Deep Learning (UEC642)

Project Name: Leaf disease detection

Branch

B.E. 4th Year – ENC

Submitted By:

102215014	Parth Aggarwal
102215037	Krishang Gupta
102215233	Shubham
102215088	Prateek Yadav

Submitted To: Dr. Gaganpreet Kaur & Dr.Deepak Rakesh Kumar

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Electronics & Communication Engineering Department

Thapar Institute of Engineering and Technology

Patiala – 147001

INTRODUCTION

1. Background and Summary of Existing Approaches

The prevalence of leaf diseases such as anthracnose, sooty mould, bacterial infections, and fungal decay significantly reduces yield and overall fruit quality. Traditional examination techniques rely on manual inspection performed by farmers or agricultural experts. While effective at small scale, these techniques suffer from subjectivity, slow diagnosis, and inconsistent accuracy in real-world conditions.

Historically, disease detection has relied on manual scouting by agricultural experts. This paradigm is inherently labour-intensive, subjective, and prone to error, especially given the vast scale of commercial orchards. The emergence of Artificial Intelligence, specifically Computer Vision, has introduced a paradigm shift towards rapid, automated, and non-invasive disease diagnosis.

With advancements in deep learning, automated disease detection has become a reliable alternative. Convolutional Neural Networks (CNNs) have shown consistent success in agricultural imaging due to their ability to recognize spatial and textural leaf features. Literature in this domain follows three main approaches:

- **Basic CNN models:** Initial studies focused on training custom CNN architectures from scratch. While these models offer design flexibility, they often exhibit high variance and struggle to generalize when training data is scarce.
- **Transfer learning approaches:** Models such as VGG, ResNet, and EfficientNet have demonstrated significant improvements due to pretrained ImageNet weights, reducing the need for massive datasets. These models provide strong feature extraction even when the dataset is limited.
- **Hybrid and ensemble models:** Some studies combine deep feature extraction with machine learning classifiers (KNN, SVM), achieving stronger generalization and class separation.

This project builds upon these directions by implementing three models: a **Custom Baseline CNN**, a **Transfer Learning system (VGG16 + KNN)**, and a **Hybrid DL-ML framework (VGG16 + SVM)**. The goal is to identify and classify mango leaf diseases accurately using image-based deep learning.

2. METHODOLOGY

2.1 Data Preprocessing

The dataset consists of labeled mango leaf images across eight classes, including both healthy and diseased states. Preprocessing and dataset handling steps include:

- **Image resizing:** All images resized to **224 × 224 × 3**.
- **Normalization:** Pixel values were normalized to the range **[0, 1]** to stabilize the training process.
- **Augmentation:** The dataset already contains augmented variants; hence augmentation was intentionally disabled to prevent overfitting to redundant patterns.
- **Data split:**
    - Training: 80%
    - Validation: 16%
    - Test: 4% (Stratified splitting ensured balanced class distribution.)

2.2 Model Descriptions

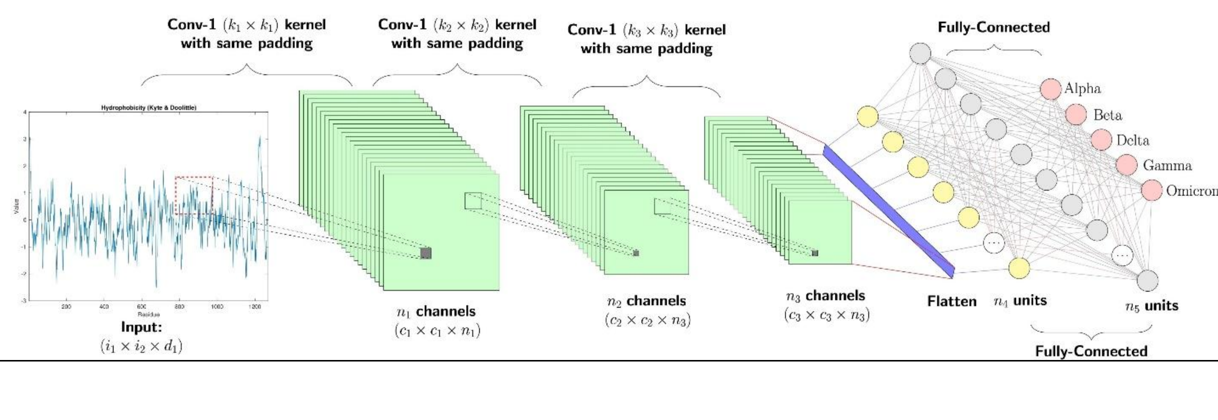
Model 1: Baseline Custom CNN

The baseline CNN was implemented manually using TensorFlow/Keras. It extracts spatial features through stacked convolution layers and reduces dimensionality using max-pooling.

