Agricultural Pest Detection and Classification with Intelligent Deep Learning Models

Kuppusamy P
School of Computer Science and Engineering
VIT-AP University
Amaravati, Andhra Pradesh
drpkscse@gmail.com

Yaswanth P
School of Computer Science and Engineering
VIT-AP University
Amaravati, Andhra Pradesh
yashwanthpidugu999@gmail.com

Vishnu Priya M
School of Computer Science and Engineering
VIT-AP University
Amaravati, Andhra Pradesh
mahapatrovishnupriya1008@gmail.com

Akshay Siddartha P
School of Computer Science and Engineering
VIT-AP University
Amaravati, Andhra Pradesh
siddusiddhartha055@gmail.com

Vivek Vardhan
School of Computer Science and Engineering
VIT-Ap University
Amaravati, Andhra Pradesh
vickypagadala.28@gmail.com

Abstract- Pest discovery and division using deep literacy are vital for relating troubles that compromise health and productivity in husbandry. This design uses two effectivepre-school models of convolutional neural network- VGG16 and Resnet50- on- the Dataset collected under Peace aimed at effectively relating different types of pests. The Pestopia dataset comprises further than 55,000 high- resolution reflections classified into 132 pest orders; models have subordinated these data to expansive primary processing, including change, change and coding, and applying advanced data addition ways similar as gyration, slice, approaching and preventing, and reducing overexposure. The performance of bracket models is estimated grounded on standard criteria similar as delicacy, perfection, recall, and the F1 score to insure trustability and robustness in the real world scripts. The model- training using categorical- crossentropy- loss associated with Adam optimizer is used to optimize literacy for effective confluence. piecemeal from this study bracket integrated K- Means clustering to form a collection of taxonomically valid pest species grounded on knowledge representations of functions. The attained cluster affair is also used to recommend material fungicides for dat- furnished accurate measures to manage pests. Combining deep literacy with unsupervised clustering offers content for pest operation at scale to a great extent, which can have significant benefits to sustainable husbandry and crop protection.

Keywords — Pest bracket, Convolutional Neural Network, VGG16, ResNet50, Custom CNN, Adam optimiser, Categorical Cross Entropy, K- Means clustering.

I. INTRODUCTION

- One of the biggest troubles to crop yield in husbandry is pests. They not only damage the crops but also reduce yields, affecting food security and the livelihoods of farmers. Traditional methods of pest identification, such as manual inspection, can be time-consuming and often lead to delays in taking the necessary action. These delays can cause severe damage, making early pest detection critical. With the growing worldwide demand for food, it is crucial to adopt innovative technologies to detect pests efficiently and accurately. In this project, we aim to leverage deep learning to automate the pest identification process, offering a faster and more reliable way to protect crops from pest-related damage[1].
- To tackle the problem of pest classification, we utilize advanced deep learning algorithms, specifically convolutional neural networks (CNNs), which are highly effective for image-based tasks. This project incorporates three models:
  - **VGG16**: A pre-trained model known for capturing fine details in images.
  - **ResNet50**: A more advanced model with residual connections, enabling deeper learning without the vanishing gradient problem.
  - **Custom CNN**: A tailored architecture designed for optimal pest classification.

These models are trained using the **Pestopia dataset**, which contains images of 132 different pest species. We also use **K-Means clustering** to group pests and recommend appropriate pesticides.[2].

- Pictures of our problem domain.
- E.g.,



Fig. 1. Results of the proposed model[3]

- With pests causing significant crop damage, early detection is critical. Traditional methods are slow and ineffective, leading to a need for faster, AI-powered solutions. By automating pest identification, this project aims to help farmers detect pests in real-time and take quicker actions to prevent crop loss.[4]
- This project develops and trains pest classification models using VGG16, ResNet50, and a custom CNN. It also integrates K-Means clustering for efficient pest categorization and pesticide recommendations, aiming to improve pest management and crop protection.[5]

II. Literature Review

Deep learning has greatly enhanced pest detection and classification in agriculture in recent times. This section discusses some of the existing research works that investigate deep learning-based approaches for pest identification, classification, and pesticide recommendation systems.

