भारतीय सूचना प्रौद्योगिकी संस्थान, पुणे
Indian Institute of Information Technology, Pune
(An Autonomous Institute of National Importance under act of Parliament)

BTP PROJECT

Topic : Plant Species Classification and Disease Detection Using Hybrid Model

Supervised By : Miss. Anupama Arun

Date : 06-11-2024

Team Members:
Aviral Tiwari (112215032)
Aditya Patil (112215007)
Ayush Kushwaha (112215034)
Dilesh Bisen (112215045)
Sahil Kumar Saroj (112215154)

OVERVIEW

INTRODUCTION

- Agriculture plays a vital role in feeding the world's population, and plays a major role in a nation's economy.
- But it is constantly challenged by plant diseases that can severely impact crop yield.
- Early detection and identification of plant diseases are critical to ensuring food security and maintaining ecosystem balance.
- Detecting the species accurately can also help us to prevent the disease by following appropriate practises.
- With advancements in deep learning, we can now assist in detecting plant species and their associated diseases. This project aims to use the power of Convolutional Neural Networks (CNN) to analyze plant images.

Fig 1[Generated by AI]

3

LITERATURE REVIEW

Sr No.	Authors	Feature Extractor	Classifier	Dataset	Limitations
1.	Bathula Nagachandrika Prasath R Praveen Joe I.R.[2] (April 2024)	Visual Geometry Group(VGG16), Variational Autoencoder (VAE), Visual Transformer (ViT)	Adaptive Convolutional Neural Network with Attention Mechanism	PlantVillage Dataset	Visual Transformer demands substantial computational resources, particularly for larger models and datasets.
2.	Hao Wu, Lincong Fang Qian Yu, Chengzhuan Yang (March 2024)	Local triangle features and local speeded-up robust features (SURF) using MATLAB Library	Euclidean distance to compute the dissimilarity of leaves	Swedish Dataset Flavia Dataset	• The Image used for prediction needs a uniform background, else we need extract leaf image from background. • Variation in dataset to adapt and learn
3.	Ying Bai, Xiaomei Bai (May 2024)	EFL-DesNet169	Transfer Learning ELU (Exponential Linear Unit)Activation Function	Self-built aquatic plant dataset	The research is limited by dataset dependency. Less Accurate Prediction

LITERATURE REVIEW

Sr No.	Authors	Feature Extractor	Classifier	Dataset	Limitations
4.	Shubham Sharma, Manu Vardhan (June 2024)	Sobel Edge Detection, Vein Morphometric Analysis	MTJNet (Multi-task Joint Learning Network)	Indian Medicinal Leaves Dataset	• Data Quality Dependency, • Adaptability
5.	Hatice Catal Reisa, Veysel Turkb (March 2024)	• Depthwise Separable Convolution (DSC) • Multi-Head Attention Mechanism	MDSCIRNet (proposed),SVM (LR), (RF),Adaptive Boosting, Hard Voting Ensemble Learning Model	• Potato Leaf Disease Dataset	Limited scope to specific species and diseases, Small, specific datasets limit generalization.
6.	Bowen Pan, Yulin Fang, Chonghuai Liu, Jianfu Jiang, Xiucai Fan, Ying Zhang, Lei Sun (December 2023)	• VGG16, • ResNet101, • ResNet50, • GoogleNet	Gradient weighted class activation mapping (Grad-CAM) technology	Self Built Images dataset from Zhengzhou Grape Garden of National Fruit Tree Accessions	• Confusion in identifying two species • Variations in image resolution may impact recognition results.

LITERATURE REVIEW

Sr No.	Authors	Feature Extractor	Classifier	Dataset	Limitations
7.	Silky Sachar , Anuj Kumar (September 2023)	• VGG-16 • ImageNet • CNN	XGBoost,NuSVC,Stacking Classifier with base learners: KNN, Decision Trees,NB. Logistic Regression	Flavia and Swedish	• Lack of a medicinal leaf dataset • No real-world habitat recognition
8.	Deepti Barhate, Sunil Pathak, Ashutosh Kumar Dubey. (March 2023)	CNN and VGG-16 models	HP-BSGD method, a tuned gradient descent algorithm	Flavia and Swedish	• The research is limited by dataset dependency. • High computational requirements. • Focus solely on leaves identification, not on disease classification.

