Sugarcane Disease Classification Using Few-Shot-Prototypical Deep Learning Techniques

John P. Q. Tomas [1], Miguel A. A. Agno [1], Paul J. M. Vale [1,*], and Marineil C. Gomez [2]

[1] School of Information Technology, Mapua University, Makati, Philippines
[2] School of Health Sciences, Mapua University, Makati, Philippines
Email: jpqtomas@mapua.edu.ph (J.P.Q.T.); maaagno@mymapua.edu.ph (M.A.A.A.); pjmvale@mymail.mapua.edu.ph (P.J.M.V.); mcgomez@mapua.edu.ph (M.C.G.)
*Corresponding author

***Abstract*—In this study, we explore the effectiveness of using a few-shot learning approach integrated with deep learning models to classify sugarcane diseases from a limited dataset. Specifically, we evaluated four models: MobileNet_V2 combined with a Prototypical Network, MobileNet_V2, ResNet50, and VGG. Our main objective was to demonstrate the potential of the Prototypical Network, paired with a lightweight model like MobileNet_V2, in addressing the challenge of limited data availability. A dataset consisting of six sugarcane disease types was used to train and test these models. The results show that the MobileNet_V2 with Prototypical Network outperformed the other models regarding accuracy, precision, recall, and overall stability, making it a strong candidate for real-world applications in resource-constrained environments. This approach demonstrated superior performance in handling small datasets and exhibited the potential for deployment in mobile or edge devices for real-time agricultural disease diagnosis. Future work will explore less significant downscaling, multi-scale processing to enhance pattern recognition, and focal loss to improve the classification of challenging samples. Additionally, incorporating explainability techniques could reveal the focus areas in model predictions, guiding feature extraction improvements for classes with subtle distinctions.**

***Keywords*—sugarcane disease classification, few-shot learning, prototypical network, MobileNet_V2, deep learning, agricultural disease diagnosis**

I. Introduction

A. Background of the Study

Amidst urgent agricultural challenges, including diseases' detrimental impact on crop yield and quality, there is a growing interest in harnessing advanced technologies like machine learning and deep learning [1]. Sugarcane disease identification remains challenging due to the variety of similar symptoms [2], posing significant challenges for farmers as even experienced farmers still fail to identify [2, 3]. Traditional methods often need to cover vast plantations, hampering timely decision-making [4]. The creation of reliable, well-coordinated diagnostic networks is a solution to resource constraints and a chance to increase the caliber and volume of these services which are provided by the pooling of technological advancements and the knowledge of experts [5].

Innovative techniques offer solutions for detecting sugarcane diseases by leveraging advanced technologies (such as deep learning) to replace the traditional disease identification method gradually [6]. Image processing can be used to analyze the patterns of diseases through large datasets and real-time processing using techniques such as Convolutional Neural Network (CNN) [7], which can identify disease symptoms in sugarcane plants. Deep learning, a subset of machine learning, processes vast datasets efficiently, transforming agriculture through precision farming techniques. Various deep learning models and techniques have emerged for sugarcane disease classification, including ensemble architectures and Bayesian learning-enhanced models [8].

B. Opportunity

Due to the nature of a disease, plant breeders and farmers are incentivized to immediately remove the diseased crop from the field, preventing the unwanted spread of the disease and promoting better crop yield. Techniques like few-shot learning hold promise but require improved datasets reflecting real-time conditions [9]. Few-shot learning approaches have also performed admirably in multiple applications in medical fields, specifically in image classification tasks [10–13].

This research addresses the limitations of existing methods for crop disease classification, which often rely on large, labeled datasets that are costly to obtain in agricultural settings. Investigating the effectiveness of few-shot learning techniques, such as Prototypical Networks, in leveraging limited labeled data for sugarcane disease classification aims to improve the efficiency and accuracy of disease detection in sugarcane farming. Coupled with a lightweight neural network such as MobileNet, this research has potential for mobile or Internet of Things (IoT) deployment. The research seeks to harness general knowledge learned from a large dataset

Manuscript received April 21, 2025; revised April 28, 2025; accepted August 15, 2025; published April 28, 2026.

so it can adapt it to certain tasks of sugarcane disease classification.

C. Problem Statement

This study aims to classify sugarcane disease and its severity. Specifically, it aims to answer the following questions:

- How does the MobileNet-Prototypical Network perform in terms of accuracy, precision, recall, F1-Score, and the Precision Recall-the Area Under Curve (PR-AUC) score for classifying sugarcane diseases with a limited number of labeled examples per class?
- How does the MobileNet-Prototypical Networks pipeline compare to baseline models like standalone MobileNet or pre-trained models, namely, ResNet and VGG in classifying sugarcane diseases?

D. Objectives

This study aims to develop a disease classification model for sugarcane using few-shot learning approaches, specifically Prototypical Networks combined with MobileNetV2, to evaluate its effectiveness in accurately identifying diseases with a limited dataset. In addition, assess the performance of the MobileNetV2 + Prototypical Network model in metrics of precision, recall, accuracy, and F1-Score, and compare it against ResNet50 and VGG16 to determine its suitability for classifying sugarcane diseases. Overall, demonstrates the effectiveness of the MobileNetV2 + Prototypical Network as a classification technique by assessing its performance on a limited dataset and highlighting its potential advantages over more complex models.

E. Scope and Limitations

The study will use neural networks, namely few-shot learning with the Mobilenetv2-Prototypical network pipeline, to extract the features and classify the dataset of sugarcane diseases.

- The data gathering will be conducted in the Philippines by the researchers;
- The images or datasets to be used are limited by the images captured by the proponents of this study;
- The dataset will be limited to sugarcane diseases in the Philippines, namely: red rot, ring spot, rust, mosaic virus and yellow leaf disease.

Furthermore, this study did not employ statistical significance testing or K-fold cross-validation during model evaluation. While performance metrics were reported on a held-out validation set, the absence of these rigorous statistical methods limits the generalizability and robustness assessment of the model's performance across different data splits or potential variations in the dataset.

F. Significance of the Study

The study focuses on classifying sugarcane diseases and optimizing a model with few-shot learning techniques. Using neural networks to extract features and classify and identify, diagnosis of sugarcane diseases will be faster, and the observations or maintenance performed by farmers and plant pathologists will be easier to achieve. By developing an accurate model, the results will benefit the following:

Farmers and Plant Pathologists: They will significantly benefit from this study as they specialize in diagnosing and treating diseases. This study will automate the identification and classification of sugarcane anomalies through visual assessment. Our study will give them more concurrent information regarding the state of their diseased crops, which will provide them with a more informed decision regarding the treatment of the crops.

Researchers: The study's outcome may provide additional or further improvements and possible advancements for future studies relating to sugarcane disease classification and other disease classification-related applications.

II. Literature Review

A. Sugarcane

Sugarcane diseases cause a significant threat to global sugar production, resulting in substantial economic losses and jeopardizing the livelihoods of millions of farmers. These diseases were classified and used for detection techniques with the help of image processing, notably the major diseases that will be discussed.

1) Sugarcane Mosaic virus

In contrast to healthy leaves (Fig. 1a), symptomatic leaves have more obvious inlay, mottles, or even stripes in colors of yellow and green that interchange with parallel veins (Fig. 1b). Leaves that are yellow or yellow white are severely infected, and the ends of new leaves get abnormally twisted (Fig. 1d) or leaves with only a few green islets or a tiny number of red punctuate necrosis (Fig. 1c). Some exhibit visible whole-leaf chlorosis. Other varieties are primarily standard green having a small number of thin pale-yellow streaks. When the temperature is high, some varieties exhibit cryptic or hazy phenomena, but when the temperature drops, the symptoms return [14].

