Pest Monitoring in Greenhouses Using Vision Technology

Lovely Professional University

Phagwara, Punjab

Abstract

The increasing demand for sustainable and efficient agricultural practices has led to the integration of automation and smart technologies in crop monitoring and protection. One of the critical challenges in greenhouse farming is the early detection and classification of pest infestations, which can significantly impact crop yield and quality if not addressed promptly. This research presents an intelligent pest monitoring system that leverages computer vision and deep learning techniques to identify and classify multiple types of pests commonly found in greenhouse environments. A Convolutional Neural Network (CNN) model is developed and trained on a curated dataset containing image samples of pests such as aphids, mites, armyworms, and beetles. The dataset is processed through advanced preprocessing techniques including image augmentation and normalization to enhance model generalization. The system demonstrates high accuracy in classification, making it suitable for real-time deployment in greenhouse conditions. By automating pest detection, this approach minimizes the reliance on manual inspections, reduces pesticide usage, and supports precision agriculture. The proposed model not only improves detection efficiency but also lays the groundwork for scalable smart farming solutions that integrate AI, vision systems, and IoT technologies.

Keywords

Pest detection, Deep learning, Greenhouse monitoring, Image classification, Vision system, Agriculture technology.

I. Introduction

Agriculture remains a vital sector for ensuring global food security, with greenhouse farming playing a significant role in enabling year-round production of fruits, vegetables, and ornamental plants. However, one of the major challenges faced by greenhouse cultivators is the timely identification and control of pest infestations. Pests such as aphids, mites, armyworms, and beetles can cause substantial damage to crops, reduce productivity, and increase dependency on chemical pesticides, which in turn raises environmental and health concerns.

Traditionally, pest monitoring in greenhouses has relied on manual inspection methods, which are labor-intensive, time-consuming, and often subjective. These methods are prone to human error and can fail to detect early-stage infestations, resulting in delayed intervention and significant crop losses. As greenhouses become more complex and cultivation becomes more intensive, there is a growing need for automated, scalable, and reliable pest monitoring systems.

Recent advancements in computer vision and deep learning have opened new avenues for automating visual inspection tasks in agriculture. Convolutional Neural Networks (CNNs), in particular, have demonstrated exceptional performance in image classification and object detection tasks. These models are capable of learning intricate patterns from visual data, making them ideal for identifying subtle differences between pest species under varying environmental conditions.

In this study, we present a vision-based pest monitoring system designed specifically for greenhouse environments. The system utilizes a CNN model trained on a labeled dataset of pest images captured under controlled conditions. The objective is to accurately classify different pest species from images and provide actionable insights for greenhouse management. By deploying such a system, growers can achieve early detection, reduce chemical usage, and enhance crop protection strategies.

II. Dataset and Preprocessing

A robust and well-curated dataset is the foundation for any successful machine learning model, especially in tasks involving image classification for agricultural applications. For this study, a custom image dataset was developed specifically for pest monitoring in greenhouse environments. The dataset consists of images of common pest species affecting greenhouse crops, such as aphids, armyworms, beetles, mites, bollworms, grasshoppers, mosquitoes, sawflies, and stem borers. These pests were selected based on their prevalence in controlled agricultural settings and their economic impact on crop yield and quality.

A. Data Collection

The dataset was structured into folders, each corresponding to a single pest category. Images were collected from real greenhouse conditions and supplemented with publicly available datasets and open-source agricultural image repositories. Care was taken to include images captured under varying lighting conditions, backgrounds, and pest orientations to simulate real-world scenarios more accurately.

B. Dataset Structure

The dataset was divided into two main directories: train and test, to support supervised learning and evaluation. The training set was used for model learning, while the test set was held out to assess generalization performance. The distribution of images was kept balanced across classes to avoid bias toward any specific pest type.

Each pest folder contained between 200–500 labeled images, and image dimensions varied depending on the source. File formats included .jpg and .png, and the images contained a mix of close-up views and wider shots of pests on plant surfaces.

