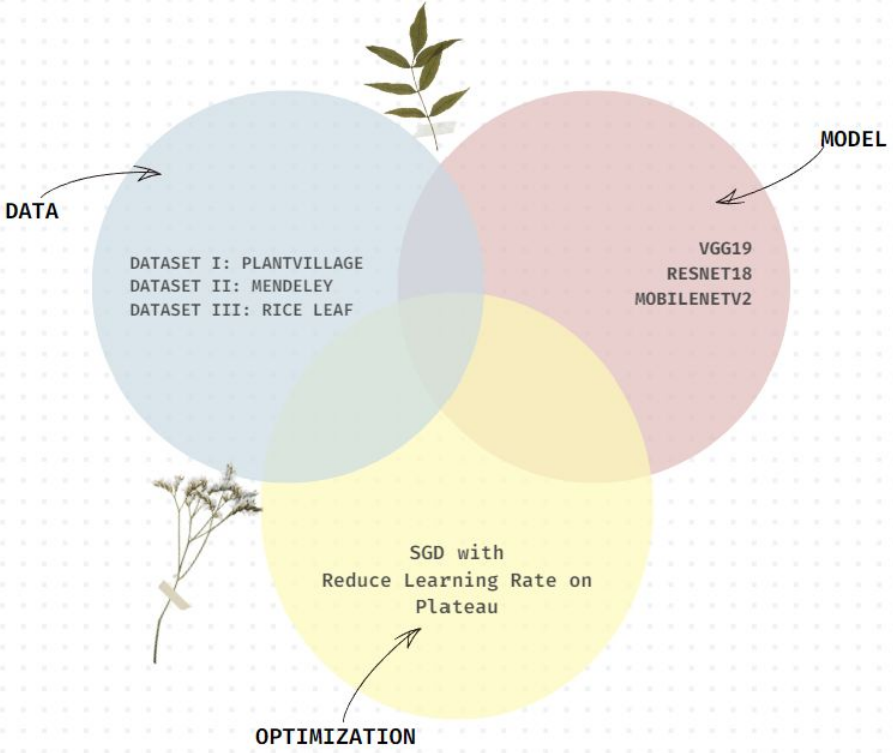
Leaf Disease Classification

COMP 6721- APPLIED ARTIFICIAL INTELLIGENCE
GROUP- B

OVERVIEW

1. CHALLENGES & GOALS
2. DATASET
3. METHODOLOGY
4. RESULT & ANALYSIS



CHALLENGES

- **Data Diversity and Quality:** Overcoming the variability in image quality
- **Complex Disease Patterns:** Handling the complexity of disease patterns and variability in visual symptoms across different plant species
- **Adaptation to New Cases:** Developing models that can generalize well to new, unseen cases of leaf diseases, not just those they were trained on.
- **Overfitting:** Preventing overfitting when models are trained on limited data
- **Computational Demands:** Managing the substantial computational resources required for training deep CNNs effectively.
- **Hyperparameter Tuning:** Identifying optimal settings for various models

GOALS

- **Performance Analysis**:CNN models—ResNet18, VGG19, MobileNetV2. Analysis on a total 11 models, 3 dataset x 3 models was conducted.
- **Dataset Complexity Impact:** Evaluate the complexity and diversity of datasets
- **Robust Model Development:** Enhancing their practical applicability in agricultural settings.
- **Methodological Improvements**: Explore and implement advanced machine learning techniques
- **Optimization and Tuning:** Fine-tune hyperparameters and explore additional optimization methods to enhance the models' accuracy and efficiency further.
- **Transfer learning**: 2 models have been analysed

DATASETS

Info	PlantVillage	Mendeley	Rice Leaf
Classes	38	22	4
Images	30.4K	24.8K	5.9K
Dimension	256*256	>400*400	300 * 300
Format	.jpg	.jpg	.jpg

Datasets:

- **PlantVillage:** 30,400 images across 38 classes; includes a variety of plant diseases.
- **Mendeley:** 24,800 images, reduced to 14,894 images across 12 classes for focused analysis.
- **Rice Leaf:** 5,932 images spread over 4 classes specific to rice leaf diseases.

Preprocessing and Data Management:

- **Uniformity:** All images resized for consistent input dimensions
- **Augmentation:** Techniques like rotation, flipping, and color adjustment used.
- **Data Splitting:** Each dataset is divided using an 80-10-10 split for training, validation, and testing
- **Goal:** To prevent overfitting and improve the models' ability to generalize, preparing them for effective real-world application.

