

YOLO-Ash: Advanced YOLO11 Modifications for Ash Gourd Leaf Diseases Detection

BY

Nusrat Jahan

ID: 232-25-013

This Report Presented in Partial Fulfillment of the Requirements for
The Degree of Masters of Science in Computer Science and Engineering

Supervised By

Dr. Md Zahid Hasan

Associate Professor

Department of CSE

Daffodil International University

Co-Supervised By

Mr. Md. Sadekur Rahman

Assistant Professor

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2025

APPROVAL

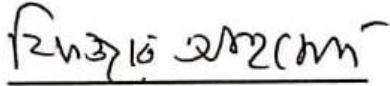
This Thesis titled “YOLO-Ash: Advanced YOLO11 Modifications for Ash Gourd Leaf Diseases Detection”, submitted by Nusrat Jahan, ID No: 232-25-013 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 11-01-2025.



BOARD OF EXAMINERS

Chairman

Dr. S.M Aminul Haque
Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Dr. Fizar Ahmed
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Dr. Md Alamgir Kabir
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University





External Examiner

Mr. Sadat Hossain
Data Scientist
Risk Management Division,
BRAC Bank Limited

DECLARATION

I hereby declare that this research has been done by me under the supervision of **Dr. Md Zahid Hasan**, Associate Professor, Department of CSE, Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by: 
Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University


Co-Supervised By
Mr. Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

Nusrat Jahan
Submitted by:
Nusrat Jahan
232-25-013
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First I express my heartiest thanks and gratefulness to Almighty Allah for His divine blessing which makes it possible to complete the final year thesis successfully.

I am really grateful and wish my profound indebtedness to **Dr. Md Zahid Hasan, Associate Professor**, Department of CSE, Daffodil International University, Dhaka, deep knowledge & keen interest of my supervisor in the field of Machine Learning to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to **Mr. Md. Sadekur Rahman, Assistant Professor**, Department of CSE, for his kind help to finish our project and also to other faculty members and the staff of CSE department of Daffodil International University.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

The detection and classification of plant diseases are critical for improving crop yields and ensuring food security. This study focuses on developing an enhanced object detection model tailored for Ash Gourd leaf diseases, a crop prone to significant yield losses due to common diseases such as Aphid infestation, Downy Mildew, Leaf Curl, and Leaf Miner. We proposed a customized YOLOv11 architecture with significant modifications, including the integration of the Convolutional Block Attention Module (CBAM), an Enhanced Spatial Pyramid Pooling-Fast (SPPF) module, the Hard-Swish activation function, and an optimized Non-Maximum Suppression (NMS) technique. These enhancements aimed to improve the model's ability to extract disease-relevant features while maintaining computational efficiency. The model was trained on a newly curated dataset of Ash Gourd leaf images, captured under real-world conditions, and classified into five categories: Healthy, Aphid, Downy Mildew, Leaf Curl, and Leaf Miner. Experimental results demonstrate that the proposed model achieves superior performance compared to baseline architectures, with a precision of 89.3%, recall of 88.8%, and mAP@0.5 of 89.3%. Incorporating attention mechanisms and feature fusion techniques significantly improved detection accuracy, particularly for subtle and overlapping disease patterns. This study highlights the potential of the customized YOLOv11 model as a robust and efficient tool for precision agriculture applications.

TABLE OF CONTENTS

Contents

CHAPTER 1	1
INTRODUCTION	1
1.1 Introduction	1
1.2 Motivation	3
1.3 Rationale of the Study	3
1.4 Research Questions	4
1.5 Expected Output	4
1.6 Project Management and Finance.....	4
CHAPTER 2	6
BACKGROUND	6
2.1 Preliminaries/Terminologies	6
2.2 Related works	6
2.3 The Problem's Scope	8
2.4 Challenges	8
CHAPTER 3	10
RESEARCH METHODOLOGY	10
3.1 Proposed Methodology/Applied Mechanism	10
3.2 Data Description.....	11
3.3 Experimental Design, Materials and Methods	15
3.4 Experimental Deep Learning Models.....	18
3.5 YOLO-Ash Customized Architecture	21
3.6 Experimental setup	31
CHAPTER 4	33
EXPERIMENTAL RESULTS AND DISCUSSION	33
4.1 Customized YOLOv11 Ablation Experiments.....	33
4.2 Performance comparison of baseline models on Ash Gourd dataset	34
4.3 Class-wise Performance of YOLO-Ash on Ash Gourd dataset.....	36
4.4 Inference Time and Processing Performance	37
4.5 Precision-Recall Trade-Off	38
4.4 Discussion	45

