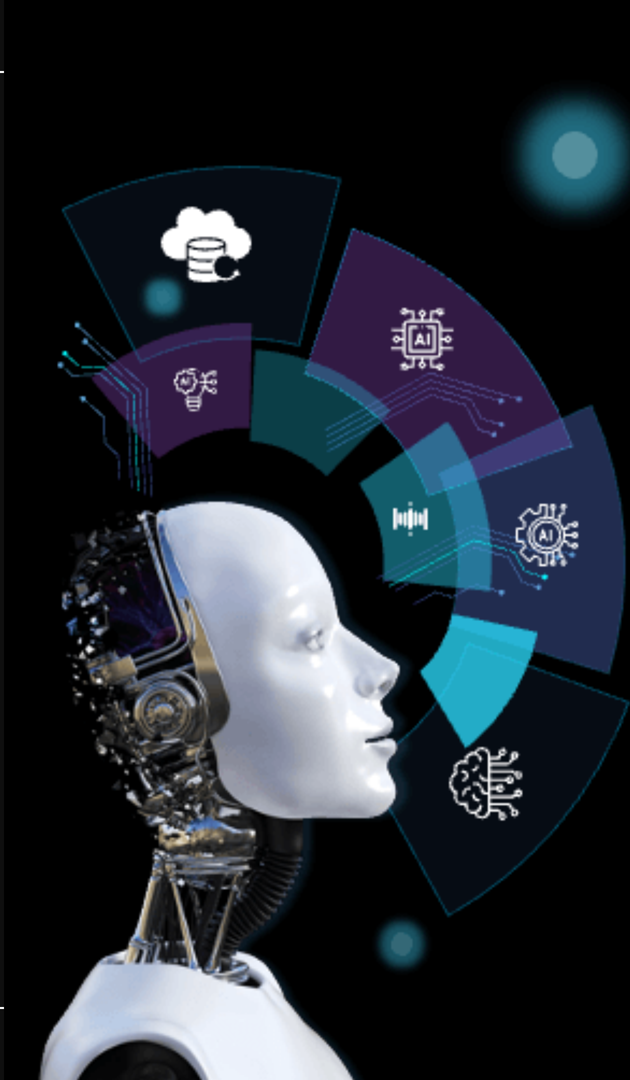
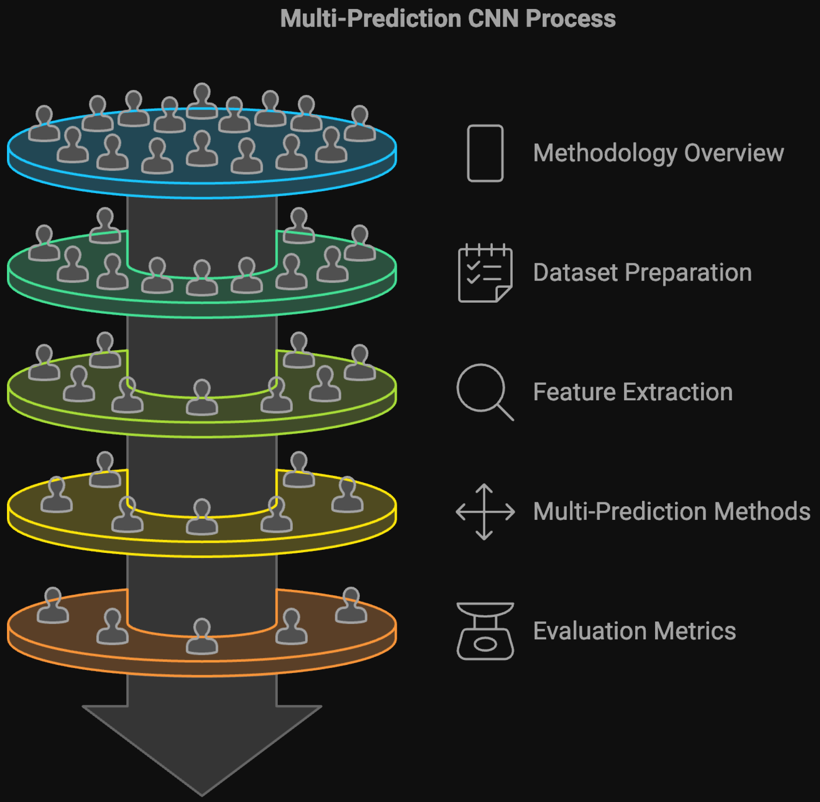
Multi-Prediction CNN Model

This document provides an overview of Multi-Model Convolutional Neural Networks (CNNs), which are advanced architectures that leverage multiple models to enhance performance in various tasks, particularly in image processing and computer vision. By integrating different CNN models, researchers and practitioners can achieve improved accuracy, robustness, and generalization capabilities in their applications.