6

RESEARCH GAP

Efficiency and Scalability

Existing models used prioritize accuracy but overlook real-time efficiency and scalability, presenting a gap in balancing accuracy with in-field performance.

Dual-task Classification

Most studies focus on either species identification or disease detection. The integration of both in a single model remains underexplored.

Background Complexity

Models often require uniform image backgrounds, limiting their use in natural, complex environments without additional image processing.

OBJECTIVES

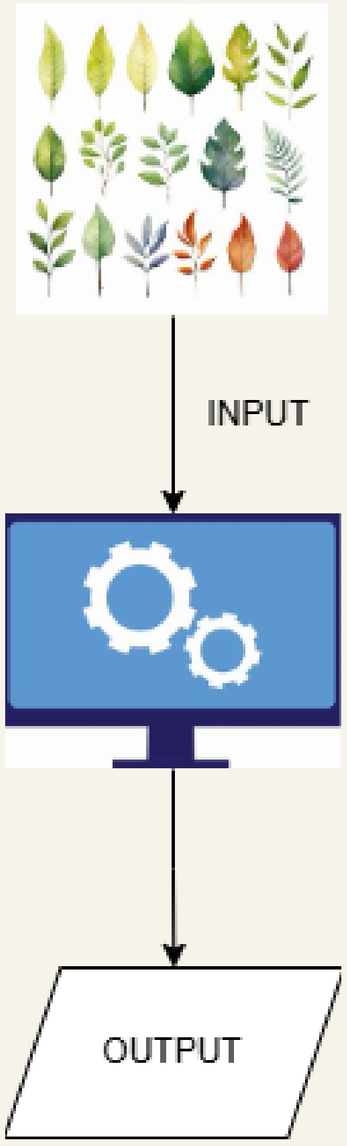

Fig 2[Generated by AI]

Develop a hybrid deep learning model leveraging VGG16, ResNet101, DenseNet121 architectures for feature extraction to identify plant species and diseases accurately.

Integrate attention mechanisms within the model to enhance focus on disease-affected areas in plant images, improving classification accuracy and model interpretability.

Optimize the model architecture for efficient training on large-scale plant image datasets, ensuring real-time disease detection in agricultural fields.

Fig 3[Generated by AI]

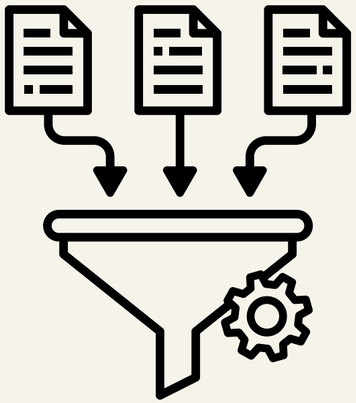

Fig 4[Generated by AI]