CHAPTER 5	47
IMPACT ON AGRICULTURE and ENVIRONMENT	47
5.1 Impact on Agriculture.....	47
5.2 Impact on Environment	47
5.3 Ethical Aspects	48
CHAPTER 6	49
CONCLUSION AND FUTURE WORK	49
6.1 Summary of the Study	49
6.2 Conclusions	50
6.3 Implications for Further Study	50
REFERENCES	52

LIST OF FIGURES

FIGURES	PAGE NO
Fig. 3.1 The working process to perform leaf disease detection	10
Fig. 3.2 above describes the distribution of images in each of the compressed files in the dataset	12
Fig. 3.3 A sample of Ash Gourd leaf diseases and a healthy set	13
Fig. 3.4 Data Collection and Processing Pipeline for Ash Gourd Dataset	16
Fig. 3.5 VGG16 architecture	19
Fig. 3.6 Architecture of Efficient-Swin	20
Fig. 3.7 YOLO-Ash Customized Architecture	22
Fig. 3.8 Convolutional Block Attention Module	24
Fig. 3.9 Enhanced SPPF	29
Fig. 4.1 The precision-recall (PR) curve for improved model.	40
Fig. 4.2 The accuracy and loss curves of YOLOv11 custom model	41
Fig. 4.3 Confusion matrix YOLOv11 custom model	42
Fig. 4.4 F1 score curves YOLOv11 custom model	43
Fig. 4.5 Precision-recall curves YOLOv11 custom model	44
Fig. 4.6 The detection score of our proposed model runs on some test images.	45

LIST OF TABLES

TABLES	PAGE NO
Table 3.1 Described the software and hardware specifications	31
Table 3.2 Described the default hyperparameters for our experiments	32
Table 4.1 Describe the performance evaluation of the proposed model and its variants	34
Table 4.2 Shows the performance comparison of the proposed model and other YOLO models.	35
Table 4.3 Present the detailed Class-wise Performance of YOLO-Ash on the Ash Gourd Leaf Disease Detection Dataset.	36
Table 4.4 Highlight the comparison of inference and processing times of different models.	38

CHAPTER 1

INTRODUCTION

1.1 Introduction

The accurate detection and classification of plant leaf diseases are critical to enhancing crop productivity and reducing the economic impact of agricultural losses. Plant diseases often manifest as subtle variations in texture, color, and structure, making manual identification a labor-intensive and error-prone process [1]. Traditional inspection methods rely heavily on human expertise, which may vary significantly depending on the skill and experience of the evaluator. Furthermore, these methods are time-consuming and unsuitable for large-scale farming operations where rapid and precise identification is essential to minimize losses. Consequently, there is a growing demand for automated and intelligent solutions for leaf disease detection [2].

Advancements in computer vision, combined with deep learning methodologies, have revolutionized the field of automated plant disease detection [3]. Modern object detection models have shown remarkable potential in identifying and classifying diseases by learning discriminative features from large datasets. Among these models, attention mechanisms and enhanced feature extraction techniques stand out for their ability to prioritize critical features while suppressing irrelevant information. Attention mechanisms, in particular, dynamically focus on disease-relevant spatial and channel-wise features, enabling models to detect fine-grained patterns that may be otherwise overlooked [4]. These mechanisms enhance the representational power of deep neural networks, especially when applied to datasets with intricate visual patterns and variable environmental conditions [5].

Incorporating attention mechanisms into feature extraction pipelines has proven effective in tackling the inherent challenges of leaf disease detection. For example, models equipped with attention modules have demonstrated improved precision and recall by identifying subtle disease symptoms, even in the presence of noise and overlapping structures [3,5]. Additionally, enhanced feature extraction techniques, such as multi-scale pooling and custom activation functions, have been employed to improve the

robustness and computational efficiency of deep learning models. These innovations are particularly valuable for agricultural datasets, where lighting, angles, and occlusions often introduce inconsistencies in the input images [2,4]

Our research addresses these challenges by customizing the YOLOv11 architecture to meet the specific requirements of Ash Gourd leaf disease detection. The proposed enhancements include the integration of the Convolutional Block Attention Module (CBAM), which refines the network's ability to prioritize spatial and channel-wise features relevant to disease detection. We further optimize the model by replacing the default activation function with Hard-Swish, a lightweight alternative that reduces computational overhead while maintaining high detection accuracy [5]. To improve multi-scale feature aggregation, an enhanced Spatial Pyramid Pooling-Fast (SPPF) module is incorporated, which leverages parallel max-pooling layers and dilated convolution to capture intricate disease patterns across varying scales. These customizations are designed to address the challenges posed by the real-world Ash Gourd leaf dataset, which exhibits complex structures, subtle disease symptoms, and varying environmental lighting [1,4]

