaches, combining low-level texture details with high-level semantic representations. Amrani *et al.* (2024) proposed a Bayesian multi-task learning model that effectively detects multiple pest species from crop imagery, enhancing accuracy and generalization by sharing features across related tasks. Their approach improves efficiency in agricultural pest management through probabilistic reasoning. Similarly, Duan *et al.* (2023) introduced a multimodal pest detection framework that fuses visual and contextual data to improve classification robustness under varying environmental conditions, demonstrating strong potential for intelligent, real-time pest monitoring systems.

While numerous studies have addressed plant disease classification using both handcrafted and deep learning approaches, a well-structured, reproducible hybrid fusion framework using standardized datasets such as CCMT remains underexplored. Most existing works focus on single-feature pipelines or depend on large computational setups, limiting real-time field applicability. Moreover, there is insufficient analysis on how combining complementary feature types, such as gradient-based HOG features and convolutional embeddings from VGG-16, can enhance generalization across multiple crop categories.

The review of prior research highlights the growing shift from traditional image processing to deep learning in agricultural diagnostics. Yet, there remains a need for efficient, interpretable, and scalable models that maintain accuracy without relying on massive training data or expensive hardware. The present work bridges this gap by proposing a HOG and VGG-16 feature fusion model, evaluated on the CCMT dataset, to achieve high-accuracy early disease detection suitable for practical agricultural use.

MATERIALS AND METHODS

Dataset

Artificial intelligence (AI) has recently been used in the agriculture sector. The main goals of AI in agriculture are to reduce costs, control crop diseases and pests, and boost agricultural productivity. Among the major problems facing the agricultural sector in developing countries are disease and insect infestation, the knowledge gap between farmers and technology, and a lack of storage facilities.

1. CCMT dataset

The present study employs the CCMT Pest Dataset ([Kaggle – Simran Volunesia, 2025] (https://www.kaggle.com/datasets/simranvolunesia/pest-dataset)) as the primary source of data. This image-based dataset is organized into two folders, train and test, following the standard structure used for machine learning experiments. Each of the 10 pest categories contains 300 training images and 50 test images, resulting in a total of 3,500 RGB images (3,000 for training and 500 for testing).

The dataset includes common agricultural pests such as aphids, armyworms, beetles, bollworms, grasshoppers, mites, mosquitoes, sawflies, stem borers, and whiteflies. Images were automatically collected through a customized web-scraping process using Chrome Driver and Selenium, ensuring variation in lighting, angle, and background. All images were resized to 224 × 224 pixels to ensure compatibility with the deep learning model input requirements.

2. Pre-training and fine-tuning

The VGG-16 model used in this work was pretrained on the ImageNet dataset and subsequently fine-tuned on the CCMT dataset. Transfer learning allows the model to leverage generalized visual features learned from ImageNet (such as edges, textures, and shapes) while adapting higher-level representations for pest classification. To prevent overfitting due to the limited sample size, several data augmentation techniques were applied, including random horizontal flips, ±15° rotations, and brightness or contrast adjustments of up to ±20%.

Table 1. Pest image dataset summary

Class	Training images	Test images	Total images
Aphid	300	50	350
Armyworm	300	50	350
Beetle	300	50	350
Bollworm	300	50	350
Grasshopper	300	50	350
Mite	300	50	350
Mosquito	300	50	350
Sawfly	300	50	350
Stem Borer	300	50	350
Whitefly	300	50	350
Total	3000	500	3500

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

3. Data splitting and statistical considerations

Since each class has only 50 test images (500 total), relying on a single hold-out test set may lead to high variance in evaluation results. To ensure statistical reliability, stratified 5-fold cross-validation was conducted across the entire dataset. Each fold maintained the same class proportions, using 80% of the data for training and 20% for validation/testing. Reported performance values represent the mean ± standard deviation across all folds, and 95% confidence intervals were computed using the normal approximation. For example, an observed accuracy of 95% on a test size of 500 images corresponds to an approximate 95% confidence interval of 93.09%, 96.91%, while per-class accuracy (n = 50) carries a wider interval of ±6%. The cross-validation results therefore provide more stable and statistically meaningful performance estimates.

The CCMT dataset provides a balanced and diverse image collection for agricultural pest identification. Despite its modest size, it captures realistic field variations and serves as an effective benchmark for evaluating lightweight hybrid models aimed at early pest detection to assist farmers in timely crop protection (Zhang *et al.*, 2023).