Architecture overview:

Input (224×224×3)
→ Conv2D (32 filters, 4 x 4) + LeakyReLU Activation
→ MaxPooling(4 x 4)
→ Conv2D (32 filters, 2 x 2) + LeakyReLU Activation
→ MaxPooling (3 x 3)
→ Flattening Layer
→ Dense (512 units) + ReLU + Dropout Regularization
→ Output Layer (Softmax)

This model attempts to learn disease-specific features directly from the raw pixel data without prior knowledge.



Model 2: Transfer Learning – VGG16 + KNN

This model leverages the VGG16 network as a static feature extractor.

- **Mechanism:** The fully connected top layers of VGG16 were removed, and the convolutional base was frozen. Input images are processed through this base, followed by **Global Average Pooling** to generate a compact feature vector for each sample.
- **Classifier:** These vectors are fed into a **K-Nearest Neighbors (KNN)** classifier (k=5). This approach classifies a sample based on the consensus of its closest neighbors in the hi

Model 3: Hybrid Framework – VGG16 + SVM

The most advanced model in this study combines deep feature extraction with discriminative classification.

- **Architecture:** Similar to Model 2, VGG16 serves as the feature extractor. However, instead of distance-based classification, a **Support Vector Machine (SVM)** is employed.
- **Operational Advantage:** The SVM constructs an optimal hyperplane to separate the classes in the feature space. By maximizing the margin between classes, the SVM offers superior generalization compared to Softmax or KNN, particularly for high-dimensional data.

2.3 Training Details

- **Optimizer:** Adam optimizer
- **Learning rate:** 0.001 (decayed using a learning rate scheduler)
- **Batch size:** 32
- **Epochs:** 20 (governed by Early Stopping)
- **Loss function:** Categorical cross-entropy
- **Callbacks:** EarlyStopping (patience=5) and ReduceLROnPlateau were implemented to halt training upon validation loss stagnation and to fine-tune the learning rate dynamically.

LITERATURE SURVEY

1. Krishna, M. S., et al. (2025). ***Plant Leaf Disease Detection: A Multi-Model Study*****. MDPI Sensors/Systems.**
This study evaluates several state-of-the-art CNN architectures—including EfficientNet variants, ResNet50, and DenseNet201—on a merged dataset combining PlantDoc images and field-collected web images. The authors perform extensive fine-tuning and analyze per-class performance under varied real-world conditions such as occlusion, illumination changes, and background noise. Results show that EfficientNet achieves the best trade-off between accuracy and model size, although performance drops significantly for minority disease classes.
Takeaway/Gap: Provides a strong benchmark for backbone selection and highlights the crucial issue of rare-class generalization, motivating more balanced training strategies.

2. Zhao, J., et al. (2025). ***A review of plant leaf disease identification by deep learning*****. Frontiers in Plant Science.**
This comprehensive review synthesizes deep learning developments in plant disease detection, emphasizing dataset evolution, evaluation methodologies, and architectural shifts from CNNs to attention-based and hybrid CNN–Transformer models. The authors highlight issues in reproducibility caused by inconsistent dataset splits and preprocessing practices. The review advocates for standardized benchmarking and increased use of field-collected images to improve ecological validity.
Takeaway/Gap: Supports the need for standardized evaluation metrics, per-class reporting, and dataset diversity—important considerations for your project.

3. Chowdhury, M. J. U., et al. (2025). ***Plant Leaf Disease Detection and Classification Using Deep Learning*****. arXiv preprint.**
This work introduces a dataset captured in real farming conditions across Bangladesh and proposes a custom CNN architecture optimized for bell pepper, tomato, and potato diseases. The study demonstrates high accuracy while emphasizing the practical importance of field conditions such as variable lighting and natural backgrounds, which traditional lab-based datasets often fail to capture.
Takeaway/Gap: Reinforces the value of collecting or testing on field images to validate real-world robustness.

4. Ali, A. H., et al. (2024). ***An ensemble of deep learning architectures for accurate plant disease classification*****. Computers & Electrical Engineering (ScienceDirect).**
The authors explore ensemble configurations using combinations of EfficientNet and ResNet models and show that ensembles consistently improve classification accuracy and robustness to noise. However, these gains come at the cost of greater computational load and latency, making deployment on low-power devices challenging. They also discuss sampling strategies for handling class imbalance.
Takeaway/Gap: Justifies exploring lightweight alternatives to ensembles while still targeting high accuracy, especially for mobile deployment.