PROPOSED METHODOLOGY

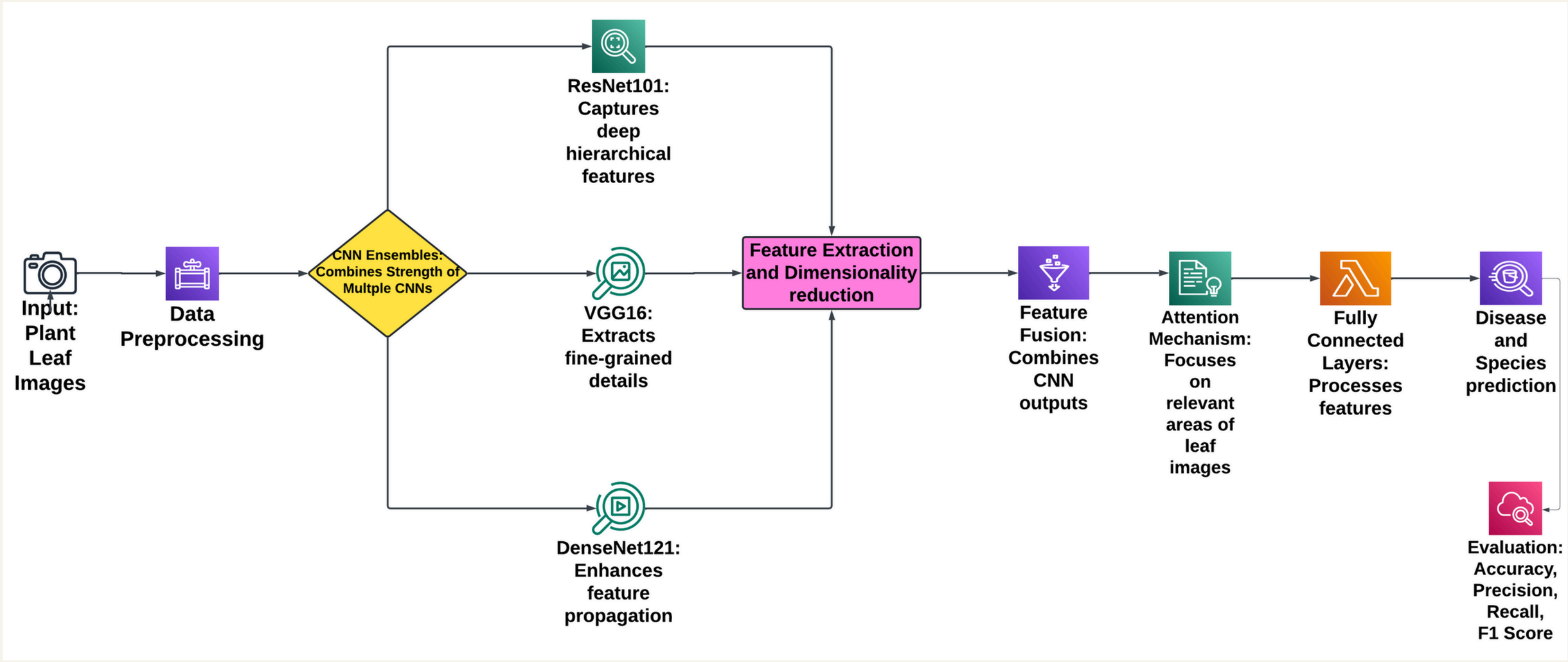

Fig 5 - Attention Mechanism

PROPOSED METHODOLOGY

Feature Extraction

- ResNet101: Uses residual connections for deep learning feature, address vanishing gradients.
- VGG16: Captures fine details with small filters, enhancing spatial representation.
- DenseNet121: Employs dense connections for efficient feature reuse and improved gradient flow.
- Combined Strengths: Leverages the unique advantages of each network for comprehensive feature representation.

Attention Mechanism

- Prioritizes Important Features: Focuses on critical areas in images.
- Improves Interpretability: Indicates which features influence predictions.
- Enhances Robustness: Adapts to variations in leaf appearance.
- Dynamic Weighting: Assigns weights based on feature significance.
- Highlights Disease Indicators: Emphasizes signs of disease for accurate detection.

DATA SET

Sr No.	Dataset	Facts About Dataset	#Classes	#Images per Classes
1.	India Soyabean Dataset[10]	• Data collected from Pune, Satara etc. from Maharashtra . • Contains healthy and insect-damaged Soyabean leaves . • Dataset is used for Disease Detection . • Comprises 6 folders with a total of 3,363 images.	6	10-200
2.	Swedish[9]	• Labelled dataset used for species classification. • Used for Species recognition . • one side Leaf Images with White Background	15	75

RESULTS

Swedish : Accuracy : 98.33%

Model Loss:

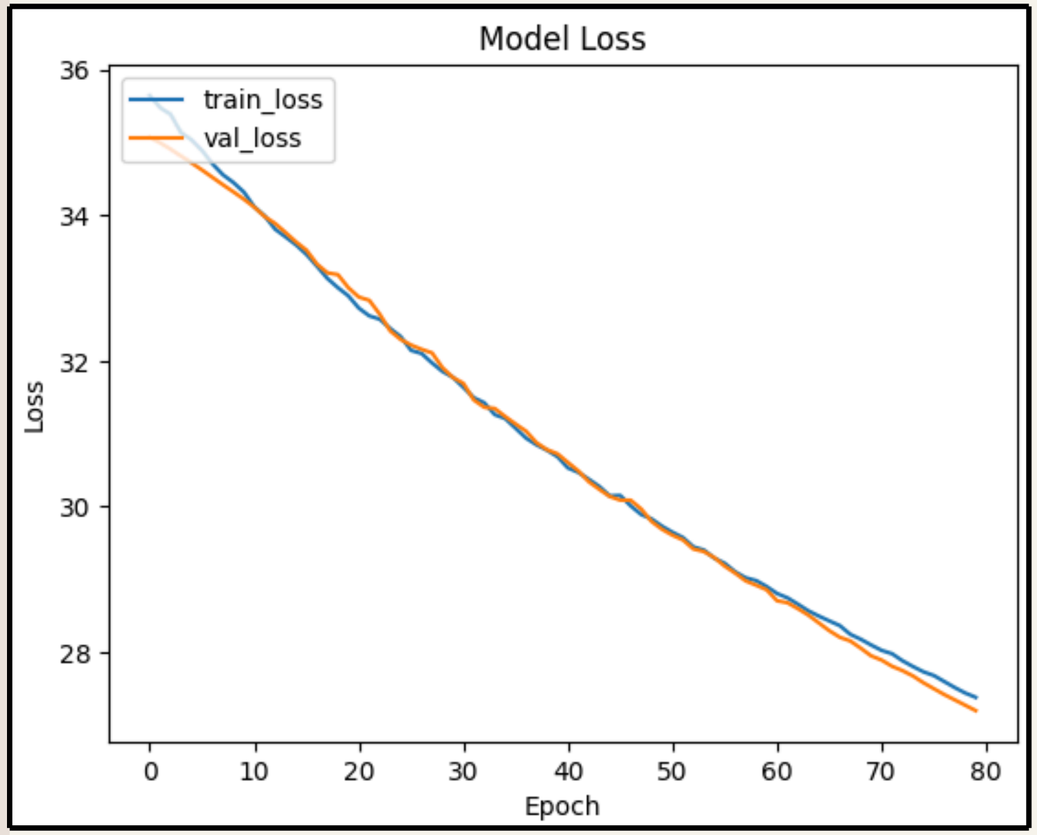

Fig 6

Model Accuracy:

Fig 7

RESULTS

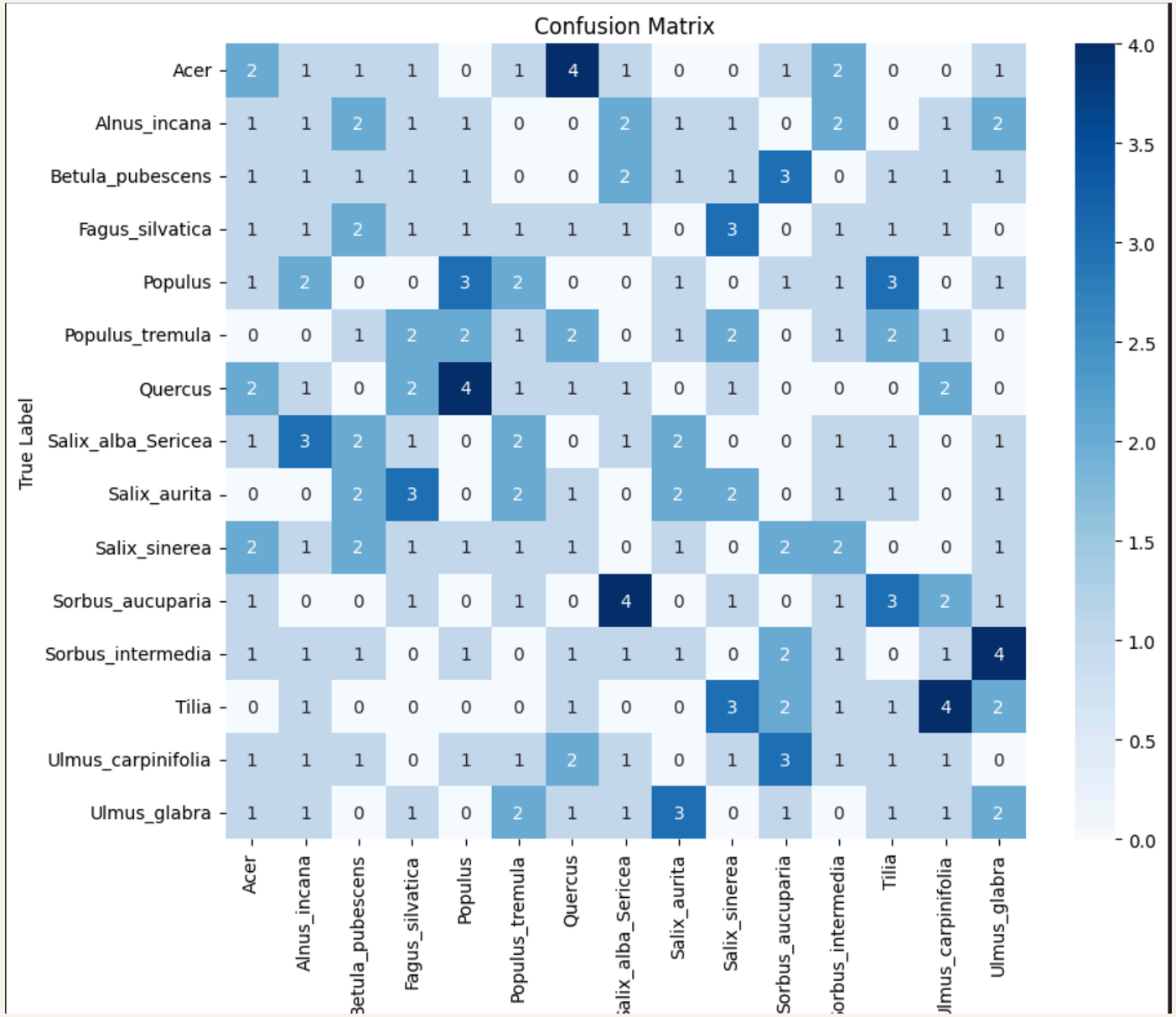

Fig 7 Swedish- Confusion Matrix

RESULTS

Indian Soyabean : Accuracy: 98.37%

Model Loss:

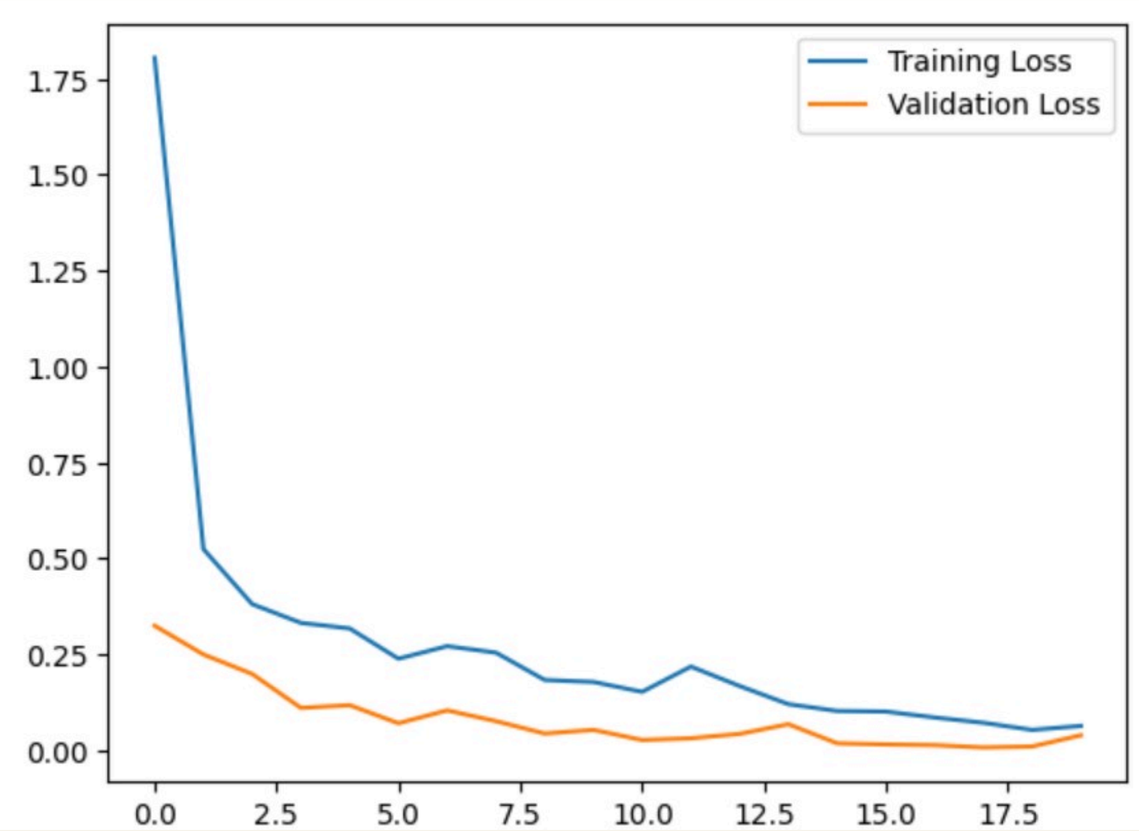

Fig 8

Model Accuracy:

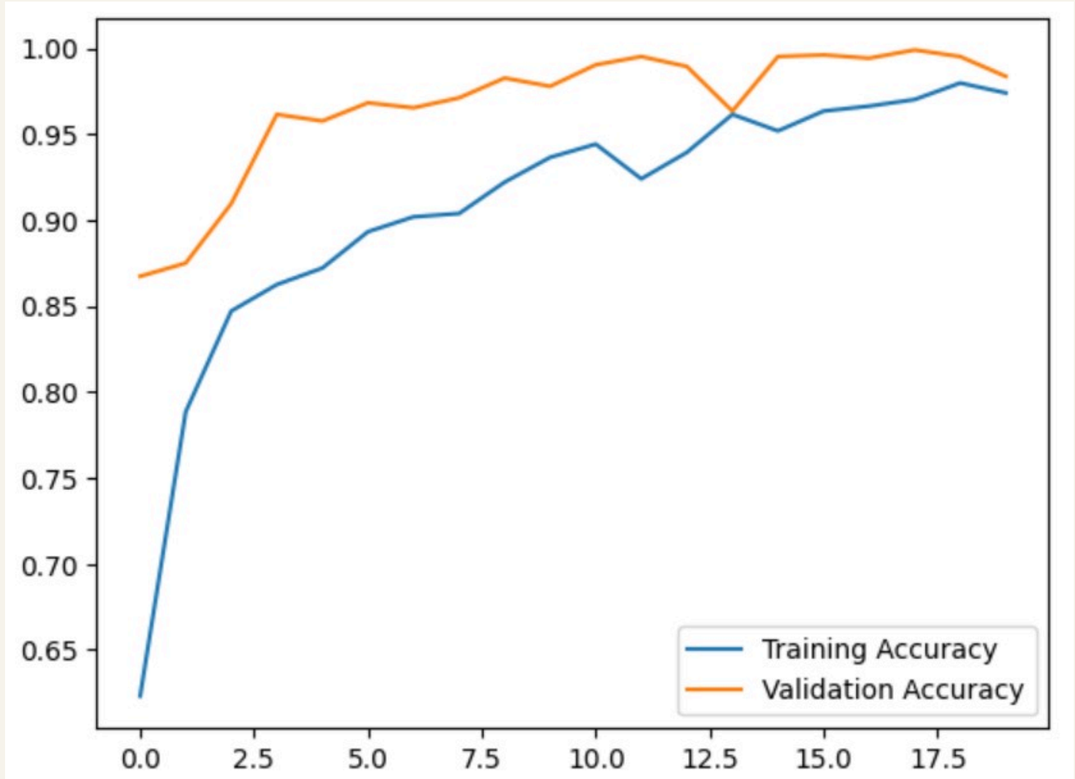

Fig 9

RESULTS

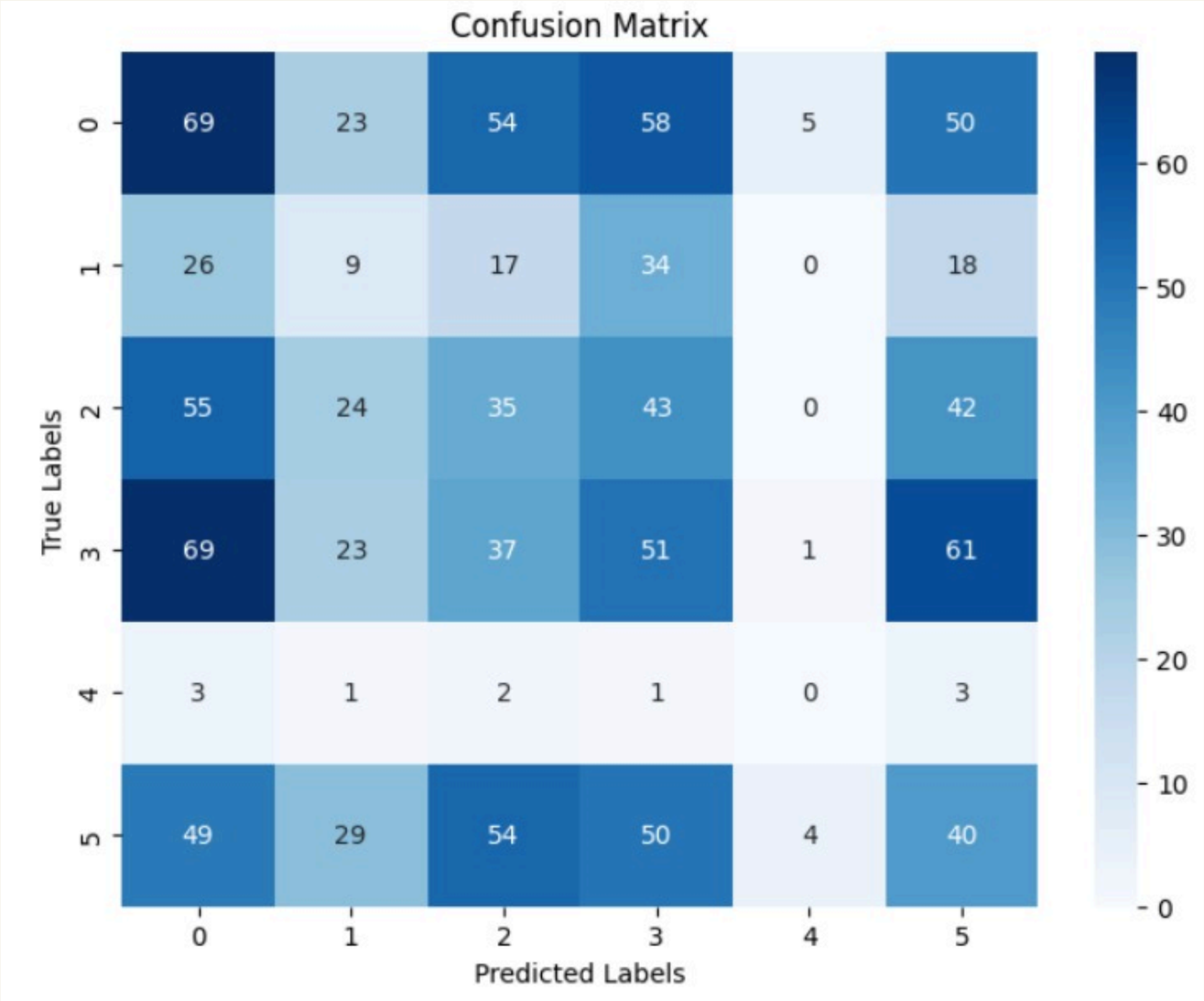

Fig 10 Indian Soyabean- Confusion Matrix

COMPARISION WITH PREVIOUS RESULTS

Swedish Dataset

Author	Method	Accuracy
Yang and Yu (2019)[11]	MFD	97.60
Hao Wu, Lincon Fang (2023) [2]	Hausdorff	98.13
Lee et al (2018) [12]	HGO-CNN	96.3
Yu et al (2021)[13]	MaskCOV	94.0
Proposed method	Hybrid Model	98.33

COMPARISION WITH PREVIOUS RESULTS

Indian Soyabean Dataset

This study marks the first comprehensive exploration of this unique dataset, establishing foundational benchmarks for future research.