By integrating these novel components, the customized YOLOv11 architecture provides a robust and efficient solution for detecting leaf diseases in agricultural settings. Our model achieves superior performance in terms of precision, recall, and mean average precision (mAP) while maintaining computational efficiency, making it suitable for large-scale deployment. This research contributes to the growing body of knowledge in automated plant disease detection by demonstrating the effectiveness of attention mechanisms and enhanced feature extraction in improving the detection of complex and subtle disease patterns [2,5]

1.2 Motivation

The rising demand for sustainable agriculture and the growing challenges posed by plant diseases underscore the need for efficient and accurate detection methods. Traditional manual inspection methods for diagnosing leaf diseases in ash gourd are not only time-intensive but also prone to errors, especially at large scales. With ash gourd being a critical crop for its nutritional and economic value, timely disease management is essential to prevent significant yield losses.

Advancements in deep learning and computer vision provide a unique opportunity to automate and enhance disease detection processes, offering solutions that are faster, more accurate, and scalable. This work is motivated by the potential to bridge the gap between cutting-edge AI technologies and practical agricultural needs, aiming to empower farmers with real-time tools for effective crop management and ultimately contribute to food security and sustainable farming practices.

1.3 Rationale of the Study

Ash gourd cultivation faces significant challenges from leaf diseases, impacting yield and quality. Traditional detection methods are labor-intensive, time-consuming, and impractical for large-scale applications. This study utilizes advanced deep learning models, including VGG-16, Eff_Swin_Model, YOLOv5, YOLOv6, YOLOv8, YOLOv9, YOLOv10 and YOLOv11l, to develop an automated, accurate, and real-time detection system. The focus on customized YOLOv11l, known for its superior detection capabilities, addresses the critical need for scalable solutions in agricultural disease management, promoting sustainable and technology-driven farming practices.

1.4 Research Questions

- RQ1 How effective is a deep learning model such as YOLOv11 in accurately detecting ash gourd leaf diseases rather than YOLOv5 and other YOLO versions?
- RQ2: What are the comparative strengths of the selected models in terms of accuracy, precision, recall, and inference speed for real-time disease detection?
- RQ3: Can the proposed YOLOv11 model provide a scalable and efficient solution for real-world applications in automated ash gourd leaf disease management?

1.5 Expected Output

- Development of a Robust Model: Creation of a scalable and accurate YOLOv11-based framework for ash gourd leaf disease detection.
- Comprehensive Model Analysis: Comparative evaluation of different YOLO models and customized YOLOv11 to determine their performance in disease classification.
- Automation in Agriculture: A practical tool for farmers enabling real-time identification and classification of leaf diseases, reducing manual labor and improving crop management.
- Foundation for Future Research: Establishment of a dataset and framework that can be extended to other crops and disease detection challenges in precision agriculture.

1.6 Project Management and Finance

The research work doesn't get fund from any individuals or organization.

1.7 Report Layout

In Chapter 1, the introduction, objectives, and key research inquiries of the study are outlined. In Chapter 2, concise synopses of the literature review are provided. In Chapter 3, the proposed methodology is described in detail. In Chapter 4, the experimental outcomes of the paper are described and examined. The fifth chapter discusses the sustainability plan, societal and environmental impacts, and ethical considerations. The sixth chapter concludes the present investigation and outlines a strategy for subsequent endeavors.

CHAPTER 2

BACKGROUND

2.1 Preliminaries/Terminologies

Modern advancements in agricultural technology have paved the way for the development of sophisticated systems to detect and classify diseases affecting crops like ash gourd. Leveraging image processing techniques, numerous studies have focused on identifying and categorizing various maladies impacting ash gourd leaf health. Accurate recognition of diseases such as aphid infestations, downy mildew, leaf curl, and leaf miner has been enabled by applying object detection and machine learning algorithms, focusing on features like shape, texture, and disease-specific patterns. These advancements represent a significant step forward in crop disease management and have the potential to revolutionize agricultural practices, promoting more efficient and sustainable farming methods in ash gourd cultivation.

2.2 Related works

The advancement of deep learning techniques has greatly influenced the field of automated plant disease detection, particularly through the application of YOLO-based architectures enhanced with attention mechanisms and advanced feature extraction methods. These methods address the limitations of traditional detection systems by improving the focus on disease-relevant features and enhancing the robustness of models against challenges such as variable environmental conditions, overlapping structures, and subtle disease symptoms. The integration of attention mechanisms has emerged as a critical component for addressing these challenges, as they dynamically refine spatial and channel-wise feature representations, ensuring better performance in detecting diseases across diverse datasets.