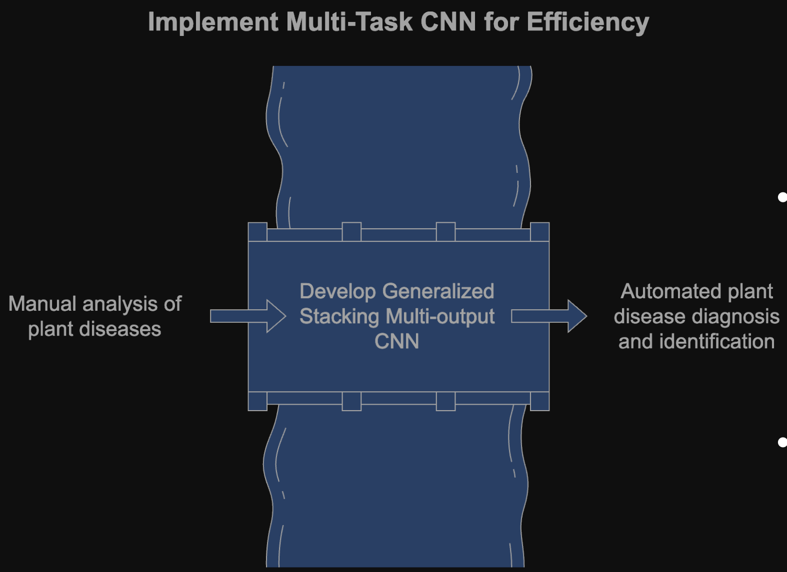


Topics

1. Introduction to Multi-Prediction CNN
2. Problem Statement
3. Methodology Overview
4. Dataset Preparation
5. CNN Backbones for Feature Extraction
6. Multi-Prediction Methods
7. Evaluation Metrics and Results
8. Conclusion and Insights



Introduction to Multi-Prediction CNN



Concept of Multi-Prediction CNNs

- **Deep Learning and Plant Pathology:-**Deep learning has significantly impacted many sectors, with agriculture and plant pathology among them. Traditional methods of identifying plant species and diagnosing diseases typically rely on manual analysis, which can be time-consuming and labor-intensive. With advancements in Convolutional Neural Networks (CNNs), image-based tasks have become more efficient and automated, particularly for applications like leaf disease detection and plant identification.

- **Leaf Images as Effective Input Data:-**Leaf images serve as effective input data for these tasks due to distinctive characteristics such as color, shape, and texture. These attributes make it possible to distinguish between plant species and detect disease markers. However, despite the advancements in CNNs, conventional models generally focus on single-task approaches, addressing either species identification or disease classification separately, rather than combining both in a unified framework.

- **Hypothesis and Multi-task CNN Model Proposal :-** This study hypothesizes that a multi-task CNN model could surpass single-task models by leveraging shared information across different tasks. The proposed Generalized Stacking Multi-output CNN (GSMo-CNN) is designed as a multi-output model capable of simultaneously handling both species identification and disease classification within a single framework. This combined approach could improve performance and provide a more efficient solution for plant pathology applications.

Critical Need for Efficient Tools

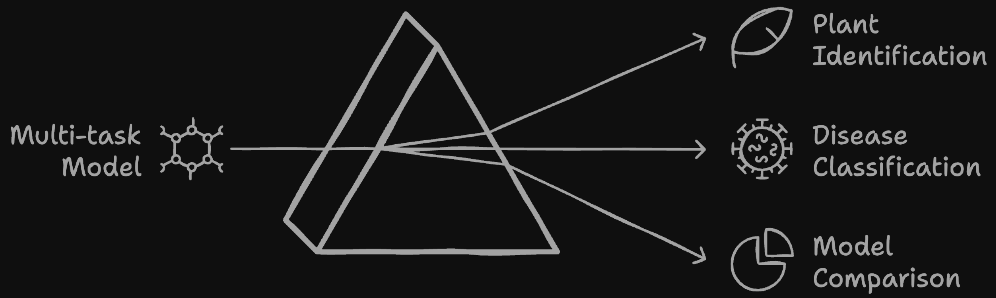
- Accurate diagnostics are essential for effective plant disease management.
- Current methods are often manual, labor-intensive, and time-consuming.
- Rapid identification of diseases can prevent crop loss and ensure food security.