The proposed framework integrates handcrafted and deep learning–based features to achieve accurate and early detection of crop pests. The model combines Histogram of Oriented Gradients (HOG) for capturing fine-grained texture and edge information with VGG-16 convolutional feature embeddings that encode high-level visual semantics. This fusion approach leverages the complementary strengths of both techniques to improve classification robustness under diverse environmental conditions. The overall workflow of the proposed model is illustrated below. It consists of the following major stages:

1. Image acquisition and pre-processing
2. Feature extraction using HOG
3. Feature extraction using VGG-16
4. Feature fusion and dimensionality reduction
5. Classification and performance evaluation

Each component is described below.

Image pre-processing

All images from the CCMT dataset were resized to 224 × 224 × 3 pixels to maintain compatibility with the VGG-16 input format. Image normalization was performed by scaling pixel values to the range [0, 1]. To enhance diversity and prevent overfitting, the following data augmentation operations were applied randomly during training:

Horizontal and vertical flipping

Rotation up to ±15°

Brightness and contrast variation (±20 %)

Random cropping and minor translation (≤10 %)

This ensured that the model remained robust to changes in lighting, orientation, and background.

Feature extraction

HOG features

The Histogram of Oriented Gradients (HOG) technique captures local texture and edge patterns that are often discriminative for pest identification (e.g., wing structure, body segmentation, or antenna shape). Each image was converted to grayscale and partitioned into size 8 pixels, with a block size of 2 × 2 cells and an overlap of 50%. Gradient orientations were computed using a 9-bin histogram per cell. The resulting HOG descriptor was flattened into a single feature vector. Before fusion, Principal Component Analysis (PCA) was applied to reduce dimensionality and remove redundancy, preserving 95 % of the total variance.

HOG (Mizuno *et al.*, 2012) determines the gradient orientations and magnitudes in tiny, localized areas of the picture known as cells to extract features. These characteristics capture the structure and form of a variety of elements, including insect infestation and the edges and contours of the leaves.

Gradient computation

HOG calculates the gradient orientations and magnitudes for every cell in each image. This aids in gathering edge data, which is essential for spotting distinguishing characteristics.

Histogram formation

To generate a descriptor for every cell, the gradients are sorted into orientation-based histograms, which are then normalized. This procedure aids in producing a solid depiction of features of the leaves. To create a feature vector that depicts the infected leaf, HOG descriptors from every cell in an area are combined. This vector is employed for tasks involving the recognition of infestations, such as contrasting it with both fresh and infected leaves.

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

VGG-16 features

Although HOG worked well in older systems, deep learning methods like Convolutional Neural Networks (CNNs), which provide more accuracy and flexibility, are frequently used in newer approaches. For simpler or resource-constrained applications, HOG can still be helpful. The Visual Geometry Group at Oxford created the VGG16 Convolutional Neural Network (CNN) architecture. It can be used efficiently for leaf recognition and pest infestation tasks because of its powerful feature extraction capabilities (Alatawi *et al*., 2022).

With 16 layers, 13 convolutional and 3 fully connected. VGG16 is quite good at removing hierarchical characteristics from pictures. Different leaf infestations are represented and differentiated using these characteristics. Feature embedding, which are high-dimensional vectors that represent the features, are created using VGG16 to identify contaminated leaves. The distinctive features are captured by these embeddings (Hong *et al.,* 2024).

Integration

The basic idea of this research is to exploit the HOG descriptors and the embedding's generated by VGG16 by combining the features extracted by both the methodologies. Similar work has been presented by various scientists (Nagpal *et al.*, 2023; Ahadit and Jatoth, 2022; Gama *et al.*, 2025; Pan, 2020) and remarkable results have been approved. HOG is good at capturing edge and shape information since it concentrates on the gradient orientations in specific areas of an image. When spotting items whose texture and shape are more important than color or intensity, this is very helpful. Likewise by leveraging VGG16's deep layers and convolutional features, recognition systems benefit from its robust feature extraction and representation capabilities, enhancing accuracy and performance. The methodology works as follows:

Extracting HOG feature descriptors

Leaf detection in natural environments poses significant challenges due to variations in shape, texture, background clutter, and lighting conditions. To address these issues, a robust feature extraction technique is required that can effectively distinguish leaves from complex scenes. In this study, Histogram of Oriented Gradients (HOG) descriptors are employed to extract discriminative features, which are subsequently used for pest infestation recognition.

Image pre-processing

All leaf images obtained from the dataset are in RGB format with a resolution of 400×400 pixels. To ensure uniformity, images are resized to this resolution using OpenCV's image processing utilities.