Fig 11 [Generated by AI]

CONCLUSION

- In this project, we have developed a robust deep-learning model capable of accurately detecting diseases from leaf images and classifying very similar-looking species using architectures such as ResNet, DenseNet, and VGG for precise classification.

- We aim to reduce the computation through feature extraction while achieving high accuracy. The model's ability to generalize various plant species and diseases demonstrates its potential for real-world agricultural applications.

- Future work may include expanding the model to incorporate a wider variety of plant species and disease types, as well as exploring real-time deployment for practical use by farmers and agricultural professionals.

Fig 12 [Generated by AI]

REFERENCES

1) Bathula, N., Prasath, R. & Praveen Joe, I.R., 2024. An automatic classification framework for identifying type of plant leaf diseases using multi-scale feature fusion-based adaptive deep network. Biomedical Signal Processing and Control, 95(Part A), p.106316 , *https://doi.org/10.1016/j.bspc.2024.106316*

2) Wu, H., Fang, L., Yu, Q. & Yang, C., 2024. Composite descriptor based on contour and appearance for plant species identification. Engineering Applications of Artificial Intelligence, 133(Part C), p.108291 , *https://doi.org/10.1016/j.engappai.2024.108291*

3) Bai, Y. & Bai, X., 2024. Deep learning-based aquatic plant recognition technique and natural ecological aesthetics conservation. Crop Protection, 184, p.106765 , *https://doi.org/10.1016/j.cropro.2024.106765*

4)Pan, B., Liu, C., Su, B., Ju, Y., Fan, X., Zhang, Y., Sun, L., Fang, Y. & Jiang, J., 2024. Research on species identification of wild grape leaves based on deep learning. Scientia Horticulturae, 327, p.112821. Available at: *https://doi.org/10.1016/j.scienta.2023.112821*

REFERENCES

5) Sachar, S., Kumar, A. A novel transfer learning-based approach for plant species prediction using leaf images. Multimed Tools Appl 83, 40323–40336 (2024). https://doi.org/10.1007/s11042-023-17311-8

6) Barhate, D., Pathak, S. & Dubey, A.K., 2023. Hyperparameter-tuned batch-updated stochastic gradient descent: Plant species identification by using hybrid deep learning. Ecological Informatics, 75, p.102094. *https://doi.org/10.1016/j.ecoinf.2023.102094*.

7) Reis, H.C. & Turk, V., 2024. Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks. Engineering Applications of Artificial Intelligence, 133 (Part C), p.108307. *https://doi.org/10.1016/j.engappai.2024.108307*

8) Sharma, S. & Vardhan, M., 2024. MTJNet: Multi-task joint learning network for advancing medicinal plant and leaf classification. Knowledge-Based Systems, 299, p.112147. *https://doi.org/10.1016/j.knosys.2024.112147*

REFERENCES

9) Oskar J. O. Söderkvist, “Computer vision classification of leaves from Swedish trees,” Master’s Thesis, Linköping University, 2001.

10) Kotwal, Jameer; Kashyap, Ramgopal (2023), “An India soybean leaf dataset,” Mendeley Data, V1, doi: 10.17632/bshkvgbzpt.1.

11) Yang, C. & Yu, Q., 2019. Multiscale Fourier descriptor based on triangular features for shape retrieval. Signal Processing: Image Communication, 71, pp.110-119. https://doi.org/10.1016/j.image.2018.11.004.

12) Lee, S.H., Chan, C.S., Mayo, S.J. & Remagnino, P., 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recognition, 71, pp.1-13. https://doi.org/10.1016/j.patcog.2017.05.015.

13) Yu, X., Zhao, Y., Gao, Y. & Xiong, S., 2021. MaskCOV: A random mask covariance network for ultra-fine-grained visual categorization. Pattern Recognition, 119, p.108067. https://doi.org/10.1016/j.patcog.2021.108067.


INDIAN INSTITUTE OF INFORMATION TECHNOLOGY PUNE
iiit
Leading Technology.
Leveraging Technology.

THANK YOU