Limitations of Current CNN Models

- Most existing CNN models are designed for single-task classification only.
- Single-task approaches overlook relationships between species and diseases.
- Complex interdependencies exist; some diseases are species-specific.

OBJECTIVE OF THE PRESENTATION

This paper develops a single CNN model that can carry out both tasks simultaneously: plant identification and disease classification. The hypothesis is that the multi-prediction model will be more effective and efficient than traditional single-task approaches. In pursuit of this goal, this study considers four different models of multi-task deep learning: multi-model, multi-label, multi-output, and multi-task, making a detailed comparison of each to identify the best-performing model.



Dataset Preparation

Methodology: Step-by-Step Pipeline for the Multi-Prediction CNN Model

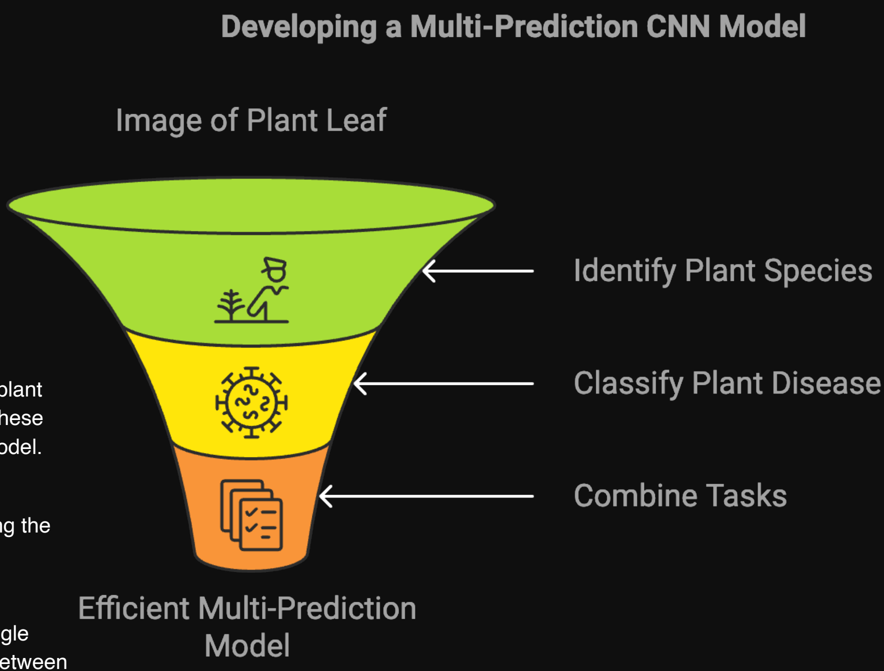
This subsection covers the methodology through which the study was conducted with step-by-step technical descriptions, so one can know how the design of pipeline identifies plant species and classifies plant disease from the images.

Background on the Problem and Approach

Objective: The objective is to develop a deep learning model that can recognize the species of a plant and classify any diseases found using only an image of a plant leaf. The challenge is to combine these tasks into a single, efficient model that performs both tasks at once, making it a multi-prediction model.

Traditional Approach Limitations : Historically, the separated model was either used in identifying the plant species or classifying the disease. Combining both tasks may result in efficiency, and better accuracy as well as saving some computational complexity.

Solution: The researchers propose a Generalised Stacking Multi-output CNN (GSMo-CNN), a single model with separate outputs for plant species and disease, but that allows learning to be shared between these tasks.



Pipeline Overview

The GSMo-CNN model uses a pipeline that involves several key stages: data preparation, CNN architecture selection, multi-task learning, and final output generation. Here's a breakdown of each step:

Step 1: Data Preparation

1. Collection of Datasets :-A set of images of leaves are collected from the dataset including Plant Village, Plant Leaves, and PlantDoc. In each of these images, information related to the species of plant is mentioned and, in case a disease affects the plant, information regarding the disease also accompanies.