In recent studies, attention mechanisms have been widely adopted to enhance YOLO-based models for agricultural applications. For instance, the TSBA-YOLO model utilizes attention modules and feature fusion techniques to detect tea leaf diseases with improved accuracy and robustness. By combining multiple attention mechanisms, this model successfully identifies subtle disease patterns in complex datasets, demonstrating the

effectiveness of attention mechanisms in capturing fine-grained features and prioritizing critical information [6]. Similarly, the YOLO-Sesame model employs an enhanced attention mechanism combined with feature aggregation to address the detection of weeds in sesame fields. This approach resolves challenges associated with overlapping structures and small targets, which are common in agricultural datasets [7].

Further advancements have been made with the integration of Convolutional Block Attention Modules (CBAM) into YOLO architectures. For example, the YOLO-CBAM model refines spatial and channel-wise attention by dynamically emphasizing disease-relevant features while suppressing noise. This integration significantly improves real-time detection capabilities, particularly in identifying invasive weeds with high precision and recall, thereby validating the role of CBAM in agricultural applications [8]. Similarly, YOLO-Papaya leverages CBAM to optimize the detection of papaya fruit diseases. By prioritizing disease-relevant features and reducing irrelevant information, this model achieves improved accuracy under varying environmental lighting conditions, making it particularly effective for real-world agricultural scenarios [9].

Moreover, models focusing on small-target detection have also incorporated attention mechanisms to overcome the challenges posed by densely packed scenarios. In one such example, researchers introduced shallow feature layers combined with attention modules into the YOLOv5 framework for wheat ear detection. This enhancement allowed the model to accurately identify small targets in complex environments, further highlighting the versatility of attention mechanisms in agricultural applications [10]. These studies underscore the importance of attention-enhanced models in addressing the unique challenges posed by agricultural datasets, where subtle patterns and overlapping structures are common.

Building upon these advancements, our study customizes the YOLOv11 architecture to improve its performance in detecting Ash Gourd leaf diseases. Unlike previous approaches, we integrate an enhanced Spatial Pyramid Pooling-Fast (SPPF) module for improved multi-scale feature aggregation, CBAM for dynamic feature prioritization, and Hard-Swish activation to enhance computational efficiency without sacrificing accuracy.

These customizations are specifically designed to address the unique challenges posed by the Ash Gourd leaf dataset, such as subtle disease patterns, overlapping structures, and varying environmental conditions. By addressing these limitations, our work bridges the gap in detection performance and computational efficiency, contributing to the broader field of automated plant disease detection.

The integration of these novel techniques into the YOLOv11 framework not only enhances precision, recall, and mean average precision (mAP) but also ensures robustness and scalability for large-scale agricultural applications. This approach represents a significant advancement in the development of attention-enhanced models for detecting plant diseases, paving the way for more accurate and efficient solutions in agricultural disease management.

2.3 The Problem's Scope

Thus, deep learning models, particularly a customized YOLOv11 architecture, are presented in this research to detect and classify diseases in ash gourd leaves, addressing the challenges associated with existing approaches. Prior research has demonstrated the successful application of various machine learning and deep learning techniques in the detection of vegetable leaf diseases. However, several challenges persist, including suboptimal detection accuracy, time inefficiencies in real-time applications, and insufficient methods for managing diverse environmental conditions in datasets. The aim of this research is to overcome these limitations by evaluating and comparing state-of-the-art detection architectures while customizing YOLOv11 to enhance its accuracy, speed, and robustness for ash gourd leaf disease detection.

2.4 Challenges

The acquisition of datasets that accurately capture the wide variety of ash gourd leaf disease manifestations across diverse environmental conditions is a significant challenge. Ensuring comprehensive and high-quality data is critical but requires extensive resources

and domain expertise to annotate symptoms effectively. Variability in environmental factors, such as changes in lighting, background complexity, and disease manifestation patterns, further complicates the development of robust datasets.

Environmental variability poses an additional challenge, as illumination differences, leaf orientations, and dynamic field conditions can reduce the accuracy of disease detection models. This variability often results in reduced reliability when these models are applied to real-world scenarios outside the controlled conditions of the training data.

Model generalization remains a critical issue, referring to the ability of developed models to operate effectively on novel and unfamiliar data. This requires overcoming the constraints of overfitting to specific characteristics of the training dataset and ensuring consistent performance across unseen samples, regions, and environmental conditions.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Proposed Methodology/Applied Mechanism

The provided diagram 3.1 a streamlined workflow illustrating the process of ash gourd leaf disease detection, from data acquisition to result analysis.