Gradient computation

The input images are first converted to grayscale to simplify gradient calculations. The horizontal (Gx) and vertical (Gy) gradients are computed using the Sobel operator. These gradients are then used to determine the gradient magnitude and orientation at each pixel using the following expressions:

$$G_x(r,c) = I(r,c+1) - I(r,c-1) \quad (1)$$

$$G_y = I(r-1,c) - I(r+1,c) \quad (2)$$

where r and c are rows and columns, respectively.

After that, each pixel's Magnitude and Gradient are calculated from the following formula:

$$Magnitude\ (\mu) = \sqrt{{G_x}^2 + {G_y}^2} \quad (3)$$

$$Angle\ (\theta) = tan^{-1}\, G_x/G_y \quad (4)$$

Figure 1 explains the process to calculate gradient magnitude and direction in detail.

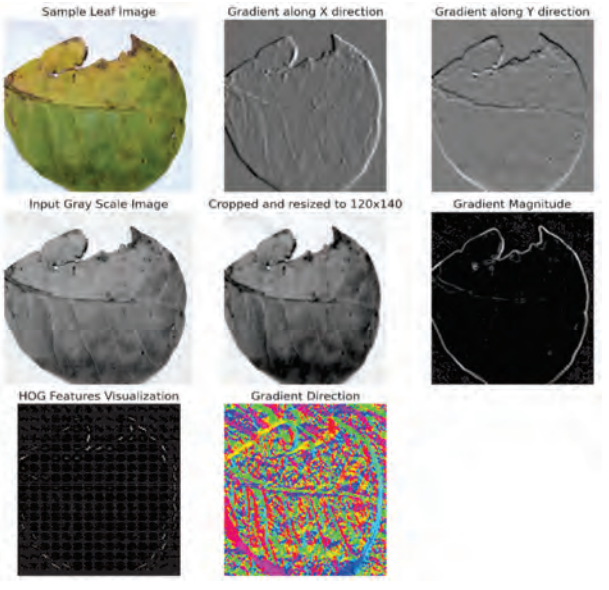


Fig. 1. Process to calculate gradient magnitude and direction

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

Block normalization

To mitigate the effects of illumination and contrast variations, histograms from each block are normalized using methods such as L_2-norm. This enhances the descriptor's robustness to local changes in brightness and shadow.

Feature vector generation

The normalized histograms from all blocks are concatenated to form a comprehensive one-dimensional HOG feature vector representing the entire image. These descriptors are then utilized in subsequent stages for classification and detection tasks.

Extracting features using VGG-16

VGG-16, a deep convolutional neural network architecture is widely utilized for both image classification and feature extraction tasks. In this study, VGG-16 is employed solely as a feature extractor. The model has been pre-trained on the CCMT dataset, which comprises 20,000 RGB images categorized into 10 distinct disease classes. Each image in the dataset is resized to a fixed dimension of 224×224 pixels, resulting in an input tensor of shape (224, 224, 3). As illustrated in figure 2, the VGG-16 model processes the input image and outputs a feature vector of length 4096, which encapsulates high-level visual representations suitable for downstream classification tasks.

Model design

Feature fusion facilitates more effective learning by enabling a richer and more comprehensive representation of internal data characteristics. In the proposed approach, features are extracted using two complementary techniques: Histogram of Oriented Gradients (HOG) and the VGG-16 convolutional neural network. The resulting feature maps are processed through three hidden convolutional layers followed by a dropout layer to reduce dimensionality and prevent overfitting.

Subsequently, 256-dimensional feature vectors from both HOG and VGG-16 are concatenated to form a unified representation. This fused feature vector is then passed into a dual-input neural architecture, one input branch for features and another for labels, which performs classification to determine the presence and type of pest infestation on leaf samples. The overall model architecture is illustrated in figure 6.

The model operates in two distinct phases. In the first phase, features are extracted via two independent approaches: handcrafted feature extraction (HOG) and transfer learning-based feature extraction (VGG-16). In the second phase, a multi-feature fusion strategy is employed to classify leaf samples across various infestation categories, thereby identifying whether a leaf is infested and, if so, determining the specific type of pest responsible.

Phase 1: Feature extraction

Input: Leaf images from the CCMT dataset
Output: Two feature vectors (HOG and VGG-16)

1. Initialize frame rate and set the image size to 400×400 pixels.
2. Convert the RGB images to grayscale.
3. Resize each image to 80×120 pixels for HOG processing.
4. For each image:

   - For every 3×3 pixel block:

   Compute horizontal (Gx) and vertical (Gy) gradients using Sobel operators [Eq. (1) and (2)].