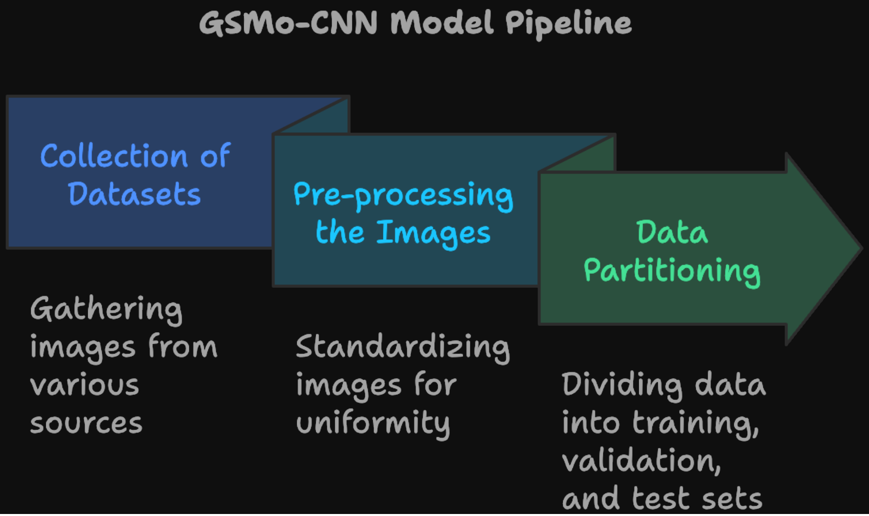
2. Pre-processing the Images :_The sizes of all the images are standardized into one dimension so that while being fed to the CNN, they become the same dimension, such as 256x256 pixels. Images are standardized in color and intensity to avoid the model from getting biased by variations in lighting or inconsistencies in the background.

3. Data Partitioning :_The dataset is divided into three segments

Training Set: 70% images, which are used for training the model.

Validation Set: 10% images, which are used for tuning model parameters.

Test Set: 20% images, which are used to evaluate final performance.



Plant Village

Description: Plant Village is one of the most popular datasets used for plant pathology research. There are 54,305 images of leaf samples in this dataset, which comprises images from 14 species of plants across 22 disease categories. The capturing conditions were controlled, like laboratory settings, with a simple plain homogeneous background and superior illumination, thus ensuring that the texture, color, and patterns of the leaves could be easily observable.

Species Diversity: Includes popular agricultural plants such as apple, blueberry, grape, and tomato.

Disease Variety: Covers common diseases like apple scab, grape black rot, and tomato early blight. Each species may have multiple disease categories, adding to the classification complexity.

Background Uniformity: All images have consistent lighting and background, making it easier for models to detect and classify features without the noise or interference present in natural settings.

Purpose: The primary usage of this dataset is for benchmarking; that is, to determine the ideal performance of a model in perfect conditions. It's uniform, thus providing CNNs with good accuracy as it has a very little amount of interference or noise due to its background.
Challenges: Although Plant Village is an excellent benchmark, the experimentally controlled setup of this cannot replicate the complexity of an agricultural field with varied lighting, background or similar environmental conditions.
Performance Benchmark: Under controlled conditions, the CNN models show the highest accuracy on the Plant Village dataset as against any other dataset. In that, the results also point out that InceptionV3 leads with *accuracy of 98.7%* with *0.97 F1-score*. Such results have reflected the capability of every model in feature extraction and classification when no noise exists.

Plant Doc

Description: PlantDoc is a collection of 2,598 images showcasing 13 plant species with 17 diseases. It was recorded in different settings, such as natural settings, variable illumination, and various orientations, considering the multiple real-world conditions that agricultural practitioners are likely to face.

Species and diseases representation: The dataset contains species such apple, cherry, corn and tomato, along with the diseases corresponding to these species.