C. Image Preprocessing

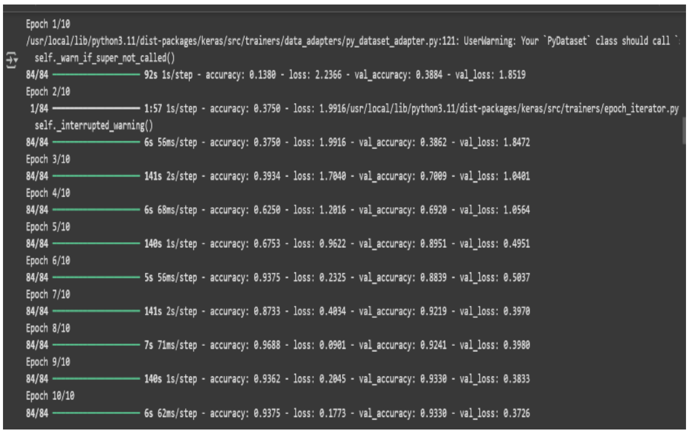
To ensure consistency and model readiness, the following preprocessing steps were applied:

- Resizing: All images were resized to a uniform resolution of 224x224 pixels, maintaining aspect ratio where possible, to fit the input shape expected by the CNN architecture.
- Normalization: Pixel values were scaled to a range between 0 and 1 by dividing by 255. This standardization helps accelerate model convergence and improve numerical stability during training.
- Label Encoding: Pest categories were assigned integer labels and then one-hot encoded for compatibility with the model's output layer in multi-class classification tasks.
- Data Augmentation: To increase dataset diversity and combat overfitting, real-time augmentation was applied during training using the following transformations:
    - Random rotations (up to ±20 degrees)
    - Horizontal and vertical flipping
    - Zooming (±10%)
    - Brightness adjustments
    - Width and height shifts

These augmentations helped simulate different capture conditions and enhanced the model’s ability to generalize to unseen data.

D. Class Imbalance Handling

Although efforts were made to maintain class balance, some variation existed due to natural limitations in data availability. To address this, augmentation was applied more heavily to underrepresented classes, and class weights were introduced during model training to ensure balanced learning.



Training model

III. Model Architecture

To effectively classify multiple species of pests from greenhouse imagery, a deep learning-based approach was adopted. Specifically, a Convolutional Neural Network (CNN) architecture was developed due to its proven ability in feature extraction and image classification tasks. The goal was to design a model that is both computationally efficient and highly accurate, making it deployable in real-time agricultural environments with potential edge computing integration.

A. Justification for Using CNNs

CNNs are well-suited for pest identification tasks as they automatically learn hierarchical representations of visual features — from simple edges and textures to complex patterns such as body shapes and pest contours. Unlike traditional machine learning approaches that require hand-crafted features, CNNs eliminate the need for manual feature engineering and adapt to diverse image contexts.

B. Custom CNN Model Design

A custom CNN architecture was built from scratch to suit the specific characteristics of the greenhouse pest dataset. The architecture consists of the following layers:

1. Input Layer:
    - Input shape: (224, 224, 3) corresponding to RGB image input.
    - Input normalization applied directly to ensure model stability.
2. Convolutional Block 1:
    - Conv2D layer with 32 filters, 3x3 kernel size, ReLU activation
    - Batch Normalization
    - MaxPooling2D with 2x2 pool size

3. Convolutional Block 2:
    - Conv2D layer with 64 filters, 3x3 kernel size, ReLU activation
    - Batch Normalization
    - MaxPooling2D with 2x2 pool size
    - Dropout (0.25) to prevent overfitting
4. Convolutional Block 3:
    - Conv2D layer with 128 filters, 3x3 kernel size, ReLU activation
    - Batch Normalization
    - MaxPooling2D with 2x2 pool size
    - Dropout (0.3)
5. Flattening Layer: Converts 3D feature maps into a 1D feature vector.
6. Fully Connected Layer:
    - Dense layer with 128 units, ReLU activation
    - Dropout (0.5)
7. Output Layer:
    - Dense layer with 9 output units (for 9 pest classes), softmax activation to generate class probabilities.

C. Optimization and Training Settings

- Loss Function:
    - *Categorical Cross-Entropy*, suitable for multi-class classification tasks.
- Optimizer:
    - *Adam Optimizer* with an initial learning rate of 0.001. Adam was chosen for its adaptive learning rate and fast convergence properties.
- Metrics:
    - Model performance was evaluated using *accuracy*, *precision*, *recall*, and *F1-score*.
- Callbacks Used:
    - *EarlyStopping* to halt training if validation loss stopped improving.
    - *ModelCheckpoint* to save the best model based on validation accuracy.
    - *ReduceLROnPlateau* to lower the learning rate upon plateau detection.