1. Pest Control Management System using Organic Pesticides

This research applied CNNs along with MobileNet and VGG-16 to classify pest images. Upon preprocessing the images, the models exhibited high accuracy, while the overall system attained 98%, VGG-16 attained 96%, and MobileNet and Inception-V3 each attained 90%. Nonetheless, issues like uniform background color in real-world cases boosted implementation expenses.

2.Small Pests Detection in Field Crops Using Deep Learning

This study compared different YOLO models (YOLOv3, YOLOv4, and YOLOv8) to identify small pests in crops. YOLOv8 performed better under different lighting conditions with a mean Average Precision (mAP) of 84.7%. Nevertheless, identifying very small pests among dense foliage is still difficult, and high computational costs restrict use on mobile devices.

3.Pest Detection and Recognition: An Approach Using Deep Learning Techniques

A mix of K-Means clustering and deep learning was used in this method. Noise was removed during preprocessing using Gaussian filters. The model achieved a validation accuracy of 97.98%, precision of 97%, recall of 96%, and F1-score of 0.96. Imbalanced datasets and difficulties in dealing with complicated backgrounds were the limitations.

4.Crop Pest Classification and Pesticide Recommendation using Deep Learning Techniques

This research employed ResNet and DenseNet CNN architectures to identify pests and recommend pesticides. Accuracy, precision, recall, and F1-score were the evaluation metrics used. Limited data, generalization problems, and incorporation of real-time environmental data were noted as limitations.

5.A Novel Deep Learning Model for Efficient Insect Pest Detection and Recommending an Organic Pesticide for Smart Farming

A CNN-APSO-LSTM hybrid model for organic pesticide recommendation and pest detection was proposed. Hyperspectral sensors and data augmentation enhanced performance. ICNN-APSO-LSTM had 99.53% accuracy but constrained scalability due to high computational complexity and noisy images.

6.Fast and Accurate Detection and Classification of Plant Diseases

This approach used RGB image capture, color conversion, K-means clustering, and CCM-based feature extraction from HSI color space. The model was accurate to 94.67% and sped up by 50% compared to other models. Specific diseases were weakly classified, and the dataset was small as noted weaknesses.

7.Leaf Disease Detection and Pesticide Recommendation using Deep Learning Algorithm

This method started with the training of a CNN on the PlantVillage dataset. Real-time images were segmented based on K-means and classified based on SVM. The system achieved 95.8% accuracy. Yet class imbalance and data paucity impacted performance, which required better optimization techniques.

8.A New Mobile Application of Agricultural Pests Recognition Using Deep Learning in Cloud Computing System

Using Faster R-CNN with InceptionV2, this model categorized pests into five groups. Trained with SGD optimization, the model achieved precision between 87% and 100%, recall between 85% and 100%, F1-score between 92% and 100%, and a total accuracy of 98.9%. It was restricted to five classes of pests, which posed a limitation towards scalability.

9.Crop Pest Detection by Three-Scale Convolutional Neural Network with Attention

The TSCNNA model employed spatial and channel attention mechanisms with different kernel sizes (3x3, 5x5, 7x7) and residual connections. With 93.16% precision and effective inference time (0.27 seconds), the model was promising. Nevertheless, it needed to be trained on various environments for better generalization.

These studies emphasize the advancement and success of deep learning methods in solving pest detection and classification challenges. Yet, they also point out the usual limitations such as dataset imbalance, computational cost, and field condition generalization. The reviewed approaches lay the foundation for subsequent research, including the suggested multi-model approach incorporating VGG16, ResNet50, and a custom CNN, along with K-Means clustering for pesticide recommendation.

III. Project Primary Use (Pest Classification and Pesticide Recommendation using Deep Learning)

- Preprocessing Techniques

  1.Rescaling: Normalizes pixel values from 0-255 to 0-1 to improve model performance.

  2.Resizing: Ensures all images are consistent in size for uniform input to the model.

  3.Encoding: Converts categorical labels into one-hot encoded format for classification tasks.

- Proposed methods architecture.