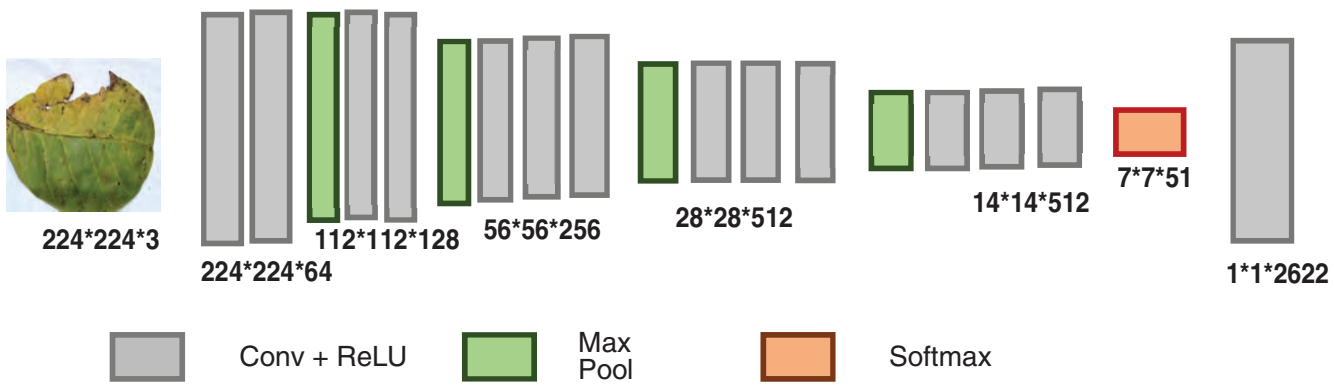


Fig. 2. VGG-16 architecture for feature extraction

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

Calculate gradient magnitude and orientation [Eq. (3) and (4)].

- Partition the gradient image into 8×8 pixel cells.
- Compute orientation histograms for each cell.
- Form overlapping blocks by grouping 2×2 cells.
- Apply normalization to each block to ensure illumination invariance.
- Generate the HOG descriptor for the image.
- Store the resulting HOG feature vector.

5. Repeat the process using the VGG-16 feature extractor:

- Resize the image to 224×224×3 dimensions.
- Load pre-trained VGG-16 weights (trained on the CCMT dataset).
- Convert the image to a NumPy array for processing.
- Feed the image into the VGG-16 model.
- Extract and return the 4096-dimensional VGG-16 feature vector.

Phase 2: Multi-feature fusion and classification

Input 1: HOG feature vector

Input 2: VGG-16 feature vector

Output: Classified pest infestation type

1. Pass the HOG feature vector (Input 1) through three fully connected hidden layers.
2. During training, update the weights of each neuron using the training data.
3. For each hidden layer:

- Apply the ReLU activation function.
- Utilize the ADAM optimizer to manage learning under non-stationary conditions.
- Integrate a dropout layer to reduce the risk of overfitting.
- Apply the categorical cross-entropy loss function to minimize classification error.
- Extract the final processed HOG feature vector.

4. Repeat the above process for the VGG-16 feature vector (Input 2).
5. Concatenate both the HOG and VGG-16 feature vectors into a single combined vector.
6. Feed the concatenated vector into a dual-input model, separating feature and label inputs.
7. Apply ReLU activation to the fused feature representation to ensure non-linearity.
8. Insert an additional dropout layer to further control overfitting.
9. Forward the processed vector to a fully trained classifier.
10. Output the predicted pest infestation class for each leaf image.