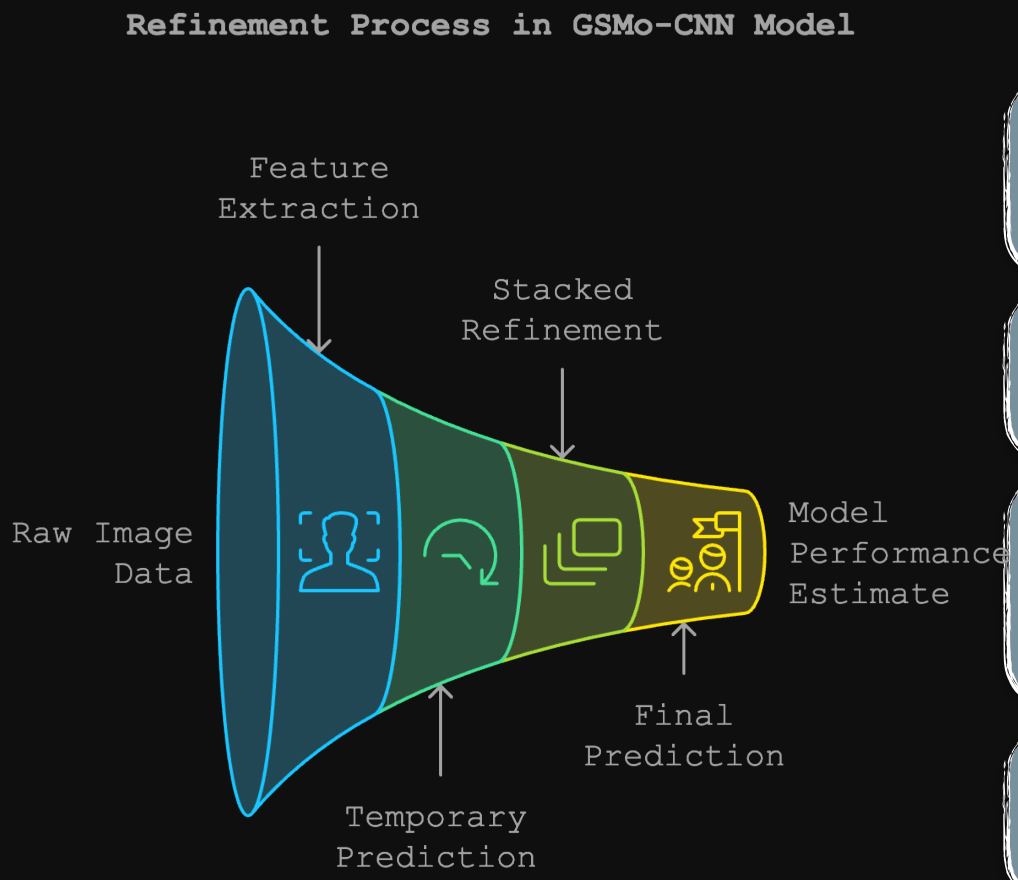
Complex Background: The images are having high background noise i.e soil, other plants and in some images farming tools which makes it difficult for CNNs to distinguish leaf features with accuracy.

Variable angles and variable lighting: The images in PlantDoc are of different angles and with different lighting than Plant Village, resulting in shadow effects, overexposed and underexposed in some images.

Another dataset suitable for evaluating the CNNs is PlantDoc, established as a testbed for the reliability and adaptiveness of the model under real-world condition. It gives a measure for the ability of CNNs to generalize out of otherwise perfect conditions seen in lab.

Issue: Environmental noise of PlantDoc such as hazy backgrounds and variable light conditions are the potential challenges for CNNs. Hence models should distinguish the leaf and its surroundings correctly and be able to adapt to the noise and be able to achieve a high classification accuracy.

Performance Benchmark: PlantDoc saw the lowest accuracy scores across all models, underscoring the dataset's complexity. InceptionV3 maintained the highest performance with 94.2% accuracy and a 0.93 F1-score, but all models experienced noticeable performance drops compared to their results on Plant Village and Plant Leaves.



Architecture of the GSMo-CNN Model

1. Feature Extraction Layers:
Convolution layers comprise the front end of GSMo-CNN which extracts salient information from images in the form of spots and textures and are used on all tasks. The elements allow for deriving the model's information on the leaf shape, the veins' patterns, and colour distribution.

2. Temporary Prediction Layers:
The model has distinct layers, which are used in generating *temporary predictions* for each task. For instance, the temporary prediction layer may first indicate that the leaf belongs to an "apple" species affected by "rust" disease.

3. Stacked Output Layers:
The temporary predictions will be added to the feature that has been extracted; thus, they are again fed into the model. For example, it improves upon its predictions because of being in the context of being in a relationship with other tasks or jobs. The model for instance might change its predicted disease when specific diseases in that species are not in the first predicted species it was able to find in the data.

4. Final Prediction Layers:
Species and disease are the outcomesof the stacked structure from which the final results come.
Only these last predictions are counted to estimate the performance of the model.