METHODOLOGY

Data Management:

- Split Ratio: Data divided into 80% training, 10% validation, 10% testing
- Preprocessing: Uniform image resizing; augmentation through rotation, flipping; color normalization

Hyperparameter Fine-Tuning:

- Number of epochs: 10
- Optimizer: SGD
- Loss Function: Cross-Entropy
- Batch Size: 32

Transfer Learning:

- Conducted on MobileNetV2 (PlantVillage Dataset) and VGG19 (Mendeley Dataset)

RESULT ANALYSIS

ResNet18:
- High accuracy across varied datasets.

MobileNetV2:
- Exceptional efficiency, ideal for real-time applications.

VGG19:
- Strong performance on less complex datasets.

Insights from Data Analysis:

- ResNet18 and MobileNetV2 show greater adaptability to different data complexities.
- VGG19 requires careful tuning to improve performance on complex datasets.

Metric	ResNet18			MobileNetV2			VGG19		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
Accuracy (%)	97	86	100	97	89	100	69	88	100
Precision (%)	97	86	100	97	80	100	71	86	100
Recall (%)	97	80	100	97	90	100	68	88	100
F1 score	97	81	100	97	89	100	67	87	100

D1: PlantVillage; D2: Mendeley; D3: Rice Leaf

COMPARISON

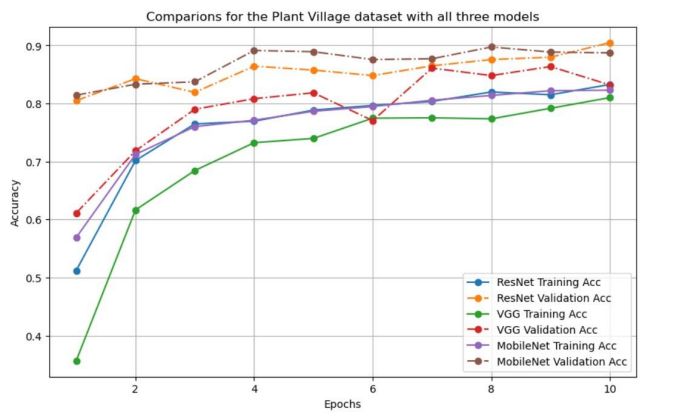


Fig 1. Comparisons for the Plant Village dataset with all 3 architectures

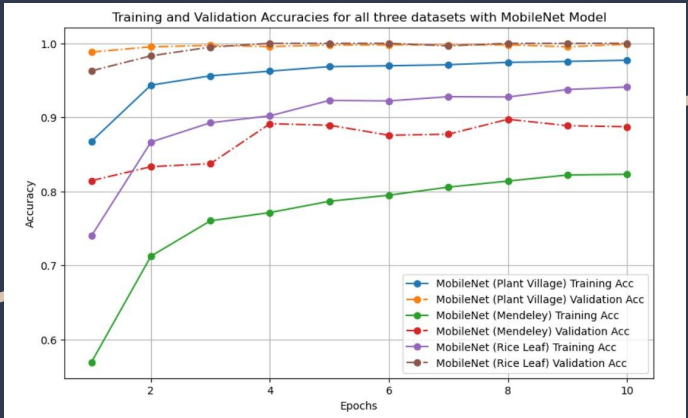
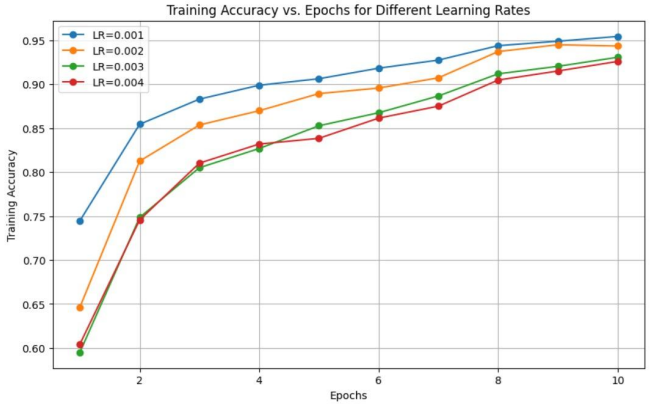
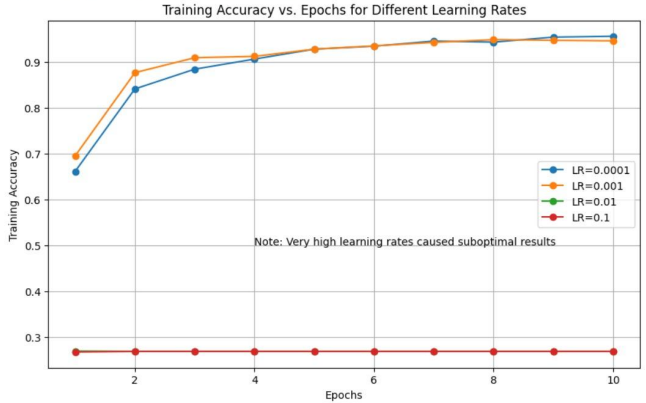