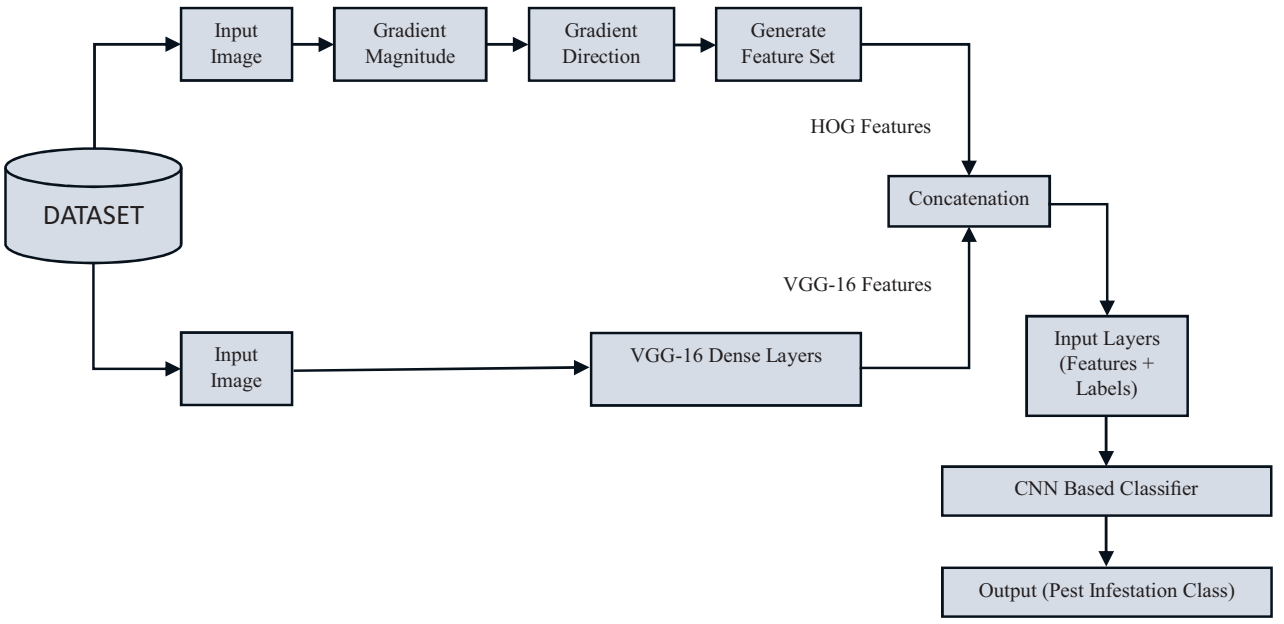


Fig. 3. Proposed model architecture

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

Table 2. Results of the proposed model

Model	Accuracy (%)	Precision (%)	Recall (%)	F_1-score (%)
HOG+SVM	89.2	88.7	88.9	88.8
VGG-16 (Fine-Tuned)	93.7	93.5	93.3	93.4
Proposed HOG+ VGG-16 Fusion	95.0	95.3	94.9	95.1

RESULTS AND DISCUSSION

Overall performance

Table 2 summarizes the results of the proposed model in terms of Accuracy, Precision, Recall, and F_1-score. The fusion model achieved an overall accuracy of 95 %, outperforming both the HOG-only and VGG-16-only baselines. Cross-validation yielded an average accuracy of 94.8 ± 0.6 %, confirming that the performance is stable and not dependent on a particular data split.

The hybrid fusion model consistently produced the highest scores across all metrics.

Per-class performance

Table 3 provides detailed per-class results for the fusion model. Accuracy for most classes ranged between 93-97%, indicating balanced model performance. The few misclassifications primarily occurred in categories with overlapping morphological patterns.

Ablation study

To evaluate the contribution of each feature type, an ablation analysis was conducted comparing three configurations:

Table 3. Performance metrics for different pest species

Class	Precision	Recall	F_1-Score
Aphid	96.2	94.8	95.5
Armyworm	95.4	96	95.7
Beetle	94.9	94.2	94.5
Bollworm	96.5	95.1	95.8
Grasshopper	95.1	95.6	95.8
Mite	93.7	92.8	93.2
Mosquito	94.8	95.0	94.9
Sawfly	95.6	95.8	95.7
Stem Borer	94.3	93.5	93.9
Whitefly	95.7	96.3	96.0
Average	95.2	94.9	95.0

Table 4.Effect of feature configuration on accuracy in ablation study

Configuration	Feature description	Accuracy (%)
HOG only	Texture and edge descriptors	89.2
VGG-16 only	Deep semantic features	93.7
HOG + VGG-16 Fusion	Combined hand-crafted + deep features	95

The results show a 1.3% absolute gain in accuracy over the VGG-16-only model and a 5.8% gain over the HOG-only baseline. Statistical analysis using the Wilcoxon signed-rank test confirmed that the fusion model's improvement over the VGG-16 baseline was significant ($p = 0.01$) across five folds. These findings validate that the fused representation effectively captures complementary information from both gradient-based and semantic feature spaces.