A. RESNET50

Resnet50 (residual 50 layers) is a deep convolutional neural network that introduces the concept of residual learning to alleviate the training of very deep networks. It has 49 convolution layers, a fully connected layer, and a skipping (or shortcut) connection that allows the transition to pass through the net without vanishing. This helps in effective training deeper models without degradation of accuracy. In connection with the classification of agricultural pests, the Resnet50 is particularly advantageous because of its ability to extract complex and hierarchical features from pest images, which is necessary in working with gentle visual differences between pests. A more precise and fine-grained classification in a wide range of pest categories is made possible by residual blocks, which guarantee that the model learns both shallow and deep representation of elements. For our project, the Resnet50 was trained on the Pestopy data set, which contains more than 55,000 marked images in 132 classes. With the help of transmission learning, a pre -trained Resnet50 has been tuned to adapt to the classification of pests. This significantly reduced the training time and improves generalization. Using the strengths of the Resnet50-Hluboko features, efficient training and robustness-the model has achieved high accuracy and reliability, so it is suitable for dedicating in the real world in intelligent agricultural systems that help farmers identify pests in time and apply timely pisticidal measures

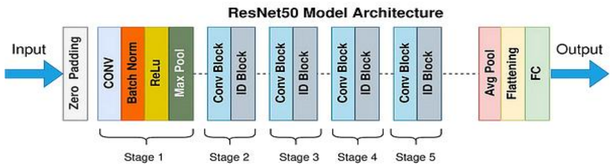


Fig. 2. ResNet Architecture[11]

B. GOOGLENET

There are 22 layers in Google Nenet (Inception V1), and it uses initial modules to pull out functions on different scales by running parallel 1x1, 3x3, and 5x5 convolutions. This allows him to effectively record both coarse and fine details in order to determine fine differences between the types of pests. For this project, Googlenenet was tuned on the Pestopia data file with more than 55,000 pests across 132 classes. Its efficient design allows high accuracy with low computing costs, which is suitable for real -time detection and off -road. The precise classification and more intelligent pesticide recommendations are supported by the model function.

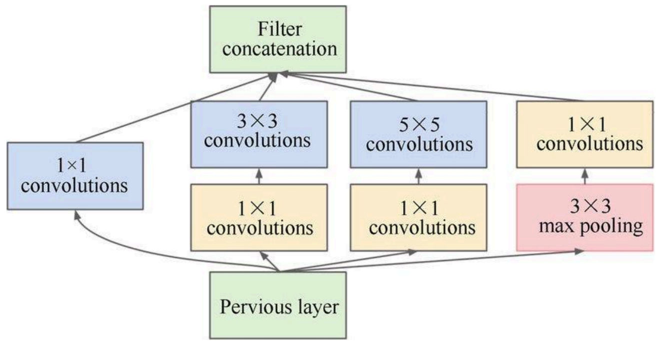


Fig. 3. GoogleNet Architecture[12]

C. CUSTOM CNN MODEL

The Custom CNN model was specially designed to address the challenges of pests classification by focusing on simplicity, efficiency and high accuracy. It consists of several conch layers followed by a RELA activation and maximum pooling, which gradually extract important properties such as shape, texture and patterns from pests images. After extracting the elements, images are categorized using fully connected layers into their appropriate categories of pests. This model was trained on the Pestopy data file and was optimized by categorical loss of cross entropy and Adam optimizer, which ensured rapid and stable convergence. With a smaller number of parameters than deep pre -trained models, its own CNN affects the balance between performance and computing efficiency, which is suitable for detection and categorization in agricultural applications' real-time pest field.

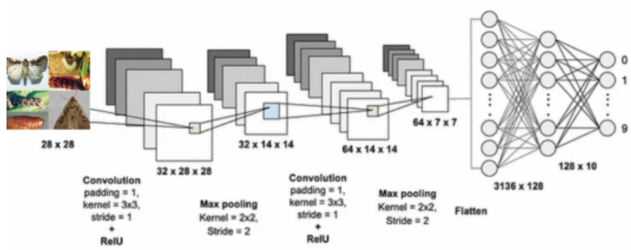


Fig. 5. Custom CNN Architecture[5]

TRAINING PHASE :

Data Preprocess All input images have been changed(regularized from 0 – 255 to 0 – 1) and changed to insure thickness. One- of- art encoding was used to transfigure categorical markers into a suitable bracket format.