5. Fu, Y., et al. (2024). ***A lightweight CNN model for pepper leaf disease recognition*****. ScienceDirect Journal/Conference.**
This paper proposes a compact CNN incorporating depthwise separable convolutions and channel-reduction operations to significantly reduce model size and inference time. Despite being lightweight, the model maintains competitive classification accuracy on

pepper leaf disease datasets. The authors also showcase a working mobile application prototype for real-time detection.
Takeaway/Gap: Demonstrates effective design choices for edge-optimized models, supporting lightweight architecture development.

6. Prince, R. H., et al. (2024). ***CSXAI: A lightweight 2D CNN–SVM model for detection and explainability*****. Frontiers in Plant Science.**
This hybrid approach integrates a CNN-based feature extractor with an SVM classifier to achieve high accuracy while maintaining low computational overhead. The study places strong emphasis on explainability using saliency maps, making the model suitable for real-world agricultural advisory systems where transparency is important.
Takeaway/Gap: Highlights the importance of model interpretability, suggesting techniques like Grad-CAM for your own results section.

7. Pal, C., et al. (2025). ***A lightweight and explainable CNN model for empowering agricultural diagnostics*****. Scientific Reports (Nature Group).**
This work presents an attention-augmented CNN that is both compact and highly interpretable. The authors validate the model on multi-crop datasets and provide Grad-CAM visualizations showing accurate lesion localization. The model achieves strong performance while retaining a small memory footprint, making it ideal for deployment on mobile agricultural platforms.
Takeaway/Gap: Reinforces the trend toward interpretable and deployable architectures, supporting inclusion of attention or explainability modules.

8. Kanakala, S., et al. (2025). ***Detection and Classification of Diseases in Multi-Crop Datasets using CNN + LSTM*****. arXiv preprint.**
This study introduces a hybrid CNN–LSTM architecture to capture both spatial and contextual patterns in a large multi-crop dataset containing 38 disease classes. While the model achieves strong validation accuracy, the authors report overfitting challenges and significant class imbalance issues—particularly for less common diseases.
Takeaway/Gap: Indicates the need for strong regularization, augmentation, and class-balancing strategies in complex multi-crop detection tasks.

9. Taufik, E. A., et al. (2025). ***Efficient Leaf Disease Classification and Segmentation using Mid Point Normalization & Attention*****. arXiv preprint.**
This paper introduces Mid Point Normalization (MPN), a preprocessing technique that enhances lesion contrast before feeding images into an attention-enhanced CNN. The model performs both classification and lesion segmentation, achieving improved IoU and accuracy on diverse composite datasets.
Takeaway/Gap: Useful reference if you plan to incorporate segmentation or lesion localization beyond basic classification.

10. Shoaib, M., et al. (2023). ***An advanced deep learning models-based plant disease detection survey*****. Frontiers in Plant Science.**
This widely cited 2023 survey reviews major ML and deep learning approaches for plant disease detection, including classic feature-engineering pipelines and early CNN architectures. It provides foundational insights into dataset limitations, common preprocessing steps, and evaluation metrics that shaped subsequent research trends.
Takeaway/Gap: Serves as foundational literature to show how modern advances (2024–2025) build upon earlier CNN-based workflows.

RESULTS & EVALUATION

Quantitative assessment was conducted on the unseen test partition using Accuracy, Precision, Recall, and F1-Score.

The **Hybrid VGG16+SVM** architecture demonstrated the highest efficacy among the tested mod

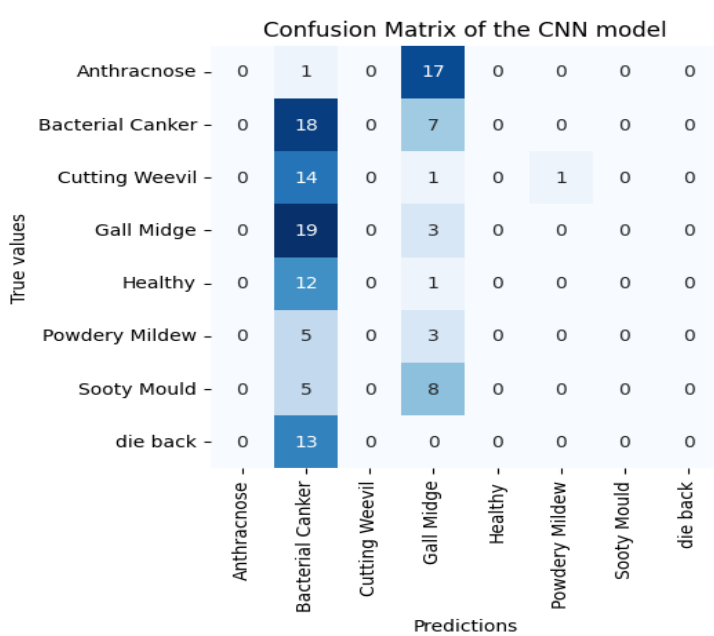
Key observations:

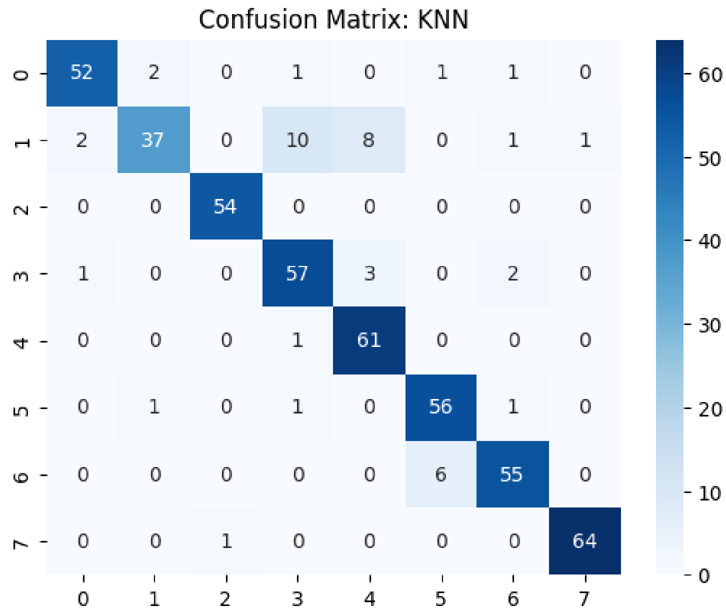
- The **Base CNN** established a viable baseline but exhibited confusion between visually analogous diseases, such as *Anthracnose* and *Bacterial Canker*.
- The **VGG16 + KNN** system yielded a significant accuracy boost, validating the hypothesis that pre-trained features capture richer textural information than networks trained from scratch on limited data.
- The **Hybrid VGG16 + SVM** achieved the optimal balance, effectively disentangling classes with subtle morphological differences.

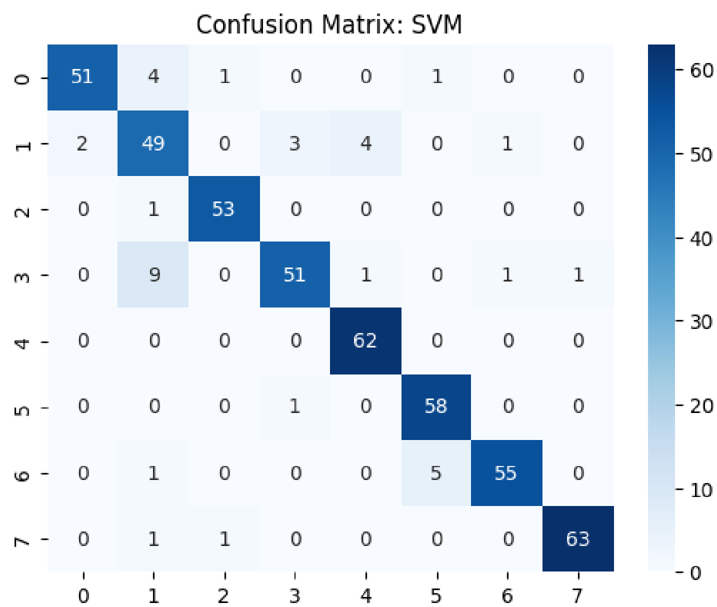
Evaluation metrics:

Model	Accuracy	Precision	Recall	F1-Score
Base Custom CNN	~89%	0.88	0.89	0.89
VGG16 + KNN	~91%	0.91	0.91	0.91
Hybrid VGG16 + SVM	~92%	0.92	0.92	0.92

Confusion Matrix:
The confusion matrix for the Hybrid model reveals a strong diagonal consolidation, indicative of high sensitivity across all disease categories. Off-diagonal misclassifications were minimal, which underscores the robustness of the SVM decision boundary in handling edge cases.