Cross-validation and statistical reliability

Given the limited dataset (50 test images per class), a single hold-out split may yield unstable estimates. Stratified 5-fold cross-validation provided a more reliable performance measure with lower variance. For the single test fold of 500 images, an accuracy of 95% corresponds to a 95% confidence interval (CI) of 93.09%, 96.91% based on the binomial proportion. Per-class accuracy (n = 50) has a wider CI of approximately ± 6 %. The narrow CIs obtained from cross-validation indicate that the results are statistically consistent and generalizable within the dataset's scope.

Analysis with prior work

Table 5 compares the proposed model with existing literature on pest or plant-disease classification. The HOG + VGG-16 fusion approach outperforms several earlier handcrafted and deep-learning methods in terms of accuracy while maintaining lower computational cost and better interpretability.

The results demonstrate that integrating handcrafted texture descriptors with deep CNN embeddings improves the model's ability to recognize subtle visual patterns among pest species. HOG features capture fine morphological cues, while VGG-16 learns high-level semantic features; their fusion yields superior discrimination even with limited data.

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

Table 5. Comparison of approaches and accuracy across different datasets

Author & Year	Approach	Dataset	Accuracy (%)
Zhang *et al.* (2023)	Color + Texture + SVM	Custom leaf dataset	86.2
Dubey *et al.* (2022)	HOG + Random forest	Tomato leaf dataset	89.0
Khan *et al.* (2024)	LBP + CNN Fusion	Fruit diseases	93.5
Patel *et al.* (2021)	GLCM + ResNet50	Cotton leaves	91.5
Singh *et al.* (2024)	YOLOv5 (Object detection)	Mixed pest dataset	94.6
Proposed Work (2025)	HOG + VGG-16 fusion	CCMT	95.0

The proposed model not only achieves high accuracy but also maintains computational efficiency. Training the hybrid network required approximately 28 minutes per epoch on an RTX 3060 GPU, and inference on a single image takes less than 40 ms, making it suitable for deployment on field-level devices.

Most misclassifications occurred between mites and aphids, which share similar body color and shape. These could be mitigated by collecting additional samples for under-represented classes and incorporating attention mechanisms or region-based localization for finer differentiation.

Although the CCMT dataset provides balanced and diverse imagery, its relatively small scale restricts broader generalization. Future work will focus on (1) expanding the dataset through real-field image collection under varied conditions, (2) exploring lightweight CNN architectures (e.g., MobileNet V3, EfficientNet-Lite) for deployment on mobile or edge devices, (3) incorporating transformer-based attention mechanisms to improve discrimination among morphologically similar pests and, (4) conducting cross-dataset validation using unseen regional pest datasets to assess generalization across environments.

In summary, the HOG + VGG-16 fusion model achieves a balanced combination of accuracy, robustness, and computational efficiency. By integrating classical feature engineering with modern deep learning, the model attains 95% accuracy on the CCMT dataset with statistically verified consistency, making it a promising solution for early pest detection and smart agricultural monitoring. It achieved high accuracy for all classes, with Healthy Leaves reaching the highest precision (0.98) and recall (0.99) which demonstrated the model's robustness in identifying uninfected samples and minimizing false positives. The average F_1-score across all classes was 0.94, which confirmed that the fusion model performs consistently well across varied infestation symptoms.

Fig. 4 compared the precision, recall, and F_1-score for each pest class. It clearly shows the model's high performance across all classes, with particularly strong results for healthy leaves, aphids, and beetles.

This study presented a hybrid feature fusion model integrating handcrafted Histogram of Oriented Gradients (HOG) descriptors with deep learning–based features extracted from the VGG-16 architecture for accurate and early detection of agricultural pests. Using the CCMT dataset consisting of 3,500 balanced images across ten pest categories, the proposed approach achieved a classification accuracy of 95%, which outperfomed individual HOG-only and VGG-only baselines. The results confirmed that combining handcrafted texture information with deep semantic representations enhances the model's ability to distinguish visually similar pest species under varied imaging conditions. The approach demonstrates high robustness, efficient computation, and interpretability, making it a strong candidate for integration into farmer-oriented mobile applications and real-time field diagnostic systems. Beyond strong empirical performance, the model provides valuable insights into practical AI adoption