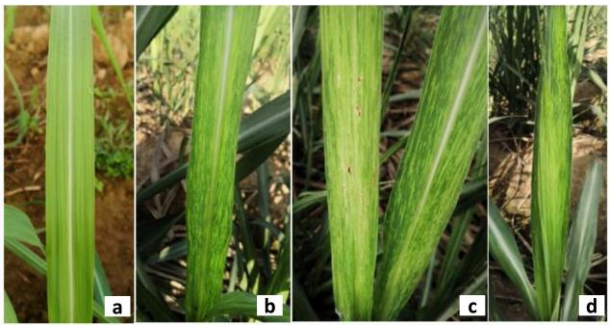

Fig. 1. Sugarcane Mosaic virus [14].

2) Red rot of sugarcane

The infection leads to changes in leaf color, with a straw hue in the middle and dark reddish-brown edges as black Acervuli develop. Infected leaves eventually hang and split at the lesions. In the midrib red rot, a red color is seen visible throughout the midrib as seen in Fig. 2. Later stages of the disease show reddening of internal tissues with white and red patches and an alcoholic scent. Symptoms

typically begin with drying and discoloration of spindle leaves, progressing from tip to base until all crown leaves wither and infected canes separate easily from nodes [15].

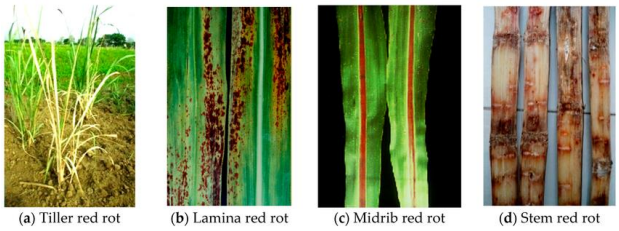

Fig. 2. Red rot of sugarcane.

3) Ring spot

Fig. 3 displays the yellow oblong patches with reddish-brown edges and a straw-colored center which are the initial symptoms of the disease. Leaf blades, leaf sheaths, and occasionally stems exhibit symptoms [16]. Sugarcane is the only known host of the ascomycete fungus Leptosphaeria Sacchari. Little bronze dots are the initial sign of the infection, which appears on the leaves of the crop's lower canopy. The spots grow longer and develop into irregularly shaped lesions with reddish-brown edges that measure 2.5 to 5 mm by 10 to 18 mm. Leaf chlorosis and necrosis might result from merging the spots. Older lesions may develop Sacchari, tiny, black spots that typically appear only on older leaves.

Fig. 3. Ring spot disease.

4) Sugarcane rust

The earliest signs of typical rust on the leaves are small, elongated, yellowish spots like the leaves in Fig. 4 [17]. The yellow spot measures 1–4 mm in length. As the disease worsens (especially on the lower leaf surface), the spots become one to three times broader and up to 20 mm longer, running parallel to the leaf venation. Furthermore, the rust pustules gradually mix and turn into flecks of orange-brown or red-brown color with a subtle but noticeable chlorotic halo. Necrotic patches form as a result of the leaf epidermis rupturing. Lesions tend to be more common toward the leaf tip and less common near the base.

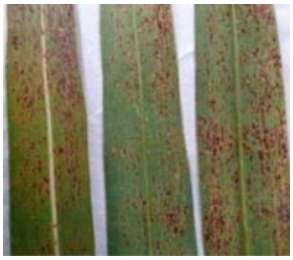

Fig. 4. Sugarcane rust.

5) Yellow leaf disease

The apparent yellowing of the leaf from the midrib, which results in necrosis, is a characteristic of ScYLD. On the elder leaves, necrosis first manifests itself. The midrib's bottom surface changes color from green to a vivid pink, reddish-brown, or yellow seen in Fig. 5. Additionally, dwarfism of the terminal internodes can be seen. The Sugarcane Yellow Leaf Virus, or ScYLV for short, is a monopartite, non-enveloped, isometric virus with a 24–29 nm diameter. It is a member of the *Luteoviridae* family of poleroviruses. The virus is composed of six Open Reading Frames (ORFs 0, 1, 2, 3, 4, and 5) with an icosahedral symmetry composed of 180 coat protein units. Its genome is a single-stranded positive sense linear RNA with 5900 nucleotides. Insect vectors, sugarcane aphids (*Melanaphis Sacchari*), effectively spread the virus in a sustained, circulative way.

Fig. 5. Sugarcane Yellow Leaf Disease (ScYLD).

B. Deep Learning for Image Classification

The study uses four deep learning models, including Inception v3, DenseNet 121, Resnet 50, and Xception, to accurately classify sugarcane aphid densities in images [18]. The objective is to develop an effective deep-learning model for pest density classification, with potential applications in mobile-based solutions and remote sensing technologies to automate pest monitoring across different crops. The outcomes of this research hold promise for improving the efficiency and precision of pest monitoring protocols in agriculture.

C. Image Preprocessing and Data Augmentation/ Enhancement Techniques

Research efforts have been directed towards disease classification in sugarcane plants [8], achieving impressive results such as an 86.53% accuracy using an ensemble deep learning model and introducing a valuable self-created database of sugarcane leaf diseases. Moreover, similar studies have focused on enhancing disease detection in sugarcane [19], emphasizing practicality for farmers. One such effort achieved a notable 96% accuracy by training a deep-learning neural network on thousands of leaf images, complemented by developing a user-friendly Android application for real-time disease prediction.

In addition, research focusing on the detection of plant diseases has extended beyond sugarcane [19, 20], addressing challenges farmers face across various crops. These studies emphasize the importance of automated disease recognition systems and offer solutions to the

difficulties associated with the accurate and timely identification of infections. They employ different deep-learning architectures and techniques to advance disease detection and classification. Deep learning's applications extend beyond sugarcane to encompass diverse crops. In tomato cultivation [21], the DbneAlexnet model achieves a 92.4% accuracy in classifying tomato plant leaf diseases, contributing to improved crop management. Deep learning models offer a promising solution for broader agriculture applications, especially for crop disease recognition [22, 23]. EfficientNet architecture demonstrates superior performance, with models achieving 99.91% accuracy and 98.42% precision on the PlantVillage dataset [22]. Srivastava *et al*. [24] introduces the PlantRefineDet strategy, combining object detection and classification to accurately localize and identify crop disorders, achieving a remarkable accuracy of 99.994%.

Furthermore, in sugarcane disease detection, deep learning frameworks utilizing various feature extractors and classifiers are evaluated, with the VGG-16 and Support Vector Machine (SVM) combination reaching an AUC of 90.2%, highlighting the efficacy of deep learning in disease identification [25]. Lastly, the potential of deep learning is harnessed to address the critical issue of crop disease recognition in uncontrolled field conditions [7]. In this study, models trained on a diverse sugarcane dataset achieved a top accuracy of 93.20% on realistic images captured under different conditions, demonstrating robustness and versatility.

Additionally, deep convolutional neural networks have proven effective in classifying various plant leaves, showcasing an average accuracy of 99.58% [23]. Automated models like the Optimal Mobile Network-based Convolutional Neural Network (OMN-CNN) [12] and Bayesian learning-enhanced deep learning [26] have also demonstrated their potential. Furthermore, innovative segmentation and feature extraction techniques have been developed for disease identification [27]. In the context of tomato leaf diseases, lightweight custom CNN models [28] and PDICNet models [29] have excelled with high accuracy.

Meanwhile, machine learning and deep learning for olive tree disease identification have opened new avenues, reaching an impressive accuracy of 96.14% [30]. Additionally, software solutions for automated disease identification [31] and novel techniques like quantum-behaved particle swarm optimization-based deep transfer learning [32] have further enriched this field. Moreover, computer vision techniques, particularly ResNet50, have enabled autonomous sugarcane lodging detection with high accuracy [33]. Spectral analysis and chemometrics offer a practical approach to differentiate between sucrose sources [34].

Finally, deep learning models have been employed for sugarcane aphid infestation classification, with Inception v3 and Xception yielding the best results [18]. These diverse contributions collectively underscore the significant progress made in leveraging deep learning for the precise and efficient classification of agricultural diseases, offering valuable insights for improving crop health and productivity.