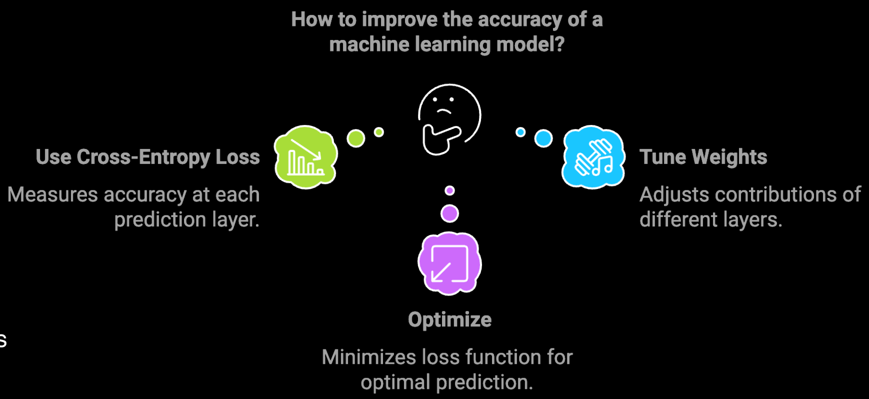
Training the GSMo-CNN Model

1. Loss Functions: At each level of prediction, a separate *cross-entropy loss function* measures the accuracy of prediction in each layer.

The model reduces these losses in order to increase accuracy at each step of the prediction pipeline.

2.Weight Tuning: The model has four balancing weights that are used to adjust how much contributions of different layers are being made. This brings the model to decide how often should it rely on temporary predictions versus final predictions.

3. Optimization: It runs several cycles with the update of various weights and minimizing loss function for optimal prediction ability employing accuracy and F1 scores as measures.



Inference and Prediction

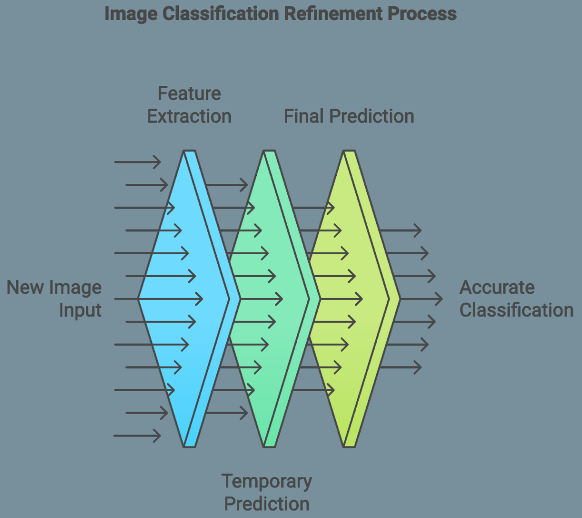
When a new image is fed into the GSMo-CNN model during inference (testing):

1.Initial Feature Extraction: The CNN extracts features from the image, just as it did during training.

2.Temporary Prediction: Temporary predictions for both tasks are made, guiding the model on possible species and disease categories.

3. Final Prediction Using Stacked Layers: The final prediction layers generate the ultimate species and disease classification by refining the temporary predictions.

Only these final predictions are used in the results, as they represent the model's most accurate assessment.



Evaluation Metrics

To assess model performance, the researchers used the following metrics:

1. **Accuracy:** Measures the percentage of correct predictions out of the total predictions.

2. **F1-score:** Balances precision (how many of the predicted positive cases were true) and recall (how many true cases were identified by the model).

3. **False Positive Rate (FPR):** Assesses the number of incorrect positive predictions, helping identify any bias in the model.

Which metric to use for assessing model performance?

Accuracy

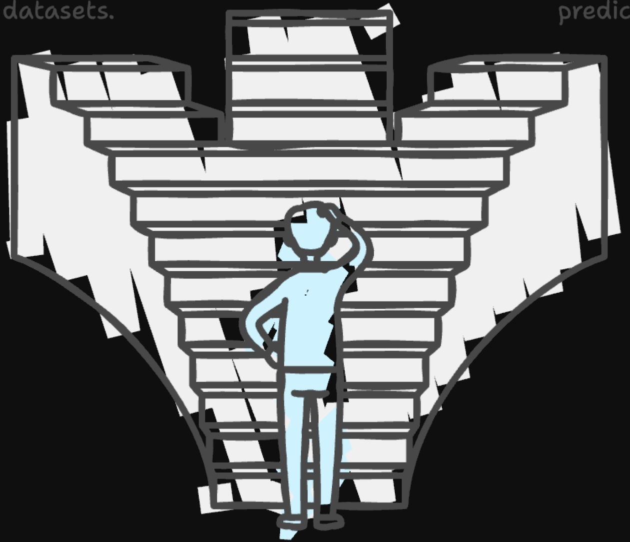
Simple and easy to understand, but may not be reliable for imbalanced datasets.