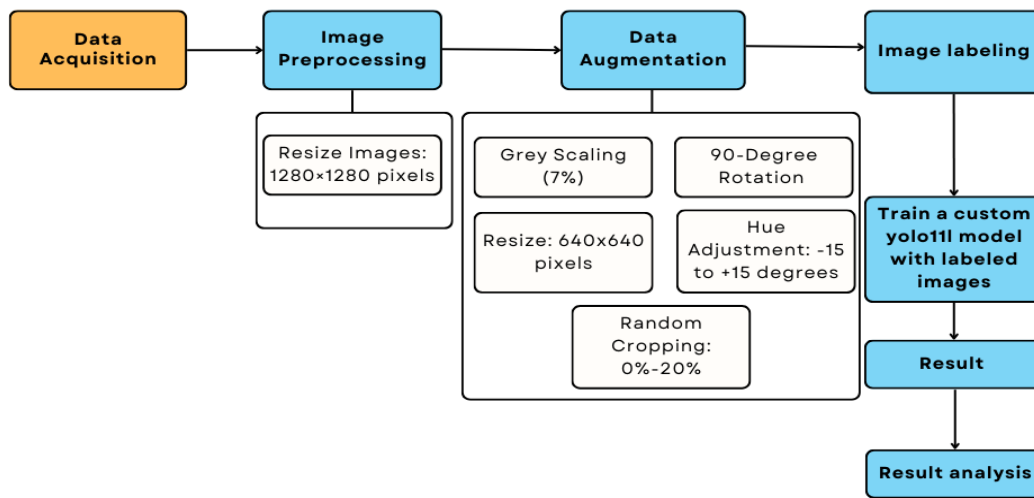


Fig. 3.1 The working process to perform leaf disease detection

In this study on ash gourd leaf disease detection, the methodology follows a structured and systematic approach to process and analyze leaf images for effective disease identification. The process begins with the acquisition of a dataset comprising high-resolution images of ash gourd leaves, capturing a range of disease conditions such as aphid infestations, downy mildew, leaf curl, and leaf miner damage. These images form the foundation for the detection framework.

The next step involves preprocessing the images to prepare them for analysis. This includes resizing the images to uniform dimensions and applying normalization to ensure consistency in pixel intensity across the dataset. These preprocessing steps are vital for

enhancing the compatibility of the data with the YOLOv11 architecture. To improve model robustness, data augmentation techniques are applied to the preprocessed images. These techniques include random cropping, grey scaling, rotation, and hue adjustments, which introduce variability and diversity to the dataset. Augmentation ensures the model is resilient to real-world variations, such as changes in lighting, orientation, and background.

The core of the methodology is the customization and training of the YOLOv11 model. This state-of-the-art object detection architecture is fine-tuned to detect and classify ash gourd leaf diseases with high precision and efficiency. The labeled dataset, created during preprocessing and augmentation, is used to train the model, enabling it to accurately recognize disease patterns in the images.

Following training, the model generates predictions for test images, identifying diseased and healthy leaves with bounding boxes and confidence scores. The results are then subjected to an in-depth analysis, where the model's performance is evaluated against ground truth annotations using metrics such as accuracy, precision, recall, and the F1 score. This evaluative step is instrumental in assessing the effectiveness of the customized YOLOv11 model and provides a foundation for further refinement and research in precision agriculture and disease management.

3.2 Data Description

The Ash Gourd dataset is organized to be helpful or used to detect and classify diseases occurring within Ash Gourd leaves. The classes of this dataset are primarily considered two categories: Healthy and Unhealthy. However, the class Unhealthy has been further divided into four specific types of diseases based on the type of infection or pest damage observed. Each category is in a separate folder with images showing typical visual symptoms associated with each kind of disease. The collection offers various ash gourd leaf images spanning five categories: Healthy, Aphid, Downy Mildew, Leaf Curl, and Leaf Miner. This approach supports the structured classification of healthy leaves or those infected by specific diseases in the dataset; it is thus suitable for model training and

validation tasks in plant disease detection. Fig.3.2 depicts the distribution of images in each disease category.