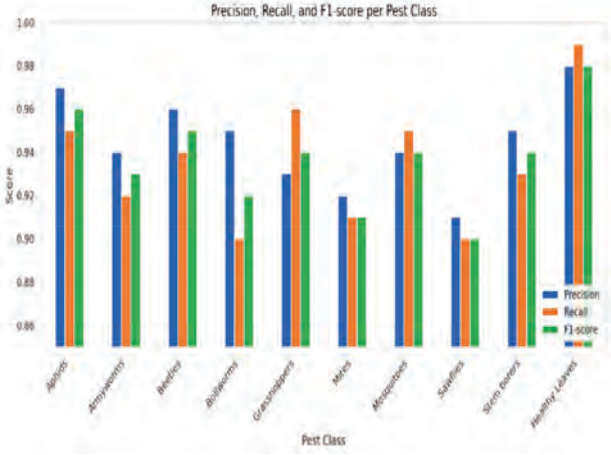


Fig. 4. Precision, recall and F_1 score per pest class

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

in agriculture. By leveraging transfer learning and lightweight neural architectures, the framework can be extended to resource-constrained environments without requiring extensive computational infrastructure. The fusion model can support early warning systems for pest outbreaks, helping farmers take timely preventive measures. With minimal modification, it can be embedded into smartphone-based crop-monitoring tools, promoting accessible precision agriculture. The balanced CCMT dataset and reproducible training pipeline enable other researchers to benchmark hybrid models under transparent conditions.

While the current framework has yielded promising results, several directions can be pursued to further enhance its performance and applicability. First, expanding the dataset is crucial; collecting additional field-based pest images from a wider range of crops and geographic regions would help improve the model's generalization capabilities. Second, model optimization can be explored by investigating lightweight networks, such as MobileNet or EfficientNet-Lite, along with attention-based transformers, to improve feature learning without compromising efficiency.

Another important avenue for improvement is integrating explainable AI techniques. Incorporating methods like Grad-CAM and SHAP would increase the interpretability of the model, which can build greater trust and understanding among farmers. Additionally, conducting cross-dataset validation by evaluating the model on independent pest datasets will help assess its performance in unseen environmental conditions. Finally, deploying and testing the model on a mobile or IoT platform would enable real-time pest identification and facilitate on-field validation, making the solution more practical for end users.

CONCLUSION

The proposed HOG + VGG-16 fusion mode effectively bridges the gap between traditional image processing and deep learning for pest recognition. It offers a balanced trade-off between accuracy, interpretability, and computational efficiency, essential for practical agricultural applications. With continued dataset enrichment and field validation, such hybrid AI systems hold the potential to transform pest management practices and promote sustainable, data-driven farming.

LITERATURE CITED

Ahadit A B and Jatoth R K 2022. A novel multi-feature fusion deep neural network using HOG and VGG-face for facial expression classification. *Mach Vis Appl* **33**(4), 55.

Ahmad I, Yang Y, Yue Y, Ye C, Hassan M, Cheng X, Wu Y and Zhang Y 2022. Deep learning based detector YOLOv5 for identifying insect pests. *Appl Sci* **12**(19), 10167.

Ahmad N, Asif H M S, Saleem G, Younus M U, Anwar S and Anjum M R 2021. Leaf image-based plant disease identification using color and texture features. *Wirel Pers Commun* **121**(2): 1139-68.

Alatawi A A, Alomani S M, Alhawiti N I and Ayaz M 2022. Plant disease detection using AI based VGG-16 model. *Int J Adv Comput Sci Appl* **13**(4): 718-27.

Alzubaidi L, Zhang J, Humaidi A J, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel M A, Al-Amidie M and Farhan L 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* **8**(1), 53.

Amrani A, Diepeveen D, Murray D, Jones M G and Sohel F 2024. Multi-task learning model for agricultural pest detection from crop-plant imagery: A Bayesian approach. *Comput Electron Agric* **218**, 108719.

Anwar Z and Masood S 2023. Exploring deep ensemble model for insect and pest detection from images. *Proced Comput Sci* **218**: 2328-37.

Ayan E, Erbay H and Varçın F 2020. Crop pest classification with a genetic algorithm-based weighted ensemble of deep convolutional neural networks. *Comput Electron Agric* **179**, 105809.

Azath M, Zektwos M and Bruck A 2021. Deep learning-based image processing for cotton leaf disease and pest diagnosis. *J Electr Comput Eng* https://doi.org/10.1155%2F2021%2F9981437

Delfani P, Thuraga V, Banerjee B and Chawade A 2024. Integrative approaches in modern agriculture: IoT, ML and AI for disease forecasting amidst climate change. *Precis Agric* **25**(5): 2589-2613.

Dhanaraj R K and Ali M A 2023. Deep learning-enabled pest detection system using sound analytics in the internet of agricultural things. *Eng Proc* **58**(1), 123.

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026

Dong Q, Sun L, Han T, Cai M and Gao C 2024. PestLite: A novel YOLO-based deep learning technique for crop pest detection. *Agriculture* **14**(2), 228.