F1-score

Balances precision and recall, suitable for imbalanced datasets.

FPR

Identifies bias in the model by assessing incorrect positive predictions.



Summary of the Methodology Pipeline

The seven main phases of the GSMo-CNN pipeline include:

Dataset Preparation: Standardization of Images and Splitting for the training, validation, and testing of the data

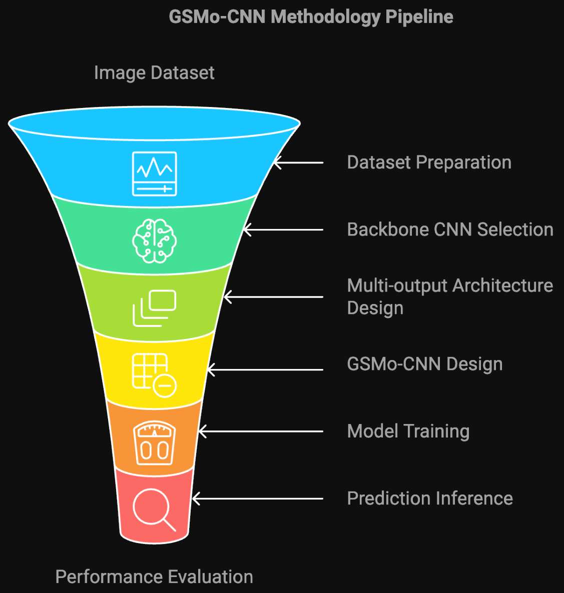
Backbone CNN: Determination of either the inceptionV3 or the resnet as it performs the tasks of feature extraction or any other multi-output architecture design: Single multi-task strategy will do, stacking more output layers increases the accuracy.

gsmo-cnn design. Utilize the following aspects; convolutional, temporary prediction, and stacked layers

Training the Model by optimization using minimum loss value and adjustment of weights values.

Infer Predictions: Obtain more accurate predictions of both species and disease.

Evaluation of Performance: Test with accuracy, F1-score, FPR.



The stacked architecture enables GSMo-CNN to treat the two tasks as one in a compact structure. The architecture of the stacked model leverages relationships between tasks hence improving accuracy. It forms a great base for further development of such applications for practical use in agricultural fields, especially in field diagnostics.

Key Observation

1.**Controlled vs. Real-world Conditions:**

Models perform best on Plant Village due to its clean, controlled setup, where there are minimal distractions, and disease symptoms are clearly visible.

Performance drops on PlantDoc highlight the additional challenge posed by real-world conditions, where models must contend with noise, shadows, and various other confounding factors.

2.**Impact of Background Complexity:**

PlantDoc's complex backgrounds present a significant challenge, especially for models not optimized for high noise levels. CNNs must differentiate disease-related features from unrelated background elements.

Plant Leaves, with moderate variability, serves as a good dataset, providing insights into a model's ability when moving from controlled set up to real-world set up.

3.Species and Disease Diversity:
Though preparing Plant Village is much more complicated and allows the preparation to achieve the greatest accuracy of species and diseases, a comparable preparation is easier. However, the models are tested somewhat under more realistic scenarios than the Plant Village does, enabling them to bridge laboratory and field data.

4. Performance Trend Across Datasets:

Since datasets are relatively uncontrolled along with the diversity in field conditions, accuracy down goes. InceptionV3 exhibits the maximum performance on all the datasets. For this type of model which trains without diversity in laboratory settings alone, it is observed at times that these might drastically fail in the real-world scenario due to some noisy backgrounds; so the requirement of variance at times of training datasets gets pertinent

Comparing CNN performance in controlled vs. real-world conditions.

The study highlights the effectiveness of the Generalised Stacking Multi-output CNN (GSMo-CNN) for simultaneous plant identification and disease classification. InceptionV3 emerged as the most robust backbone model, achieving superior performance across various datasets, including controlled and real-world scenarios. Its ability to handle complex features with high accuracy makes it particularly suitable for practical applications in precision agriculture, thereby enhancing plant pathology diagnostics.

Thank you.