III. Materials and Methods

A. Data Collection

To create the dataset for the sugarcane leaf images for disease classification, we used the following settings under Table I for capturing the images with a Canon EOS 600D Digital SLR camera with a sensor width, that depends on the model of the camera, of 22.3 mm (s).

TABLE I. Specification for Image Acquisition

Camera Setting	Range	Specifications Used
Mode	A-DEP/M/Av/Tv/P	A-DEP
Focal length (f) (Camera lens)	18–55 mm	40–55 mm
Aperture Setting (f-stop)	f/3.5 to f/6	f/4.5 to f/6
Lighting (ISO)	ISO 100–6400	ISO 100

Setting the camera's mode to A-DEP, Auto-Depth of Field, allows the researchers to capture the best aperture for the object and adjust the shutter speed accordingly. Adjusting the camera's setting with the focal length of 40–55 mm (f) allows capturing the length of the sugarcane leaf ranging from 15–30 cm (L) to ensure the view of the disease symptoms and the aperture setting(f-stop) from f/4.5 to f/6 that is accounted from the feature of the A-DEP mode under the natural sunlight, around 11 am to 3 pm, wherein it can differ such as the clouds making the lighting dim. As for the sensor width, it is 22.3 mm (s) which is needed for determining the distance on taking a photo of the sugarcane leaf using the formula below.

$$D = \frac{L \times f}{s} \tag{1}$$

The images were taken outdoors under natural sunlight using ISO 100 to prevent overexposure and maintain sharpness in sunny conditions (Fig. 6). Each photo quality was captured at 3456×5184 pixels resolution and 72 dpi, however, varying in shutter speed from slow to fast (1/500–1/800 s), accounting for the aperture setting, ensuring detailed visual information on the disease for classification. Manuela Samaco conducted visual validations. Table II represents the number of images used as the dataset for the Mobilenetv2 + Prototypical Network model.

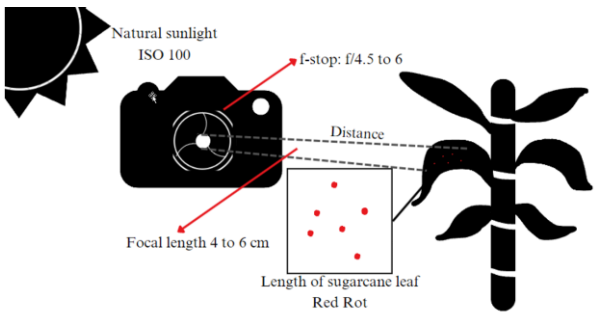

Fig. 6. Visual representation of the setup.

TABLE II. SUGARCANE DATA DISTRIBUTION

Sugarcane Disease	No. of Images	Percentage (%)
Ring Spot	30	16.67
Rust	30	16.67
Mosaic	30	16.67
Yellow Leaf Disease	30	16.67
Red Rot	30	16.67
Healthy	30	16.67
Total	180	100

Fig. 7 provides representative samples of six categories in the dataset: Ring Spot, Rust, Mosaic, Yellow Leaf Disease, Red Rot, and Healthy leaves. Each example captures characteristic symptoms that differentiate one class from another, aiding in the disease classification process. These samples highlight the visual diversity within the dataset and demonstrate the unique challenges in differentiating between similar symptoms across disease classes.

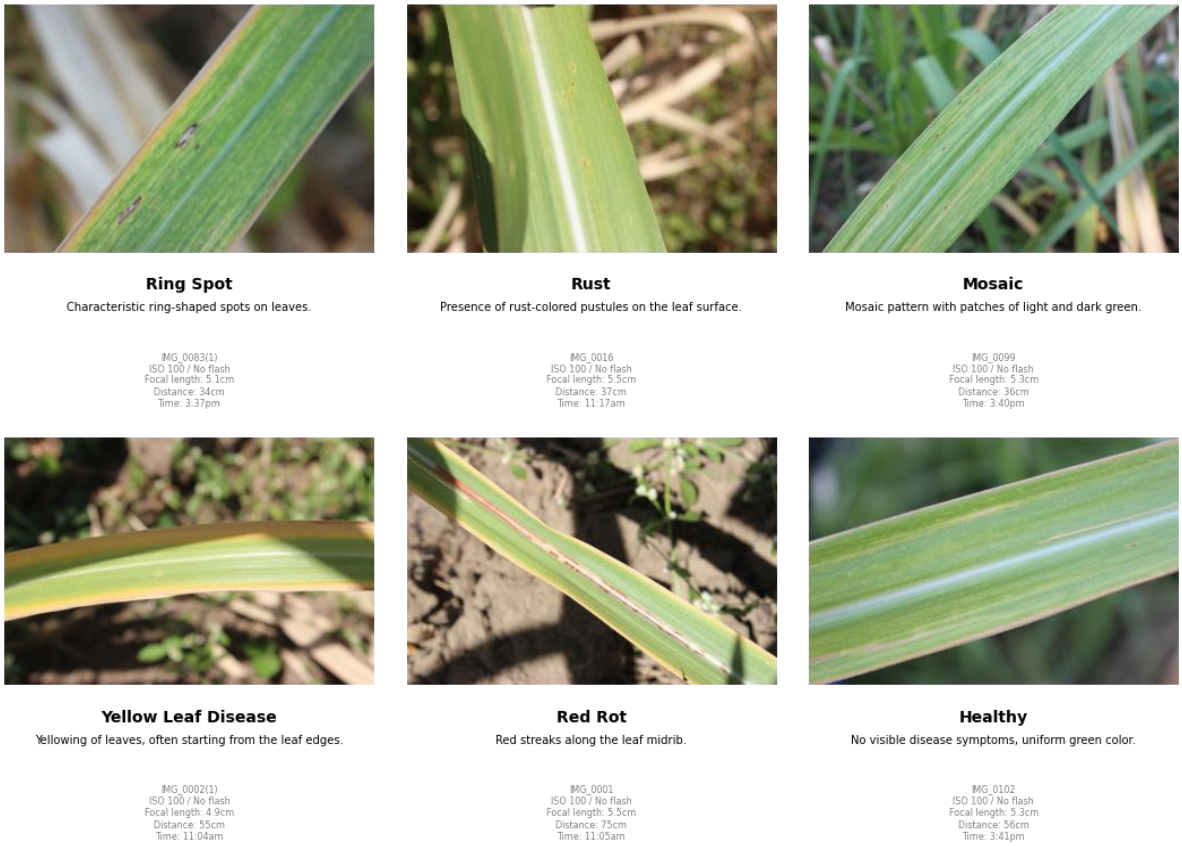

Fig. 7. Example of each class.

B. Data Preprocessing

To ensure dataset quality and consistency, the following preprocessing steps will be carried out:

Image Resizing—The original images were shrunk to 96×96 pixels. This was done to lessen the computational burden and preserve important features for disease classification. Fig. 8 displays the side-by-side comparison of the original and downscaled images.

Normalization—The pixel values in each image were scaled to fall between (0, 1). This was achieved by dividing pixel values by 255, which helped to stabilize the training process and enabled the model to converge more successfully.

Class Labeling—Every picture inside the collection was given a name based on the disease class it belongs to. Numerical values were used to encode the labels to meet the output criteria of the model.

C. Conceptual Framework

This framework is the foundation for a thorough investigation of using a hybrid model for sugarcane disease classification.

Fig. 9 outlines the framework of the prototypical network, which hinges on the feature extractor that will utilize a convolutional neural network to create the Prototypes. From which the feature of the query can depict the class probabilities using the Euclidean distance.

D. Model Development

1) Few-shot learning techniques

Few-shot learning remains a cornerstone of this research methodology, emphasizing the model's ability to generalize from a limited dataset and make accurate predictions for unseen data. In the context of sugarcane disease classification, Few-Shot Learning techniques will enable the model to effectively identify and classify different types of sugarcane diseases using only a small number of labeled examples.

Fig. 8. Original and downscaled images side by side.