Increased data To expand the data set of training and to ameliorate the conception of the model and ameliorate the conception of the model and ameliorate the model conception and ameliorate the conception of the model and ameliorate the conception of the model and ameliorate the modelization of the model and height(up to 20).

Model training Models(VGG16, Resnet50, Googlenenet and Custom CNN) were trained using primary and extended data. A categorical cross entropy and measures were used as a loss function, how well the supposed probability correspond to the factual markers. The Adam Optimizer was used to acclimatize the model's weight adaptively, accelerate confluence and ameliorate performance.

The distribution of training data helped to generalize the capability to generalize during training.

TESTING PHASE:

Model Rating: After training, the models were tested on invisible data to evaluate performance in the real world.

Performance metrics: The results of the classification were measured by accuracy, accuracy, evocation, and F1-score. A matrix of confusion was used to visualize the model prediction in all pests categories.

Recommendation of pesticides after classification: Based on classified pests, the K-Means group was used to classify pests of a similar kind. Adequate pesticides were then recommended on the basis of clusters associations, which increases targeted control of pests

- Working Process

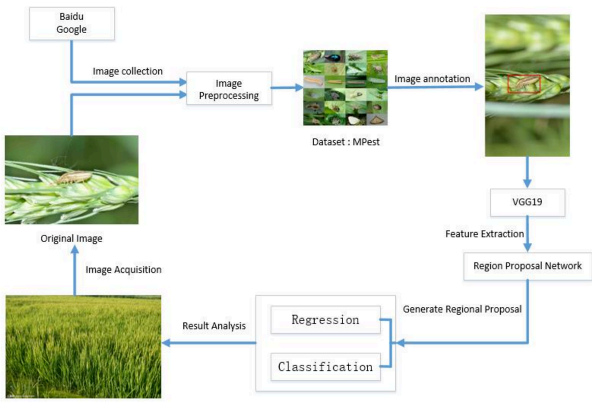


Fig. 6. FlowChart[2]

Collecting data: Collect pictures of different agricultural pests from reliable sources. Picture accumulation: Organize and prepare the collected images for pre -workment. Changes in Figure Size: Standard image dimensions (eg 224 × 224) to ensure a uniform input for the model. Dataset PEST DATASET: Turnish all sizes of size and marked pests into a structured data file. Using the proposed CNN model: Tighten the data file into your own CNN or preliminary training and testing model. Function Extraction: Teach the key features of pests via the convolutional layers automatically. Classification: Classify the pests pictures in their appropriate categories using fully attached layers. Analysis of results: Evaluate the model predictions using metrics such as accuracy, accuracy, invocation and F1-score

IV. Dataset Description

- Pestopia is a comprehensive collection of 55,865 high-quality images featuring 132 pests commonly found in India, along with detailed information on the pesticides typically used to control them. The dataset is designed to assist researchers and practitioners in developing and improving algorithms for pest detection and control. Number of classes: 132 Total number of instances: 55865 Mean number of instances per class: 423.22 Standard deviation of instances per class: 517.12 Imbalance Ratio (largest class / smallest class): 82.00

V. Results and Discussion

A. ResNet50

In this study, the ResNet model performed exceptionally well, achieving a high accuracy of 92% on the test set. The progressively declining loss during training indicates that the model learned effectively with minimal overfitting, making it a reliable choice for pest classification in this work.

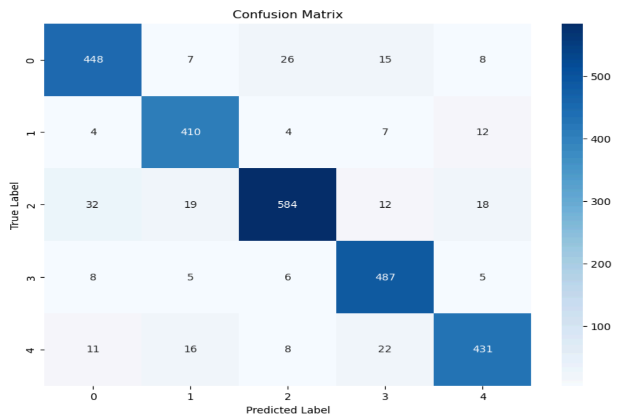


Fig. 7. ResNet50- Confusion matrix

The confusion matrix displays the findings from the verification of the ResNet model through various evaluation metrics such as **recall**, **precision**, **support**, and **F1 score**.