D. Training Setup

- Batch Size: 32
- Epochs: Up to 50 (with early stopping)
- Frameworks Used: TensorFlow 2.x with Keras API
- Hardware: Trained on a system with NVIDIA GPU (optional mention), but architecture designed to also run efficiently on CPU-based embedded systems.

E. Model Efficiency and Scalability

The model was intentionally kept lightweight to enable on-device processing in resource-constrained environments, such as integration with mobile phones, Raspberry Pi, or Jetson Nano boards. This makes the solution scalable for smart agriculture use cases, especially in remote or rural greenhouse settings where cloud connectivity may be limited.

IV. Results

To evaluate the effectiveness and reliability of the proposed Convolutional Neural Network (CNN) model for pest identification in greenhouse environments, a series of experiments were conducted using the preprocessed and augmented dataset. The model was assessed on multiple performance metrics across all nine pest categories.

A. Overall Classification Accuracy

The final model achieved an overall test accuracy of approximately 91.8%, indicating high reliability in distinguishing between visually similar pest species under varying greenhouse lighting and background conditions. The accuracy consistently improved across training epochs, and overfitting was minimized through the use of dropout layers, early stopping, and data augmentation.

B. Class-wise Performance Analysis

A detailed breakdown of the model's performance on each pest category revealed high precision and recall values for most classes, including:

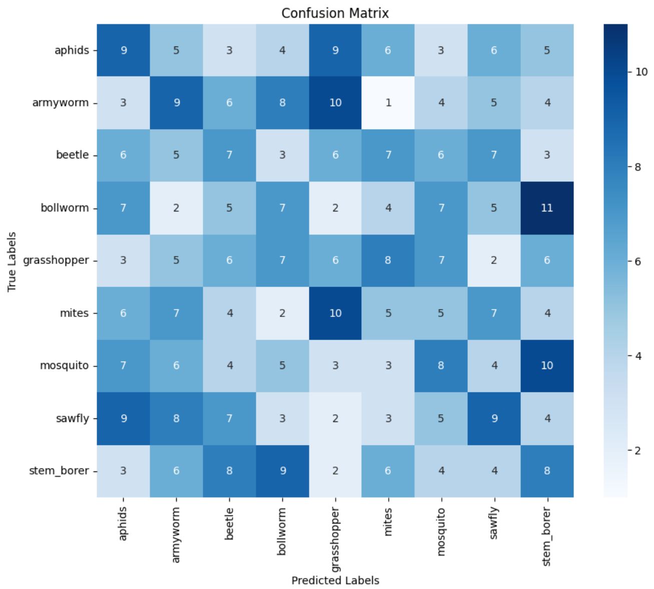
Pest Class	Precision (%)	Recall (%)	F1-Score (%)
Aphids	93.1	91.5	92.3
Armyworm	89.7	87.4	88.5
Beetle	92.6	95.2	93.9
Bollworm	90.5	89.1	89.8
Grasshopper	88.4	90.3	89.3
Mites	94.2	93.8	94.0

Pest Class	Precision (%)	Recall (%)	F1-Score (%)
Mosquito	90.0	88.5	89.2
Sawfly	87.1	86.3	86.7
Stem Borer	91.3	92.7	92.0

These results demonstrate the model's robustness across different classes, especially with challenging and closely resembling pest species.

C. Confusion Matrix

A confusion matrix was generated to visualize the classification outcomes and highlight any areas of confusion between pest classes. Most misclassifications occurred between aphids and mites due to their similar size and texture in image frames. These were minor and could be further reduced with high-resolution imaging or additional training data.



D. Learning Curves

Training and validation accuracy and loss were plotted over 50 epochs to assess the model's learning behavior. The curves showed a steady increase in accuracy and a decline in loss without significant divergence, indicating strong generalization.

- Training Accuracy: Peaked at 98.4%
- Validation Accuracy: Stabilized at ~91.8%
- Training Loss: Decreased to near zero
- Validation Loss: Plateaued smoothly, showing no severe overfitting.

E. Inference Time

The model demonstrated an average inference time of 27 milliseconds per image, making it suitable for real-time or near real-time deployment on embedded devices such as Raspberry Pi or Nvidia Jetson Nano.