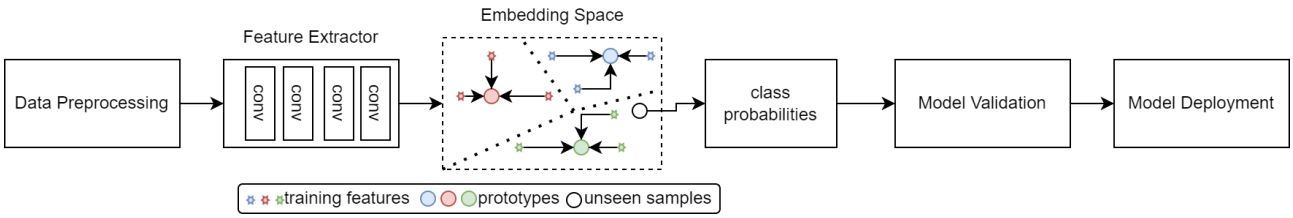

Fig. 9. Conceptual framework.

2) Few-shot problem definition

In machine learning, it is often to learn from experience E (datasets) to classes of task T (object recognition), with the quality of its performance evaluated by P (accuracy) [35]. Few-shot learning represents a specific category of machine learning where T is provided by the lacking information given by E.

$$ED_{support} = \{(x_i, y_i)\}_{i=1}^{I=N\times K} \tag{2}$$

$$ED_{query} = \{(x_i)\}_{j=1}^{J=N\times Q} \tag{3}$$

In this equation for few-shot learning, ED is separated into $ED_{support}$ and ED_{query} shown above, I is commonly a small integer that comes between 1–5. On the first equation, $ED_{support}$ is based on N categories with K samples per category. While ED_{query} contains samples from the same N category and Q samples per category. The purpose of this equation is to classify ED_{query} into N categories based on the limited supervised data from $ED_{support}$ [36].

3) MobileNet

Fig. 10 shows a diagram of MobileNetV1 and MobileNetV2's Convolutional Blocks. MobileNetV1 is made up of two levels. By applying a single convolutional filter to every input channel, the depth wise convolution in the first layer carries out lightweight filtering. A 1×1 convolution, referred to as a pointwise convolution, makes up the second layer. It computes linear combinations of the input channels to produce new features.

Table III showcases the general layers of MobileNetV2. MobileNetV2 consists of two different kinds of blocks. The first is a stride of one while the other is a residual block. An additional one is a shrinking block comprised of

a stride of two. There are three levels for both varieties of blocks. The first layer will then be used in a 1×1 convolution, known as ReLU6. The depth-wise convolution is the second layer. Another 1×1 convolution, without any non-linearity, makes up the third layer. If ReLU is used again, the deep networks have the power of a linear classifier solely on the non-zero volume portion of the output domain. Moreover, there is an expansion factor t. For every primary experiment, $t = 6$. If the input had 64 channels, the internal output would have $64 \times t = 64 \times 6 = 384$ channels.

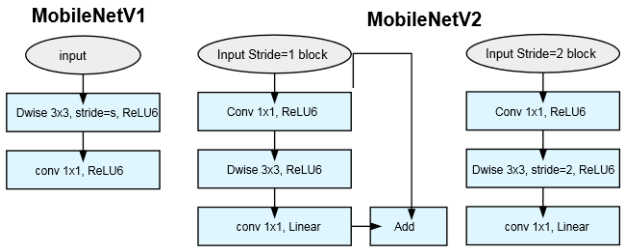

Fig. 10. MobileNet convolutional blocks.

TABLE III. MOBILENETV2 LAYERS

Input	Operator	Output
$h \times w \times k$	1×1 conv2d, ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3×3 dwise s = s, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times tk$	linear 1×1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

TABLE IV. MOBILENETV2 OVERALL ARCHITECTURE

Input	Operator	t	C	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	35	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1×1	-	1280	1	1
$7^2 \times 1280$	avgpool 7×7	-	-	1	-
1×1×1280	Conv2 1×1	-	k	1	

Table IV displays the overall architecture of MobileNetV2. The variables t, C, n, and s represent the expansion factor, repeating number, and stride. Spatial convolution is performed with 3×3 kernels. The leading network (width multiplier 1224×224) typically employs 3.4 million parameters and 300 million multiply-adds in computing. (MobileNetV1 introduces the width multiplier.) Further investigation is done into the trade-offs between performance for width multipliers ranging from 0.35 to 1.4 and input resolutions from 96 to 224. When the model size varies from 1.7M to 6.9M parameters, the computational cost of the network can reach 585M MAdds. 16 GPUs with a batch size of 96 are used to train the network.

4) Parameters

Table V displays all the parameters used in the MobilenetV2 Prototypical Network-Pipeline which will be explained here.

TABLE V. PARAMETER SETTINGS FOR MOBILENETV2 PROTOTYPICAL NETWORK PIPELINE

Parameter	Details
Input Shape	(96, 96, 3)
Embedding Dimension	64
Batch Size	8
Epochs	10
Optimizer	Adam

The embedding dimensions define how an input image is encoded as a feature vector. The output of the MobileNetV2 model is a vector of this dimension, capturing the most important features of the image while discarding irrelevant details. A very high embedding dimension would cause the model to overfit the training data because it would have too much capacity to memorize the limited data. A lower embedding dimension, like 64, ensures that the model generalizes better by compressing the information in the image without capturing too many unnecessary details.

Adam, also known as Adaptive Moment Estimation, is an optimizer that combines the advantages of AdaGrad and RMSProp. Adam can adapt to various feature scales during training since it can handle sparse gradients and dynamically change the learning rate for each parameter independently. Rapid convergence was essential because it allowed the models to learn effectively in fewer epochs, which shortened the training period. Adam adjusts the learning rates dynamically by calculating the gradients' first (mean) and second (uncentered variance) moments, which aids the model in navigating intricate features and enhances its learning procedure.

5) Training and validation

Data Preparation: Images are loaded from a specified directory and resized to 96×96 pixels. They are then normalized by dividing by 255.0 to bring the pixel values into the [0, 1] range. The labels are encoded into integers to be compatible with the model's loss function.

Model Definition: A pre-trained MobileNetV2 model is used as the base encoder, with the top classification layers removed (include_top = False). To stop its weight from changing while training, this base model is frozen. Then a custom dense layer is added to the encoder to produce embeddings of size 64.

Prototypical Network Training: The training process involves computing prototypes (centroids) for each class using the mean of the feature vectors (embeddings) of samples from that class. The model computes the distance between each sample's embedding and the prototypes during each epoch. The model then calculates a custom loss (prototypical_loss) based on these distances and adjusts its weight using gradient descent. Training accuracy is computed per batch to monitor how well the model learns.

Validation: The model evaluates the validation set after each epoch. The validation loss and accuracy are computed using the same prototypical network approach. This helps to track the model's performance on new data that is not part of the training.

Plotting: The training and validation accuracies and losses are plotted to help analyze and visualize the learning progress that the model made over epochs.

Training Setup: The researchers trained four models: MobileNetV2, MobileNetV2 combined with Prototypical Networks, ResNet50, and VGG. Using a limited dataset, these models were chosen to compare performance on the classification of six sugarcane diseases. The dataset was preprocessed by resizing it to 96×96 pixels, and the images were normalized. The MobileNetV2 + Prototypical Network model was trained using the Adam optimizer with a fixed learning rate of 0.001 over 10 epochs. Each batch consisted of 8 samples. The pre-trained weights were obtained on ImageNet using the module "tf.keras.applications" from the TensorFlow library (version 2.17.0). The MobileNetV2 backbone of the Prototypical Network was initialized with pre-trained weights from ImageNet, allowing the model to leverage general image features for improved classification under few-shot conditions.

IV. Result and Discussion

The researchers present performance metrics and loss curves for the analysis to show the performance of all *models.*

A. Performance Parameters

The efficiency of various deep-learning models for the classification of sugarcane diseases was evaluated in this study using metrics precision, recall, accuracy, F1-Score, and PR-AUC. To shed light on the advantages and disadvantages of each strategy, the findings for these measures across models MobileNetV2, Prototypical Network with MobileNetV2, ResNet50, and VGG16 are examined.