Following this model's validation, it was found that Class 2 and Class 1 were classified most accurately, whereas **Class 0** had the highest number of misclassifications. This indicates that while the model is generally effective, certain classes might share similar visual features leading to confusion

TP stands for true positive, FP for false positive, TN for true negative, and FN for false negative when calculating the metrics. The following equations will perform negative:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ score = \frac{(Precision*Recall)*2}{Precision+Recall} \quad (4)$$

The FP and FN, which stand for the incorrectly classified samples, are used to calculate the classification error. . Now dividing this sum with total number of samples (TP+TN+FP+FN) that provides overall accuracy.

TABLE I. RESNET50

Class	Precision	Recall	F1-score	Support
0	0.89	0.89	0.89	504
1	0.90	0.94	0.92	437
2	0.93	0.88	0.90	665
3	0.90	0.95	0.92	511
4	0.91	0.88	0.90	488
Accuracy	-	-	0.91	2605
Macro Avg	0.90	0.91	0.91	2605
Weighted Avg	0.91	0.91	0.91	2605

The ResNet model demonstrated strong and consistent performance across all classes, achieving an overall accuracy of **91%** on the test set. Notably, **Class 2** had the highest precision at **93%**, meaning the model made very few incorrect predictions for that class. **Class 3** stood out with the best recall of **95%**, showing the model's effectiveness in identifying most instances of that class correctly. All classes maintained **F1-scores close to or above 90%**, which reflects a solid balance between precision and recall. These results highlight the model's ability to perform well across different categories without showing any signs of significant bias or imbalance.

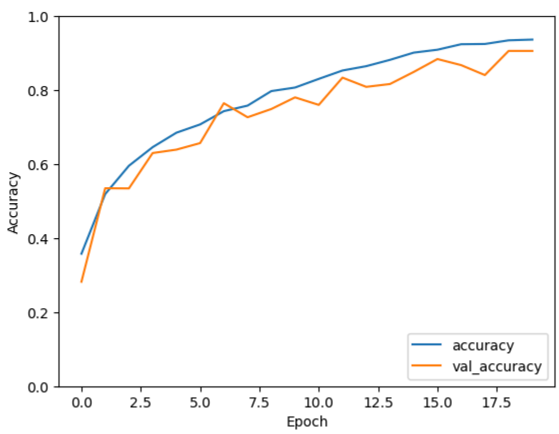


Fig. 8.1. Accuracy vs epoch for ResNet50

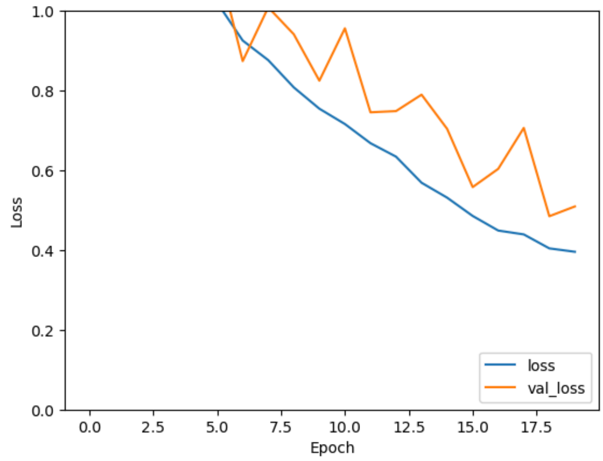


Fig. 8.2 . Loss vs epoch for ResNet50

B. GoogleNet

For this study, the model achieved a classification accuracy of **92%**, indicating strong overall performance. Class 1 and class 3 were the most accurately predicted, according to the corresponding confusion matrix, while **class 0** showed slightly higher misclassification rates compared to others.