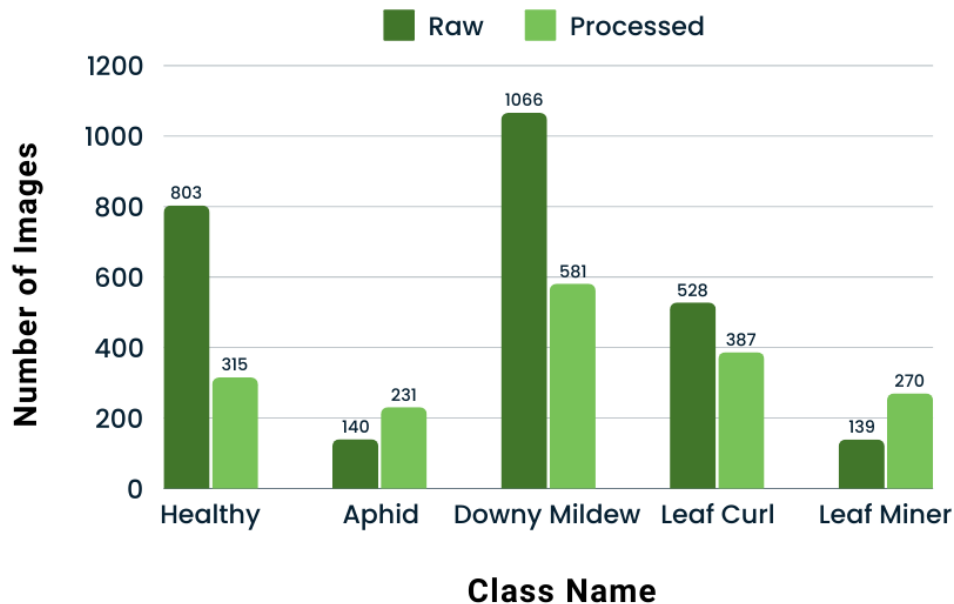


Fig. 3.2 Above describes the distribution of images in each of the compressed files in the dataset.

Images in the dataset of ash gourd leaf have been captured using an 8 MP smartphone camera with a telephoto lens, f/2.0-aperture, 1.0 μ m pixel size, PDAF, and OIS. The camera supports up to 3.7 times optical zoom and 5 times hybrid zoom with enhanced digital processing and captures highly detailed pictures even when zoomed to further distances.

Images of leaves were captured at standardized and static distances of 20 to 90 cm to ensure close-up shots of the leaves while also giving slightly wider shots. Indeed, for it to be an image dataset that most accurately describes the real world, pictures are shot in a real-life environment from several angles and with various natural light, including sunlight, diffused light, and partial shade. Since these are variations in imaging conditions, the model training and testing against disease symptoms and leaf positions will stand on robust ground, strengthening the standpoint of the dataset's applicability for trustworthy and adaptable ash gourd disease detection systems.

Fig.3.3 displays samples of Ash Gourd leaves affected by various diseases (Downy Mildew, Leaf Curl, Leaf Miner, Aphid) and a healthy set, showing both the front and back sides for each condition.