Key Findings and Model Selection: MobileNetV2, with its lightweight architecture and efficient use of linear bottlenecks and depth-wise separable convolutions, offered competitive performance and was suitable for real-time applications. The integration of a Prototypical Network further enhanced its accuracy and robustness, particularly in few-shot learning scenarios where training data was limited. ResNet50, with its residual connections, excelled at capturing intricate patterns and detecting subtle variations, but its high computational cost and larger memory usage limited its efficiency. On the other hand, VGG16 struggled across all metrics, with slower inference times and increased processing costs due to its lack of advanced architectural features, making it less effective compared to the other models.

Model Limitations and Future Considerations: While the Prototypical Network with MobileNetV2 provided the best overall performance, certain limitations need to be acknowledged:

- Lower Accuracy on Complex Images: The baseline MobileNetV2 and its Prototypical Network variant may still struggle with high-resolution images or those containing intricate patterns.
- Scalability Concerns: While effective for simpler tasks or smaller datasets, these models may not scale well to very large datasets or more complex tasks without significant modifications.
- Moderate Computational Requirements: Even with reduced parameters, models like ResNet50 exhibit high computational costs and may not be suitable for environments with limited hardware capabilities.

Moving forward, the choice of input size, model architecture, and training strategies should be carefully balanced to optimize computational efficiency and performance metrics.

B. Performance Analysis

Table VI compares the performances of all the models. The MobileNetV2 + Prototypical Network model achieved a validation accuracy of 80.56%, indicating that combining MobileNetV2's feature extraction with Prototypical Networks enhances classification performance, particularly with limited data. Precision (0.8413) and recall (0.8056) are well-balanced, reflecting the model's ability to handle both false positives and false negatives effectively. The F1-Score of 0.8019, calculated as the mean of precision and recall, indicates consistent performance across classes. The PR-AUC (0.8091), representing the area under the precision-recall curve, reflects the model's effectiveness in distinguishing between classes, even when features overlap.

The standalone MobileNetV2 model achieved a validation accuracy of 61.11%, lower than the Prototypical Network variant, but demonstrating MobileNetV2's capability in feature extraction for classification tasks. Precision (0.6153) and recall (0.6111) suggest some challenges in distinguishing between classes with similar features. The F1-Score (0.5893) reflects the model's balance between precision and recall, though it is lower than that of the Prototypical Network variant, indicating a decrease in prediction consistency. PR-AUC (0.7787) shows moderate performance in separate classes.

TABLE VI. Performance Analysis

Model	Validation Accuracy	Precision	Recall	F1-Score	PR-AUC
MobileNetV2+Prototypical Network	0.8056	0.8413	0.8056	0.8019	0.8091
MobileNetV2	0.6111	0.6153	0.6111	0.5893	0.7787
ResNet50	0.2500	0.2212	0.2500	0.2024	0.2600
VGG	0.1667	0.0278	0.1667	0.0476	0.5833

ResNet50 had a validation accuracy of 25%, indicating difficulties in generalization and likely overfitting, as seen in the confusion matrix analysis. Precision (0.2212) and F1-Score (0.2024) are low, reflecting frequent misclassifications across classes. The recall (0.2500) aligns with the accuracy, showing that the model identified only a small fraction of true positive cases. The PR-AUC (0.2600) further indicates ResNet50's limited ability to separate classes, suggesting that its depth and complexity may not suit this dataset with limited, balanced samples.

VGG achieved a validation accuracy of 16.67%, similar to ResNet50. Precision (0.0278) and F1-Score (0.0476) indicate challenges in correct classification, with significant misclassification observed in the confusion matrix. Like ResNet50, VGG overfit the training data and predicted a single dominant class for all samples in the test set. The PR-AUC (0.5833) suggests a limited capacity to distinguish between classes, though it performed slightly better than ResNet50. These results indicate that VGG's deeper architecture, like ResNet50, is not suitable for datasets with limited examples.

- C: Correct Classification;
- M: Misclassified as another class (Table VII).

MobileNetV2 achieved 100% accuracy for Healthy (Class 6) but misclassified Rust (Class 2) and Mosaic (Class 3) as Healthy, suggesting overlapping features. It performed well in Red Rot (Class 5), correctly predicting 5 out of 6 samples. The Prototypical Network addition improved classification, eliminating misclassifications for Healthy and Yellow Leaf Disease (Class 4), but Mosaic remained difficult to separate, often confused with Rust and Red Rot, as shown in Fig. 11.

ResNet50 and VGG16 suffered from severe overfitting, predicting Red Rot (Class 5) and Rust (Class 2) for all samples, respectively, indicating an inability to generalize with the limited dataset. The confusion matrix highlights distinct misclassification trends, with Ring Spot (Class 1) lacking strong features and Mosaic frequently misclassified due to feature overlap.

False positives were common for Healthy, while Rust and Mosaic had high false-negative rates, suggesting inadequate feature representation. Adjusting classification thresholds or loss functions could help mitigate these issues. Overall, MobileNetV2 + Prototypical Network provided a more balanced classification, reducing extreme misclassifications and improving class separation compared to standalone MobileNetV2.

TABLE VII. CONFUSION MATRIX ANALYSIS

Model	Class 1 (Ring Spot)	Class 2 (Rust)	Class 3 (Mosaic)	Class 4 (Yellow Leaf Disease)	Class 5 (Red Rot)	Class 6 (Healthy)
MobileNetV2	2 (C)/4 (M)	2 (C)/4 (M)	3 (C)/3 (M)	4 (C)/2 (M)	5 (C)/1 (M)	6 (C)/0 (M)
MobileNetV2+Prototypical	5 (C)/1 (M)	3 (C)/3 (M)	5 (C)/1 (M)	5 (C)/1 (M)	5 (C)/1 (M)	6 (C)
ResNet50	1 (C)/5 (M)	2 (C)/4 (M)	6 (M)	2 (C)/4 (M)	4 (C)/2 (M)	6 (M)
VGG	6 (M)	6 (M)	6 (M)	6 (M)	6 (C)	6 (M)