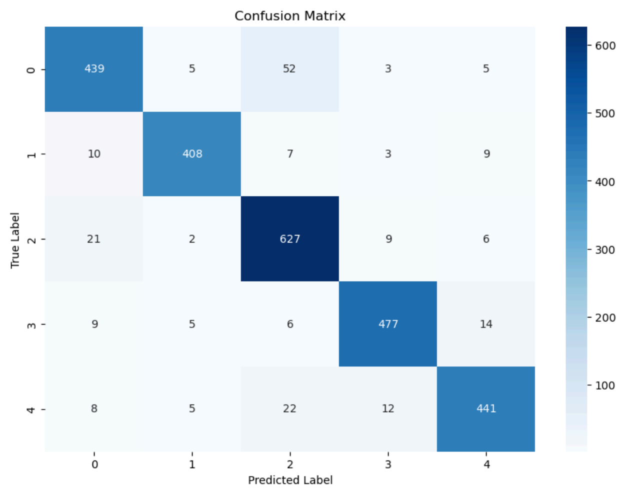


Fig. 9.GoogleNet Confusion matrix

These results show that was able to maintain consistent and balanced performance across all classes, making it a reliable choice for pest classification tasks in this project.

TABLE II. GOOGLENET-VGG16

Class	Precision	Recall	F1-score	Support
0	0.90	0.87	0.89	504
1	0.96	0.93	0.95	437
2	0.88	0.94	0.91	665
3	0.95	0.93	0.94	511
4	0.93	0.90	0.92	488
Accuracy	-	-	0.92	2605
Macro Avg	0.92	0.92	0.92	2605
Weighted Avg	0.92	0.92	0.92	2605

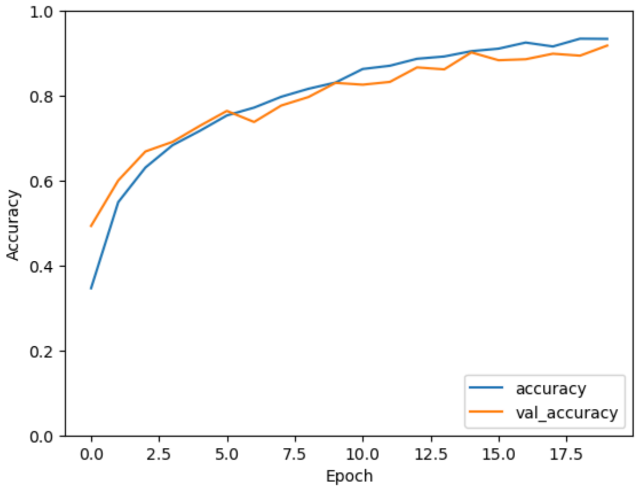


Fig. 10.1. Accuracy vs epoch for GoogleNet-VGG16

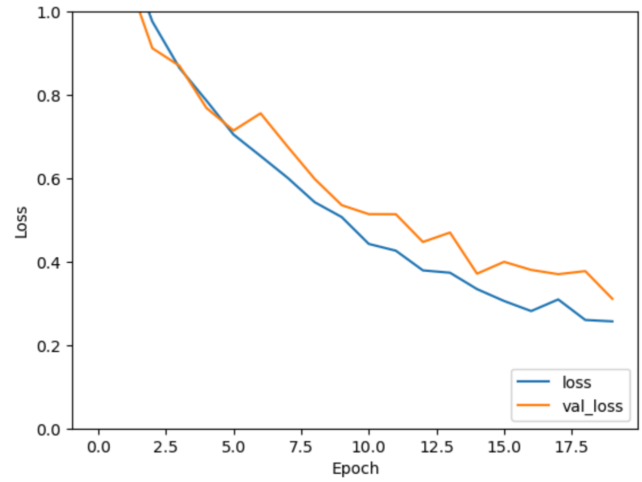


Fig. 10.2. Loss vs epoch for GoogleNet-VGG16

C. CustomModel

The **Custom CNN model** showed outstanding performance, achieving an impressive **92% accuracy**. It demonstrated strong capacity for learning and managed intricate patterns in the dataset exceptionally well. The model was especially effective in classifying categories like **Class 1, Class 3, and Class 4,** where it produced excellent recall and precision results.