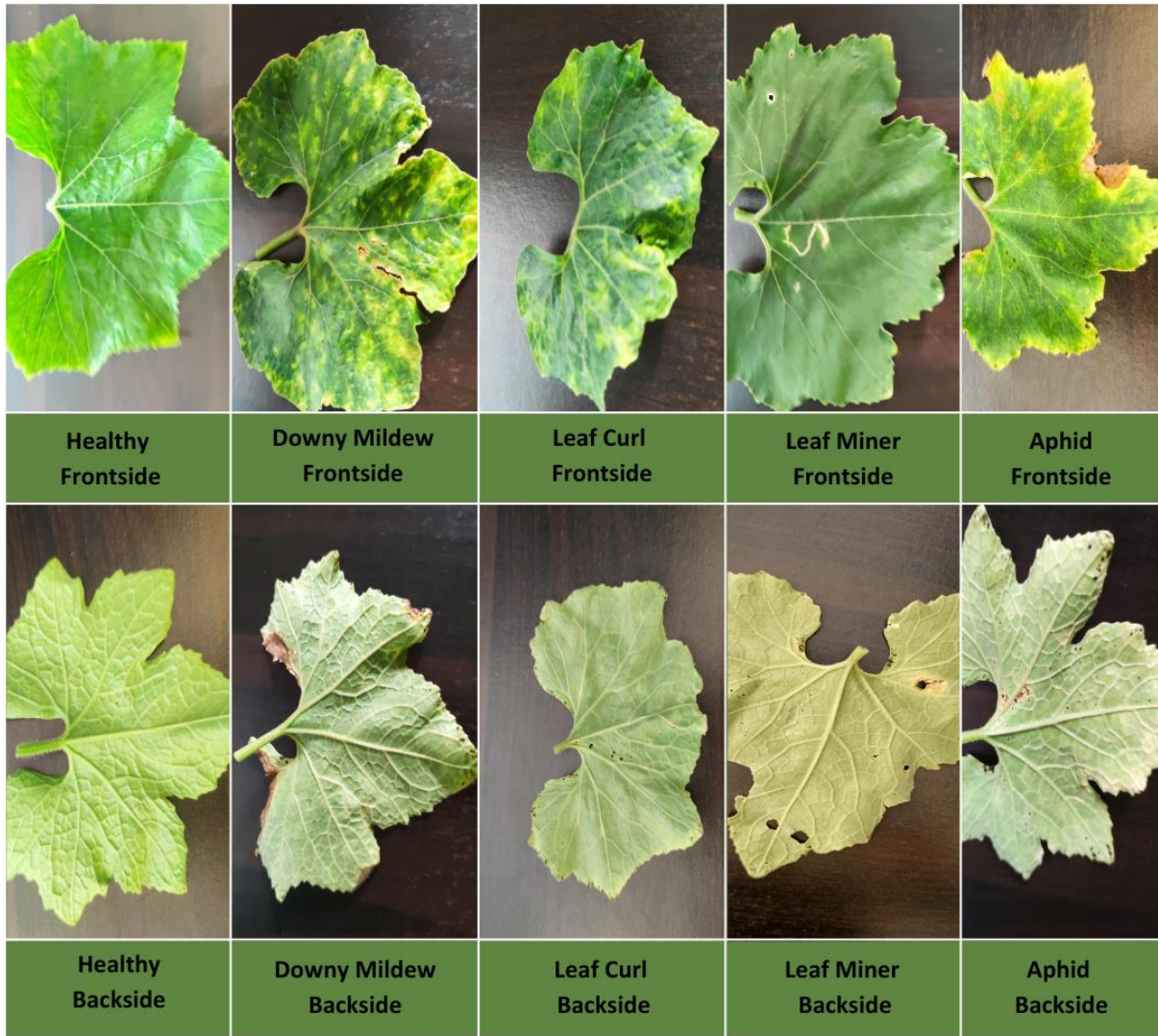


Fig. 3.3 A sample of Ash Gourd leaf diseases and a healthy set.

Healthy: This class consists of images of normal ash gourd leaves free of diseases, pests, and all other kinds of damage. The leaves are of normal shape and without any blemishes, which provides a basis for matching against the leaves of diseased ones. These include even coloration, no breakage of the leaf structure, no visible spots,

discoloration, curling, or any kind of deformity. It is used as a control group to create contrast in classification, either for a healthy or diseased leaf.

Downy Mildew: This class includes images showing leaves infected with the fungal disease Downy Mildew. These infections thrive in moist and humid conditions and typically spread on the undersides of leaves, causing distinctive mottling and lesions. Yellow to pale green patches on the upper leaf surface and downy, greyish-white fungal growth on the underside. Affected areas may turn brown or necrotic as the disease progresses. Downy mildew reduces photosynthesis, weakens overall plant growth, and can severely affect crop yield if untreated.

Leaf Curl: Leaves in this category show leaf curl, often resulting from viral infections or pest vectors. The virus disrupts the leaf's cellular structure, causing it to twist or curl. Puckering or twisting of leaf margins is often accompanied by discoloration and thickening of leaf tissue. Leaves may also appear leathery, distorted, or stunted. Leaf curl hinders photosynthesis and normal plant growth, potentially leading to stunted development and reduced fruit production.

Leaf Miner: This class represents leaves affected by leaf-mining insects. The larvae of these insects burrow between leaf layers, leaving visible feeding trails. Winding or blotchy marks across the leaf surface, often appearing as light-colored, translucent pathways. Affected areas may die off and turn brown as the damage advances. Leaf miner infestation reduces the photosynthetic area of the leaf, weakening the plant. Severe infestations can cause premature leaf drop and significantly lower plant productivity.

Aphid: These leaves are suffering from an infection caused by aphids. Aphids are tiny insects that feed on sap, causing serious damage to the leaf surface and generally affecting the plant's vitality. Small punctures, yellowing of leaves, and a mottled appearance. Leaves may appear distorted and develop sticky honeydew patches. Aphids can weaken the plant, reduce photosynthetic ability, and make the plant more vulnerable to secondary infections.

Each class in this dataset corresponds to a unique visual pattern, enabling machine learning models to distinguish between healthy and various disease states in ash gourd leaves. This granular categorization supports early disease detection and pest management in ash gourd cultivation.

3.3 Experimental Design, Materials and Methods

3.3.1 Data Acquisition

The acquisition was done through field collection of ash gourd leaf images using a high-resolution smartphone camera. Working with agricultural pathologists, we identified both the diseased and healthy ash gourd leaves, documenting them. On each visit to a session, pathologists helped diagnose the specific diseases affecting each leaf for proper identification and subsequent classification and labeling of images.