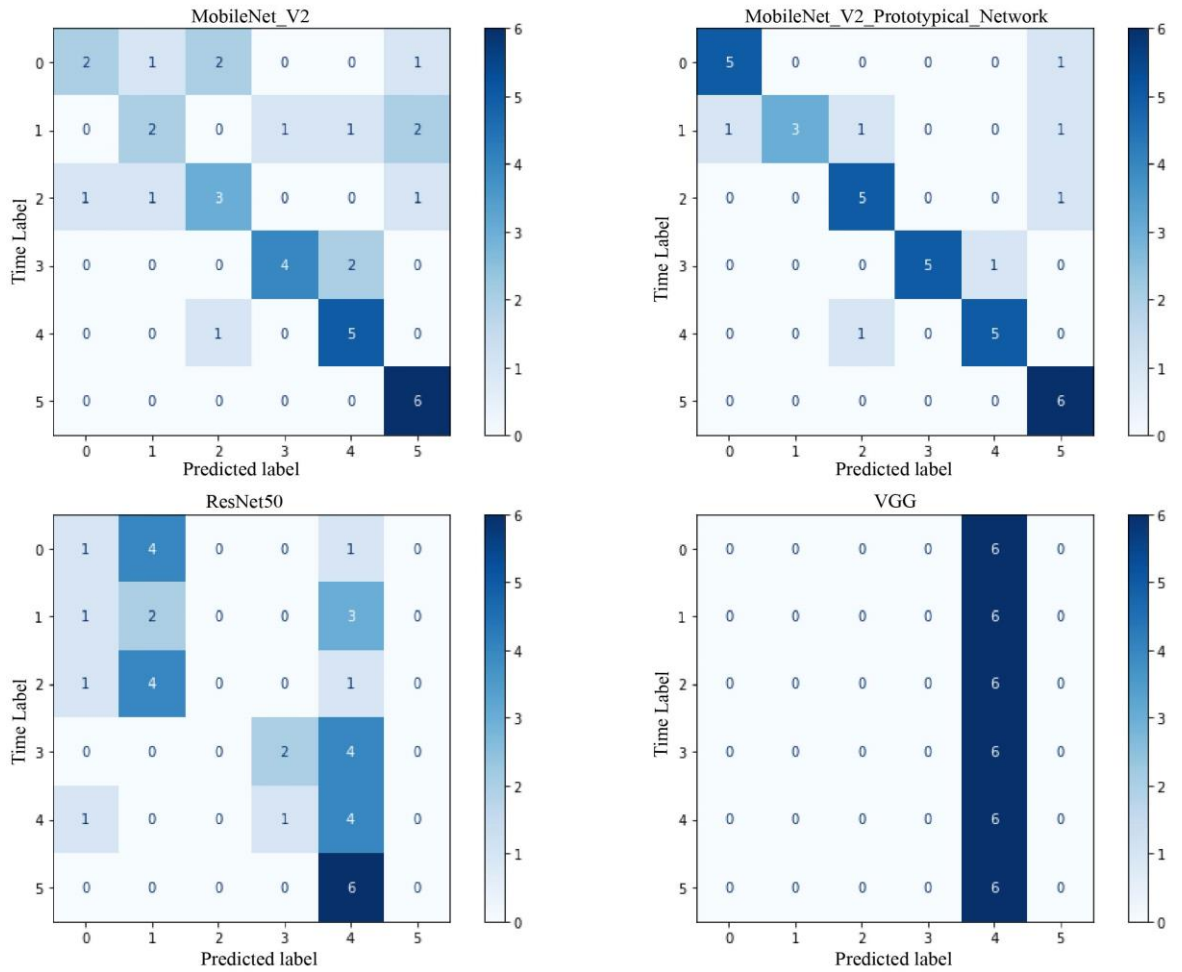

Fig. 11. Confusion matrix of all models.

C. Visualized Analysis

Fig. 12 illustrates the distinct behaviors of different models in sugarcane disease classification. MobileNetV2 achieves 67% validation accuracy but shows widening gaps between training and validation loss, indicating overfitting. In contrast, the MobileNetV2 + Prototypical Network model reaches 80% validation accuracy with a m ore stable loss curve, suggesting better generalization.

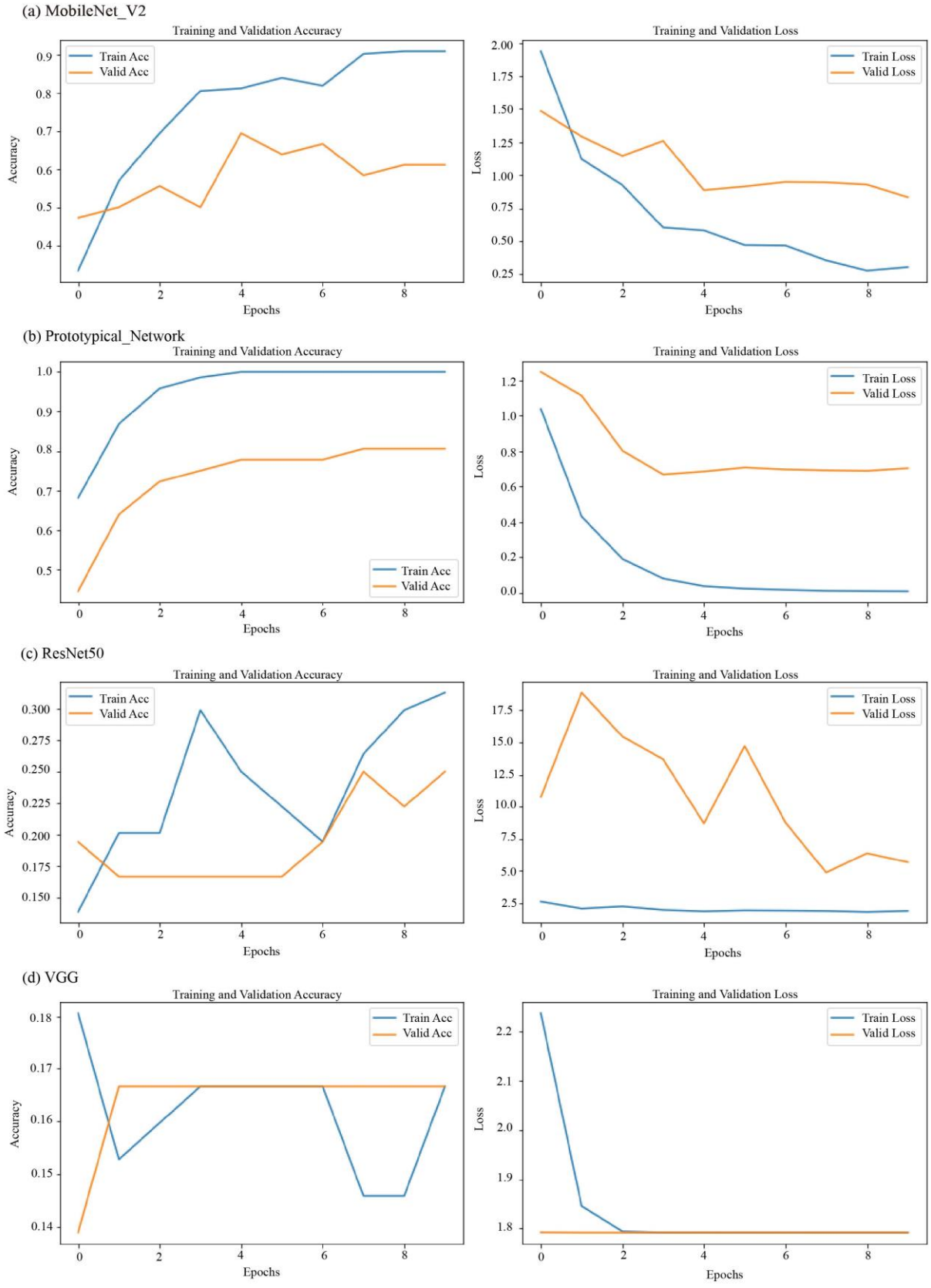

Fig. 12. Training, validation loss, and accuracy of all models.

ResNet50 exhibits high fluctuations in accuracy, settling at only 25%, while VGG16 performs the worst at 16.67%, both struggling to capture meaningful disease features. Their erratic loss patterns highlight difficulties in learning from the dataset, reinforcing their unsuitability for this task.

MobileNetV2 + Prototypical Network consistently achieves higher accuracy and steadier convergence, unlike ResNet50 and VGG16, which show weak generalization. The model confidently classifies Rust and Ring Spot (>90%) but struggles at lower confidence (~40%), misclassifying Red Rot as Mosaic and Rust. Healthy samples are sometimes labeled as Rust, likely due to subtle artifacts like shadows or blemishes. These findings highlight the need for improved feature differentiation, especially for diseases with overlapping symptoms.

The 2D t-SNE (t-Distributed Stochastic Neighbor Embedding) visualization in Fig. 13 reduces the high-dimensional embeddings generated by the MobileNet_V2 Prototypical Network to two dimensions, providing a clearer view of the relative distances between disease classes. t-SNE preserves local relationships, which allows for observing clustering patterns and areas of overlap between disease classes. Each point in Fig. 13 represents an image sample from the validation set, color-coded by disease class. Misclassified samples are marked with red "X"s to show where the model's predictions differ from actual labels.

The distribution of clusters reflects how the model interprets similarities and differences between classes in feature space. Distinct clusters, such as for Healthy and Red Rot, indicate separability between these classes. Areas of overlap, particularly among Mosaic, Yellow Leaf Disease, and Rust, suggest shared visual features that present challenges for classification. Misclassifications are mainly located within these overlapping areas, indicating limitations in feature separation. This visualization points to the need for further improvements in class separability, especially for classes with similar visual characteristics.