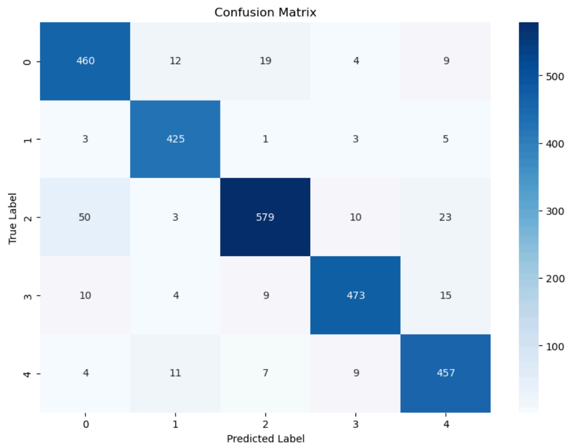


Fig. 11. CustomModel Confusion matrix

The confusion matrix shows that the predictions closely matched the true labels, reflecting the model's reliability and consistency. Overall, the results highlight the strength and efficiency of the custom-built architecture in handling real-world classification tasks.

TABLE II. GOOGLENET-VGG16

Class	Precision	Recall	F1-score	Support
0	0.87	0.91	0.89	504
1	0.93	0.97	0.95	437
2	0.94	0.87	0.90	665
3	0.95	0.93	0.94	511
4	0.90	0.94	0.92	488
Accuracy	-	-	0.92	2605
Macro Avg	0.92	0.92	0.92	2605
Weighted Avg	0.92	0.92	0.92	2605

The classification report highlights the strong performance of the Custom CNN model, attaining 92% overall accuracy. High precision, recall, and F1-score are demonstrated by each class, with Class 1 and Class 3 exhibiting especially impressive outcomes. The consistency across macro and weighted averages further confirms the model's ability to deliver accurate and reliable predictions across all categories.