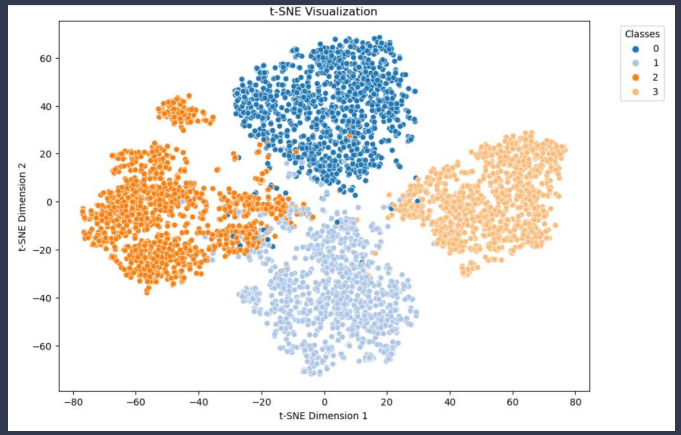
Fig 2. Training and Validation Accuracy for all three datasets with MobileNetV2

HYPERPARAMETER TUNING

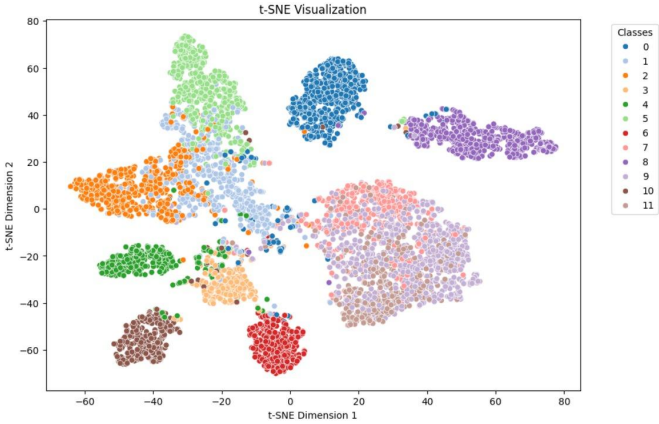




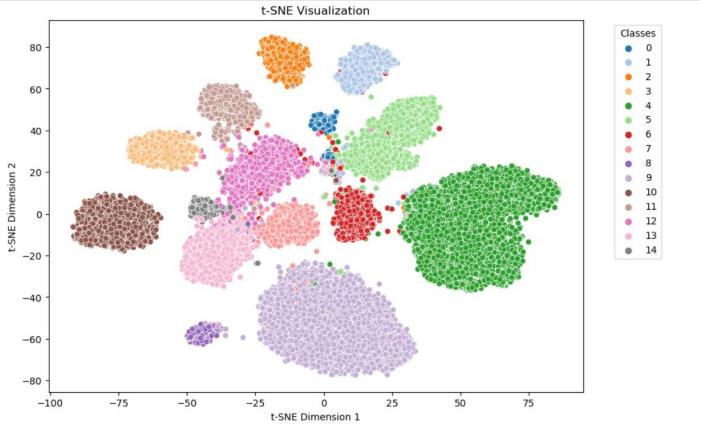
t-SNE



MobileNetV2 RiceLeaf Dataset



VGG19 Mendeley Dataset



ResNet18 PlantVillage Dataset

REFERENCES

[1] Mohanty, S.P. PlantVillage Dataset. Available at: https://github.com/spMohanty/PlantVillage-Dataset

[2] Mensah Kwabena, P. et al. Dataset for Crop Pest and Disease Detection. DOI: 10.17632/bwh3zbpkpv.1. Available at: https://data.mendeley.com/datasets/bwh3zbpkpv/1

[3] Sethy, P.K. Rice Leaf Disease Image Samples. DOI: 10.17632/fwcj7stb8r.1. Available at: https://data.mendeley.com/datasets/fwcj7stb8r/1

[4] Sirenjeevi, P. et al. Leaf Disease Identification using ResNet. DOI: 10.1109/ICECONF57129.2023.10083963. Available at: https://ieeexplore.ieee.org/document/10083963.

[5] Elfatimi, E. et al. Beans Leaf Diseases Classification Using MobileNet Models. DOI: 10.1109/ACCESS.2022.3142817. Available at: https://ieeexplore.ieee.org/document/9681085.

[6] Nguyen, T.-H. et al. A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease. Available at: https://www.mdpi.com/2624-7402/4/4/56.