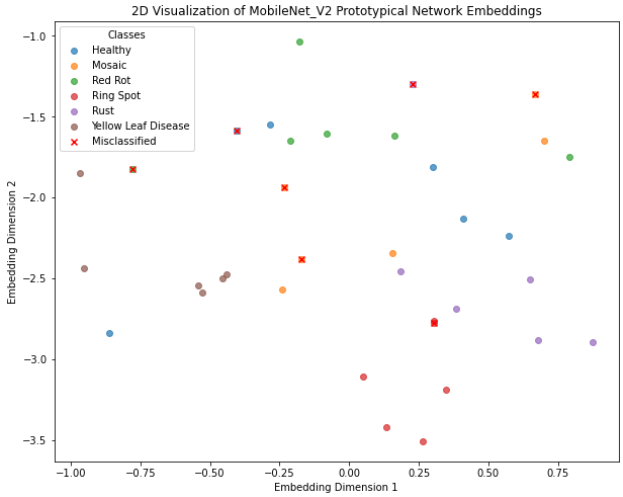

Fig. 13. 2D t-SNE visualization of MobileNetV2-Prototypical network embeddings for sugarcane disease classes.

D. *Analysis of Error Sources*

The Prototypical Network, combined with MobileNet_V2, demonstrated significant challenges in distinguishing between visually similar sugarcane diseases. Although it classified Red Rot (Class 5) with perfect accuracy, it showed notable misclassifications between classes such as Ring Spot (Class 1) and Rust (Class 2), and between Mosaic (Class 3) and Yellow Leaf Disease. These errors, as reflected in the confusion matrix, can be attributed to overlapping visual symptoms that are difficult for the model to differentiate. Furthermore, environmental factors during data collection, such as inconsistent lighting and background variations, affected the clarity of key visual features needed for accurate classification.

The model's performance was influenced by several methodological choices that contributed to these misclassifications. The preprocessing stage of downscaling images to 96×96 pixels significantly reduced subtle yet critical color and texture details, particularly for diseases like mosaic and Yellow Leaf Disease. Future work will explore less aggressive downscaling or multi-scale processing to preserve finer visual cues and enhance the model's ability to differentiate between visually similar classes. Additionally, the Prototypical Network's reliance on Euclidean distance as a similarity metric treats all dimensions equally, potentially overlooking nuanced relationships crucial for distinguishing between classes with minor visual distinctions. Investigating alternative similarity metrics, such as cosine distance, could improve class separability, especially for overlapping clusters observed in t-SNE. Lastly, the model's current architecture, particularly MobileNetV2's Depthwise Conv2D layers, constrained spatial filtering. Incorporating attention mechanisms could enable the model to dynamically focus on the most discriminative regions of the image, thereby improving feature extraction and reducing misclassifications

Episodic training, a core principle of many few-shot learning protocols, is designed to simulate the few-shot inference scenario during training. However, this study adopted a strategy of computing full prototypes for each class per epoch. This approach, while a reasonable trade-off for the initial exploration, deviates from strict episodic sampling. This methodological choice may impact the model's generalization capabilities to truly novel few-shot tasks. While the t-SNE visualization (Fig. 13) offers insights into the clustering of embeddings, the current analysis lacks advanced interpretability tools such as Grad-CAM. The absence of such visualizations limits the ability to precisely understand which specific image regions or features the model prioritizes for classification.

Future work could investigate the effects of incorporating episodic training to better align with advanced few-shot learning methodologies and potentially enhance the model's robustness. Furthermore, a parallel priority will be the integration of explainability techniques to provide clearer insights into model decision-making, moving beyond just performance metrics to truly understanding the model's behavior.

V. Conclusion

In this study, we conducted a comparative analysis of models for sugarcane disease classification, integrating few-shot learning techniques with deep learning. Four models were evaluated: MobileNetV2, MobileNetV2 with a Prototypical Network, ResNet50, and VGG16. Results demonstrated that combining MobileNetV2 with a Prototypical Network achieved the highest accuracy (80.56%) and balanced performance across all metrics. The study confirms that few-shot learning, when paired with a lightweight architecture, effectively addresses data scarcity in agricultural disease classification while maintaining computational efficiency. It is important to restate that these results are exploratory and should not yet be generalized to broader contexts without further validation, given the limited and geographically specific nature of the dataset.

Future research should explore optimizing image resolution and multi-scale feature extraction to improve classification performance and explore transfer learning and domain adaptation techniques to enhance model generalizability across different geographic regions and sugarcane cultivars. Implementing focal loss could help address misclassification in visually similar disease classes. Moreover, incorporating more robust validation strategies, including K-fold cross-validation and statistical significance testing, to ensure the robustness and reliability of the model's performance across varied data subsets. To further enhance the statistical robustness of future evaluations, reporting performance metrics with associated confidence intervals or standard deviations is recommended. This would provide a more comprehensive understanding of the variability and reliability of the model's performance across multiple runs or different data splits. Additionally, incorporating explainability techniques such as Grad-CAM may provide insights into model decision-making, aiding in refining feature extraction for challenging cases. False positives or negatives may lead to incorrect treatment decisions. We recommend future work to focus on model calibration and integration with expert systems to reduce this risk.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Miguel Abhram A. Agno: In charge of handling the experimental setup, writing the related literature, methodology, analyzing the data, developing the prototype, and discussing the results. Paul Jericho M. Vale: Collaborated in the design and implementation of the experimental setup, co-authored the literature review and methodology, participated in data preprocessing, led in the prototype development, contributed to result analysis, refinement of the manuscript, and visual presentation of findings. John Paul Q. Tomas: Served as the research adviser, providing guidance throughout the study, offering critical feedback on the methodology and analysis, and overseeing the overall progress and direction of the research. Marineil C. Gomez: Supervision, review and editing, funding acquisition. All authors had approved the final version.

Acknowledgment

The researchers would like to express their sincere gratitude to the following individuals and groups who contributed to the successful completion of this thesis:

To our thesis adviser, Prof. John Paul Q. Tomas, for their consistent guidance, valuable insights, and support throughout every stage of this research.

To the faculty of the School of Information Technology (SoIT), including the CS Program Chair, Prof. John Paul Q. Tomas, and the Technical Assistant to the Dean, Ms. Sally Zara, and Dean Ariel Kelly D. Balan, for their support and encouragement, as well as for providing us with the resources necessary to complete this study.

To the Plant Pathologist, Manuela Samaco, and Sugarcane Farmers for providing the sugarcane disease dataset used in this research, and for their collaboration in the data collection process.

Lastly, to Mapúa University, for providing us with a rigorous Computer Science program that has honed our skills in artificial intelligence and deep learning, enabling us to apply these concepts effectively in solving real-world challenges.