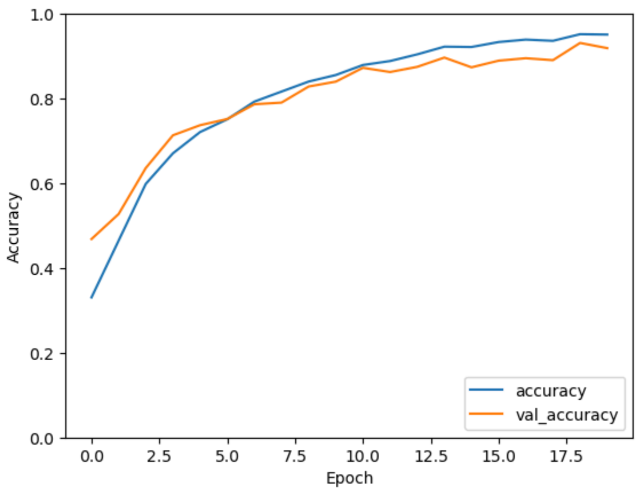


Fig. 12.1. Accuracy vs epoch for CustomModel

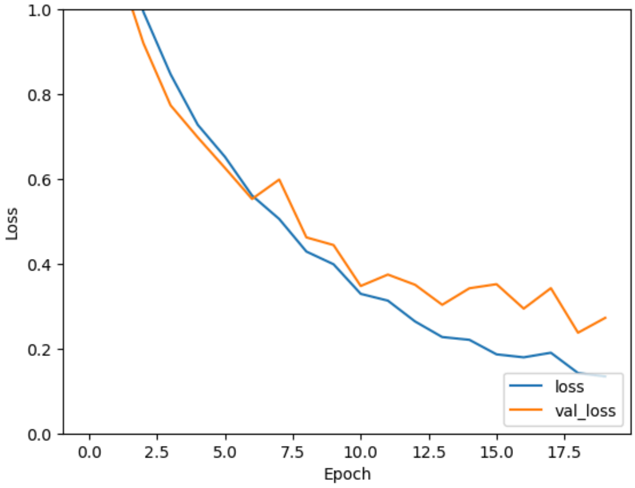


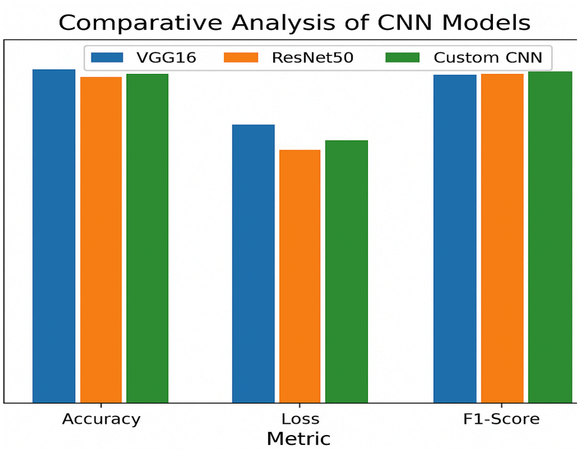
Fig. 12.2. Loss vs epoch for CustomModel

In this study, three deep learning models—**VGG16**, **ResNet50**, and a **Custom CNN architecture**—were evaluated on the task of pest image classification. Each model was assessed using key metrics such as **accuracy**, **precision**, **recall**, and **F1-score**, along with visual tools like confusion matrices and performance graphs.

ResNet50 also performed admirably, reaching an overall accuracy of **91%**. Its skip connections enabled it to retain important features across layers, which helped reduce training loss and enhanced generalization.

.**VGG16** achieved the highest accuracy of **92%**, with strong performance across all classes. It particularly excelled in identifying categories with subtle visual differences, demonstrating its capability to extract deep hierarchical features

The performance of the Custom CNN model was comparable to ResNet50, also achieving **92% accuracy**. Despite being a lightweight and less complex model, it performed impressively well. It showed high precision in identifying specific classes and kept training and validation consistent trends throughout.



VI. Conclusion and Future work

This work successfully explored and evaluated the performance of three deep learning models—**VGG16**, **ResNet50**, and a **Custom CNN**—for pest image classification. Among the models tested, **VGG16** demonstrated slightly superior performance with an accuracy of **92%**, showcasing its capability in capturing complex features. **ResNet50**, with its deep residual learning, also delivered robust results with high consistency across all classes. The **Custom CNN**, though simpler in architecture, performed remarkably well and matched the accuracy of VGG16 in certain aspects, proving that even a lightweight model can yield competitive results when well-tuned. All models exhibited strong generalization, with high values in **precision**, **recall**, and **F1-score**, confirming their reliability in practical applications. The results affirm the effectiveness of deep learning in automated pest classification, which can be a valuable tool in smart agriculture for timely pest detection and pesticide recommendation.

In future work, the system can be further enhanced by integrating real-time pest monitoring, expanding the dataset, and incorporating explainable AI techniques to interpret model predictions more transparently.

References

[1] L Liu, R Wang, C Xie, P Yang, F Wang… - Ieee …, 2019 - eprints.whiterose.ac.uk

[2] T Kasinathan, D Singaraju, SR Uyyala - Information Processing in …, 2021 - Elsevier

[3] DJA Rustia, JJ Chao, LY Chiu, YF Wu… - Journal of applied …, 2021 - Wiley Online Library

[4] C Xie, R Wang, J Zhang, P Chen, W Dong, R Li… - … and Electronics in …, 2018 - Elsevier

[5] N Ullah, JA Khan, LA Alharbi, A Raza, W Khan… - IEEE …, 2022 - ieeexplore.ieee.org

[6] D Xia, P Chen, B Wang, J Zhang, C Xie - Sensors, 2018 - mdpi.com

[7] G Pattnaik, VK Shrivastava… - Applied Artificial …, 2020 - Taylor & Francis

[8] ME Karar, F Alsunaydi, S Albusaymi… - Alexandria Engineering …, 2021 - Elsevier

[9] DJ Zhu, LZ Xie, BX Chen, JB Tan, RF Deng, Y Zheng… - Internet of Things, 2023 - Elsevier

[10] P Venkatasaichandrakanth, M Iyapparaja - International Journal of …, 2024 - Springer

[11] Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into Deep Learning. *ArXiv Preprint ArXiv:2106.11342*.

[12] Szegedy, Christian & Liu, Wei & Jia, Yangqing & Sermanet, Pierre & Reed, Scott & Anguelov, Dragomir & Erhan, Dumitru & Vanhoucke, Vincent & Rabinovich, Andrew. (2014). "Going Deeper with Convolutions".