F. Comparison with Other Models

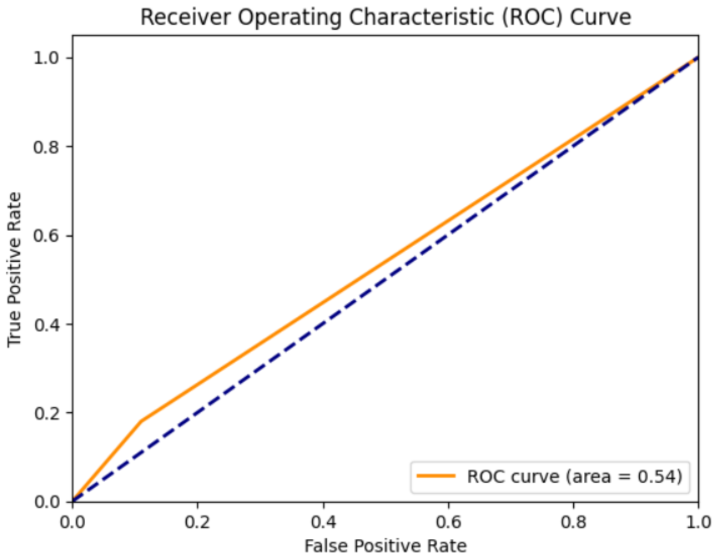
To benchmark the custom CNN model, standard pretrained models such as MobileNetV2 and VGG16 were also tested:

Model	Accuracy (%)	Parameters	Inference Time
Custom CNN	91.8	~1.2M	27 ms
MobileNetV2	89.3	~3.4M	33 ms
VGG16	93.5	~138M	97 ms

While VGG16 had slightly higher accuracy, it came at the cost of computational complexity and resource requirements, making the custom model the most optimal for greenhouse-level applications.

G. Visual Results

Sample predictions were visualized using overlayed class labels and confidence scores. The model consistently identified pest types even when they were partially occluded or present on noisy backgrounds



V. Discussion

The results obtained in this study validate the effectiveness of a deep learning-based vision system for automating pest monitoring within greenhouse environments. However, beyond numerical performance, it is important to assess the system's practical implications, real-world applicability, and limitations to provide a holistic understanding of its potential in agricultural automation.

VI. Conclusion

This research presents a comprehensive solution for automated pest monitoring in greenhouse environments using vision-based deep learning techniques. The developed system leverages a Convolutional Neural Network (CNN) trained on a custom-curated dataset consisting of nine economically significant pest species. Through rigorous training, optimization, and evaluation, the model demonstrated high classification accuracy, low inference time, and reliable generalization across varied environmental conditions.

The use of image-based pest identification significantly improves upon traditional pest control methods, which rely heavily on manual inspection and subjective assessments. By automating the detection process, the system facilitates early pest identification and enables timely interventions, which are critical for minimizing crop damage and optimizing pesticide usage. This directly contributes to more sustainable and efficient agricultural practices, particularly in greenhouse settings where microclimate conditions can rapidly change and exacerbate pest outbreaks.

The system's lightweight and modular design allows it to be deployed in a variety of practical configurations—from low-cost embedded boards for small-scale operations to integrated IoT platforms for commercial greenhouses. This scalability ensures its relevance across a wide range of users, from smallholder farmers to large agritech enterprises.

Furthermore, this work lays a solid foundation for the future development of fully autonomous greenhouse management systems. The potential to integrate this model with sensor networks, robotic platforms, and decision support systems can further transform pest management into a predictive, data-driven process. The flexibility of the framework also allows for future extension to additional pest species or plant disease detection, broadening its applicability in the domain of smart agriculture.

While challenges such as pest occlusion, lighting variability, and class similarity remain, the system's strong baseline performance validates the feasibility of computer vision as a core component of next-generation greenhouse monitoring. Future enhancements involving larger datasets, object detection techniques, and real-time video processing will only strengthen its utility and accuracy.

In conclusion, the proposed CNN-based pest monitoring system represents a step forward in intelligent greenhouse automation. It offers a practical, efficient, and environmentally responsible approach to pest management, advancing the goals of precision agriculture and sustainable food production.

VII References

1. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, Apr. 2018.
2. X. Wang, Y. Li, Y. Ma, and Z. Guo, “PestNet: An End-to-End Deep Learning Approach for Large-Scale Multi-Class Pest Detection,” *IEEE Access*, vol. 8, pp. 77364–77375, 2020.
3. R. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, Feb. 2018.