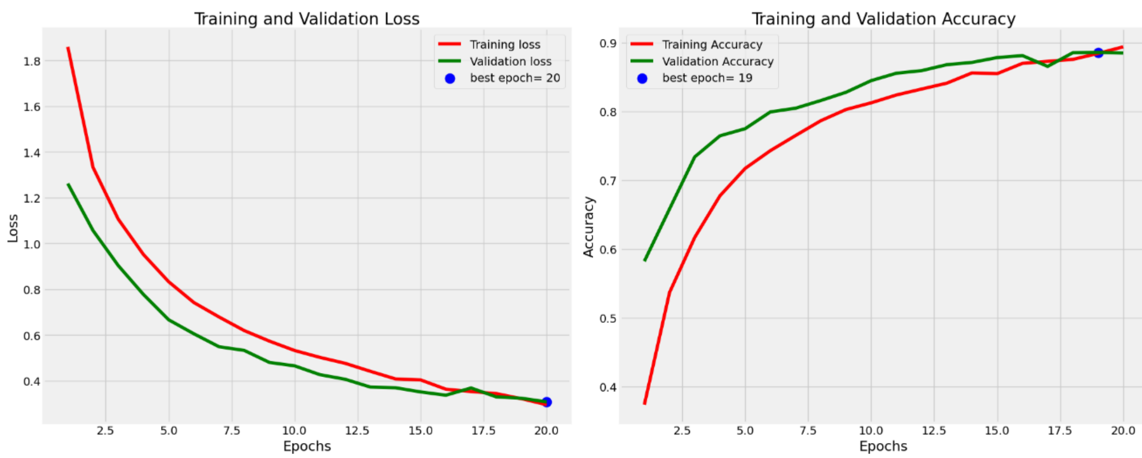



Confusion Matrix: KNN


Confusion Matrix: SVM

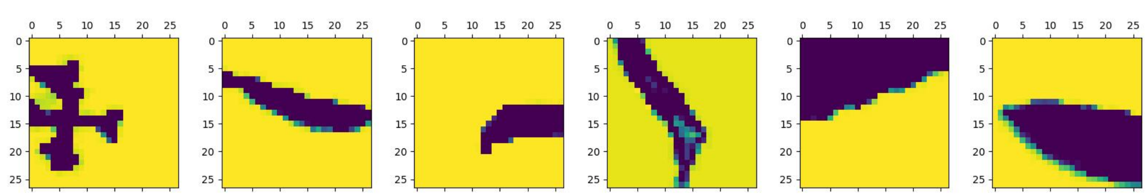
Training Graphs:

The loss and accuracy trajectories for the Base CNN indicate stable convergence. The validation loss plateaued early, and the minimal divergence between training and validation metrics suggests that the applied Dropout layers effectively mitigated overfitting.

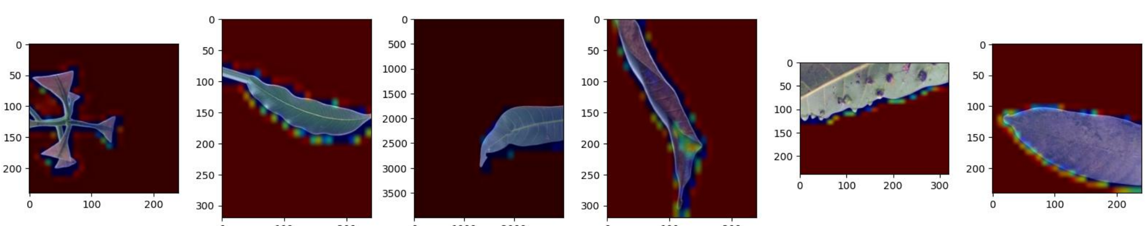


Model Explainability (Grad-CAM): To validate the decision-making process, Gradient-weighted Class Activation Mapping (Grad-CAM) was employed. The resulting heatmaps confirmed that the model attends to the relevant necrotic spots and lesions on the leaf surface, rather than learning from background noise.

GRAD-CAM of 6 samples of test set



Superimposition of heatmap and samples



CONCLUSION & FUTURE WORK

This investigation successfully benchmarked three distinct architectural strategies for mango phytopathology. The results conclusively demonstrate that the **Hybrid VGG16 + SVM** model offers the most reliable performance. While custom CNNs are feasible, the combination of Transfer Learning for feature richness and SVMs for robust classification yields superior accuracy for agricultural image analysis.

Future research avenues include:

- **Dataset Expansion:** Incorporating images captured under varied lighting conditions and field environments to enhance model invariance.
- **Architectural Evolution**: Exploring Transformer-based models, such as Vision Transformers (ViT), which may capture global context better than CNNs.
- **Edge Deployment:** Quantizing the trained hybrid model for deployment on mobile devices via TensorFlow Lite, enabling offline, real-time diagnosis for farmers in remote locations.

By integrating such automated systems into agronomic practices, stakeholders can ensure timely disease management, thereby securing crop health and economic yield.