References

[1] R. Akhter and S. A. Sofi, "Precision agriculture using IoT data analytics and machine learning," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 8, pp. 5602–5618, 2022. doi: 10.1016/j.jksuci.2021.05.013

[2] J. G. A. Garcia, "A review of the main challenges in automatic plant disease identification based on visible range images," *Biosyst. Eng.*, vol. 144, pp. 52–60, 2016. doi: 10.1016/j.biosystemseng.2016.01.017

[3] A. Tripathi, U. Chourasia, P. Dixit, and V. Chang, "A survey: Plant disease detection using deep learning," *Int. J. Distrib. Syst. Technol.*, vol. 12, no. 3, pp. 1–26, 2021. doi: 10.4018/IJDST.2021070101

[4] S. P. Mohanty, D. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, 2016. doi: 10.3389/fpls.2016.01419

[5] S. A. Miller, F. D. Beed, and C. L. Harmon, "Plant disease diagnostic capabilities and networks," *Annu. Rev. Phytopathol.*, vol. 47, pp. 15–38, 2009. doi: 10.1146/annurev-phyto-080508-081743

[6] X. Li, X. Li, S. Zhang *et al*., "SLViT: Shuffle-convolution-based lightweight vision transformer for effective diagnosis of sugarcane leaf diseases," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 6, 101401, 2023. doi: 10.1016/j.jksuci.2022.09.013

[7] H. S. Malik *et al*., "Disease recognition in sugarcane crop using deep learning," *Advances in Intelligent Systems and Computing*, vol. 1133, pp. 189–206, 2020. doi: 10.1007/978-981-15-3514-7_17

[8] S. D. Daphal and S. Koli, "Enhancing sugarcane disease classification with ensemble deep learning: A comparative study with transfer learning techniques," *Heliyon*, vol. 9, no. 8, e18261, 2023. doi: 10.1016/j.heliyon.2023.e18261

[9] L. M. Tassis and R. A. Krohling, "Few-shot learning for biotic stress classification of coffee leaves," *Artif. Intell. Agric.*, vol. 6, pp. 55–67, 2022. doi: 10.1016/j.aiia.2022.04.001

[10] M. Li *et al*., "Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging," *npj Digit. Med.*, vol. 3, no. 1, 2020. doi: 10.1038/s41746-020-0255-1

[11] J. Lin *et al*., "Few-shot learning for the classification of intestinal tuberculosis and Crohn's disease on endoscopic images: A novel learn-to-learn framework," *Heliyon*, vol. 10, no. 4, e26559, 2024. doi: 10.1016/j.heliyon.2024.e26559

[12] G. García *et al*., "Circumpapillary OCT-focused hybrid learning for glaucoma grading using tailored prototypical neural networks,"

Artif. Intell. Med., vol. 118, 102132, 2021. doi: 10.1016/j.artmed.2021.102132

[13] B. Matias *et al*., "Transformers, convolutional neural networks, and few-shot learning for classification of histopathological images of oral cancer," *Expert Syst. Appl.*, vol. 241, 122418, 2024. doi: 10.1016/j.eswa.2023.122418

[14] G. H. Lu *et al*., "Sugarcane mosaic disease: Characteristics, identification and control," *Microorganisms*, vol. 9, no. 9, p. 1984, 2021. doi: 10.3390/microorganisms9091984

[15] M. I. Hossain *et al*., "Current and prospective strategies on detecting and managing colletotrichum falcatum causing red rot of sugarcane," *Agronomy*, vol. 10, no. 9, p. 1253, 2020. doi: 10.3390/agronomy10091253

[16] N. R. Laurel, R. L. De, J. S. Mendoza, M. Angelo, and F. M. Dela, "Identification of Epicoccum sorghinum and its effect on stalk sugar yield," *Sugar Tech.*, vol. 23, no. 6, pp. 1383–1392, 2021. doi: 10.1007/s12355-021-01017-y

[17] R. Manavalan, "Efficient detection of sugarcane diseases through intelligent approaches: A review," *Asian J. Res. Rev. Agric.*, vol. 3, no. 1, pp. 174–184, 2021.

[18] R. Karthickmanoj, J. Padmapriya, and T. Sasilatha, "A novel pixel replacement-based segmentation and double feature extraction techniques for efficient classification of plant leaf diseases," *Mater. Today Proc.*, vol. 47, pp. 2048–2052, 2021. doi: 10.1016/j.matpr.2021.04.416

[19] C. Sarkar, D. Gupta, U. C. Gupta, and B. B. Hazarika, "Leaf disease detection using machine learning and deep learning: Review and challenges," *Appl. Soft Comput.*, vol. 145, 110534, 2023. doi: 10.1016/j.asoc.2023.110534

[20] S. Dahiya, T. Gulati, and D. C. Gupta, "Performance analysis of deep learning architectures for plant leaves disease detection," *Meas. Sensors*, vol. 24, 100581, 2022. doi: 10.1016/j.measen.2022.100581

[21] A. Umamageswari, S. Deepa, and K. Raja, "An enhanced approach for leaf disease identification and classification using deep learning techniques," *Meas. Sensors*, vol. 24, 100568, 2022. doi: 10.1016/j.measen.2022.100568

[22] V. Tiwari, R. C. Joshi, and M. Dutta, "Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images," *Ecol. Inform.*, vol. 63, 101289, 2021. doi: 10.1016/j.ecoinf.2021.101289

[23] Y. Alqahtani, T. Nazir, A. Javed, F. Jeribi, and A. Tahir, "An improved deep learning approach for localization and recognition of plant leaf diseases," *Expert Syst. Appl.*, vol. 230, 120717, 2023. doi: 10.1016/j.eswa.2023.120717

[24] S. Srivastava, P. Kumar, N. Mohd, A. Singh, and F. S. Gill, "A novel deep learning framework approach for sugarcane disease detection," *SN Comput. Sci.*, vol. 1, no. 2, p. 87, 2020. doi: 10.1007/s42979-020-0094-9

[25] G. Sachdeva, P. Singh, and P. Kaur, "Plant leaf disease classification using deep convolutional neural network with Bayesian learning," *Mater. Today Proc.*, vol. 45, pp. 5584–5590, 2021. doi: 10.1016/j.matpr.2021.02.312

[26] S. Reddy, G. Varma, G. P. S. Varma, and R. L. Davuluri, "Resnet-based modified red deer optimization with DLCNN classifier for plant disease identification and classification," *Comput. Electr. Eng.*, vol. 105, 108492, 2023. doi: 10.1016/j.compeleceng.2022.108492

[27] M. Badiger and J. Mathew, "Tomato plant leaf disease segmentation and multiclass disease detection using hybrid optimization enabled deep learning," *J. Biotechnol.*, vol. 374, pp. 101–113, 2023. doi: 10.1016/j.jbiotec.2023.07.011

[28] M. Pathan, N. Patel, H. Yagnik, and M. Shah, "Artificial cognition for applications in smart agriculture: A comprehensive review," *Artif. Intell. Agric.*, vol. 4, pp. 81–95, 2020. doi: 10.1016/j.aiia.2020.06.001

[29] H. El Akhal, A. Yahya, N. Moussa, and A. El, "A novel approach for image-based olive leaf diseases classification using a deep hybrid model," *Ecol. Inform.*, vol. 77, 102276, 2023. doi: 10.1016/j.ecoinf.2023.102276

[30] T. Tamilvizhi, S. Rajendran, K. Anbazhagan, and R. Kalimuthu, "Quantum behaved particle swarm optimization-based deep transfer learning model for sugarcane leaf disease detection and classification," *Math. Problems Eng.*, vol. 2022, no. 1, 3452413, 2022. doi: 10.1155/2022/3452413

[31] R. Manavalan, "Efficient detection of sugarcane diseases through intelligent approaches: A review," *Asian J. Res. Rev. Agric.*, vol. 3, no. 1, pp. 174–184, 2021.

[32] R. U. Modi *et al*., "State-of-the-art computer vision techniques for automated sugarcane lodging classification," *Field Crops Res.*, vol. 291, 108797, 2023. doi: 10.1016/j.fcr.2022.108797

[33] H. Erikliođlu *et al*., "Classification and quantification of sucrose from sugar beet and sugarcane using optical spectroscopy and chemometrics," *J. Food Sci.*, vol. 88, no. 8, pp. 3274–3286, 2023. doi: 10.1111/1750-3841.16674

[34] I. Grijalva, B. J. Spiesman, and B. McCornack, "Image classification of sugarcane aphid density using deep convolutional neural networks," *Smart Agric. Technol.*, vol. 3, 100089, 2023. doi: 10.1016/j.atech.2022.100089

[35] T. Mitchell *et al*., "Machine learning," *Annu. Rev. Comput. Sci.*, vol. 4, no. 1, pp. 417–433, 1990. doi: 10.1146/annurev.cs.04.060190.002221

[36] Y. Chen, Z. Liu, H. Xu, T. Darrell, and X. Wang, "Meta-baseline: Exploring simple meta-learning for few-shot learning," in *Proc. Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 9062–9071.