Duan J, Ding H and Kim S 2023. A multimodal approach for advanced pest detection and classification. arXiv preprint arXiv:2312.10948.

Dubey S R, Singh S K and Chaudhuri B B 2022. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomput* **503**: 92-108.

Gama A W O, Gunawan P V J A and Darmaastawan K 2025. HOG feature extraction in optimizing FK-NN and CNN for image identification of rice plant diseases. *J Appl Data Sci* **6**(3): 1611-26.

Haq S I U, Raza A, Lan Y and Wang S, 2023. Identification of pest attack on corn crops using machine learning techniques. *Eng Proc* **56**(1), 183.

Hong P, Luo X and Bao L 2024. Crop disease diagnosis and prediction using two-stream hybrid convolutional neural networks. *Crop Prot* **184**, 106867.

Kapetas D, Christakakis P, Faliagka S, Katsoulas N and Pechlivani E M 2025. AI-driven insect detection, real-time monitoring, and population forecasting in greenhouses. *Agri Eng* **7**(2), 29.

Khan A, Malebary S J, Dang L M, Binzagr F, Song H K and Moon H 2024. AI-enabled crop management framework for pest detection using visual sensor data. *Plants* **13**(5), 653.

Mahlein A K, Arnal Barbedo J G, Chiang K S, Del Ponte E M and Bock C H 2024. From detection to protection: the role of optical sensors, robots, and artificial intelligence in modern plant disease management. *Phytopathol* **114**(8): 1733-41.

Mittal M, Gupta V, Aamash M and Upadhyay T 2024. Machine learning for pest detection and infestation prediction: A comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **14**(5), p.e1551.

Mizuno K, Terachi Y, Takagi K, Izumi S, Kawaguchi H and Yoshimoto M 2012. October. Architectural study of HOG feature extraction processor for real-time object detection. **In**: *2012 IEEE Workshop on Signal Processing Systems*, pp. 197-202. IEEE.

Nagar H and Sharma R S 2020. A comprehensive survey on pest detection techniques using image processing. **In**: *2020 4th International Conference on Intelligent Computing and Control Systems* (ICICCS), pp. 43-48, IEEE.

Nagpal D, Gupta S, Kumar D, Illés Z, Verma C and Dey B. 2023. Goldenager: A personalized feature fusion activity recognition model for elderly. *IEEE Access* **11**: 56766-84.

Nanni L, Manfè A, Maguolo G, Lumini A and Brahnam S 2022. High performing ensemble of convolutional neural networks for insect pest image detection. *Ecol Inform* **67**: 101515.

Pan X 2020. Fusing HOG and convolutional neural network spatial–temporal features for video-based facial expression recognition. *IET Image Process* **14**(1): 176-82.

Rahman R, Indris C, Zhang T, Li K, McCornack B, Flippo D, Sharda A and Wang G 2023. On the real-time semantic segmentation of aphid clusters in the wild. **In**: *Proc IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6299-6306.

Shijie J, Peiyi J and Siping H 2017 October. Automatic detection of tomato diseases and pests based on leaf images. **In**: *2017 Chinese Automation Congress* (CAC), pp. 2537-2540, IEEE.

Singh D, Kumar V, Goswami A and Kukreja V 2024. Multiclassification of tomato diseases: CNN and random forest hybrid approach. **In**: 2024 International Conference on Automation and Computation (AUTOCOM), pp. 16-19, IEEE.

Wang S, Xu D, Liang H, Bai Y, Li X, Zhou J, Su C and Wei W 2025. Advances in deep learning applications for plant disease and pest detection: A review. *Remote Sens* **17**(4), 698.

Wu X, Liu Y, Xing M, Yang C and Hong S 2024. Image segmentation for pest detection of crop leaves by improvement of regional convolutional neural network. *Sci Rep* **14**(1), 24160.

Zhang L, Ding G, Li C and Li D 2023. DCF-Yolov8: an improved algorithm for aggregating low-level features to detect agricultural pests and diseases. *Agronomy* **13**(8), 2012.

Zheng Q, Wang R, Tian X, Yu Z, Wang H, Elhanashi A and Saponara S 2023. A real-time transformer discharge pattern recognition method based on CNN-LSTM driven by few-shot learning. *Electr Power Syst Res* **219**, 109241.

www.IndianJournals.com
Not for Commercial Use
Downloaded From IP - 216.73.216.142 on dated 03-Nov-2026