The images have been taken in natural daylight to represent the diseased and healthy leaves in their natural appearance. Photography of leaves at different angles and distances captures detailed features relevant to disease identification. In total, five classes of distinct image categories were collected: Healthy and four disease-based classes: Aphid, Downy Mildew, Leaf Curl, and Leaf Miner.

Fig 3.4 illustrates a data collection and processing pipeline for an Ash Gourd dataset, detailing stages of image acquisition, preprocessing, augmentation, and organization into disease categories.

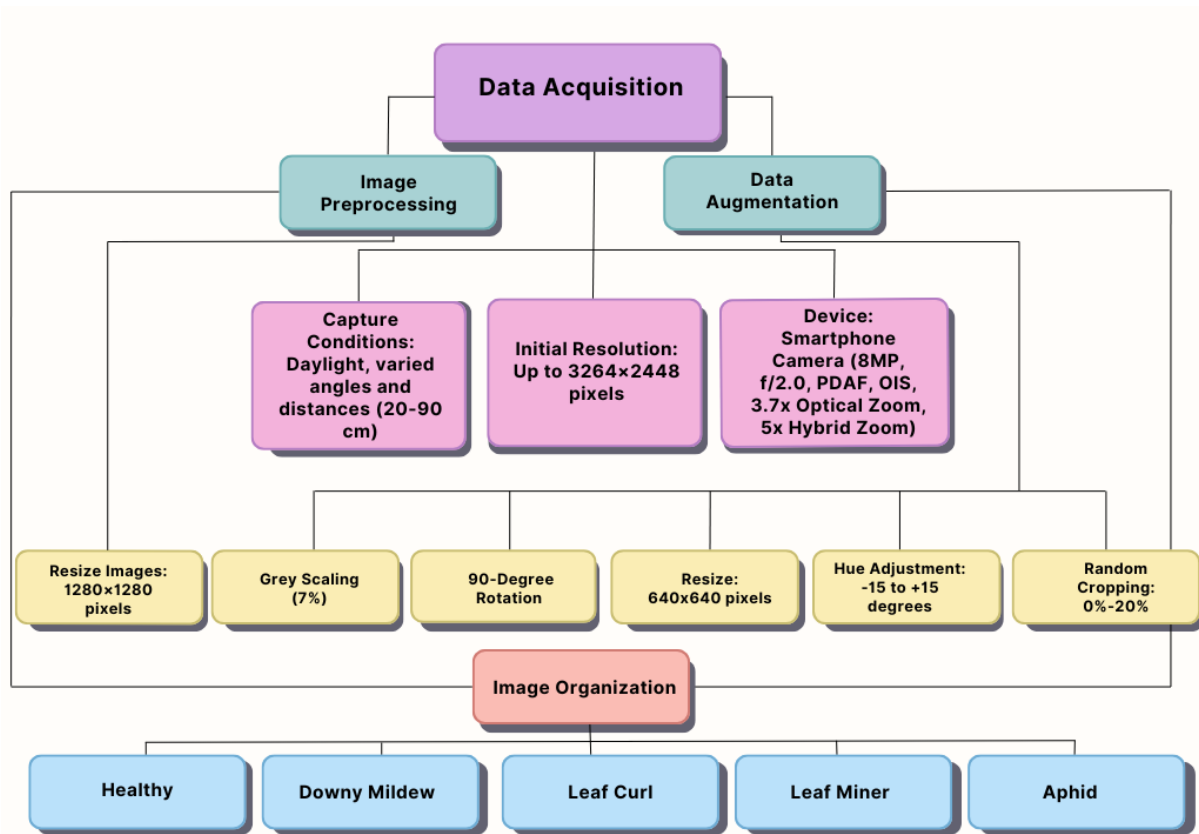


Fig. 3.4 Data Collection and Processing Pipeline for Ash Gourd Dataset

3.3.2 Data pre-processing

It includes all the images in this dataset, resized to a fixed resolution of 1280×1280 pixels. This ensures that the images will all exactly have the same size, fitting into specific-sized requirements for the input of any neural network. The sizing of images should be uniform in deep learning models, which take input of specific dimensionality for their processing. All the images are resized to 1280 × 1280. Hence, it speeds up data processing and blesses the model with better learning and generalizing of the pattern irrespective of the different sizes of images in the dataset.

3.3.3 Data augmentation

During data augmentation, each image undergoes a series of transformations designed to improve the model's robustness and adaptability. This process incorporates conventional augmentation techniques commonly employed in research for plant disease identification. This first undergoes a 90-degree rotation clockwise or counterclockwise, adding positional variability, which is good for plant leaf images that may appear in different orientations in real-life scenarios, hence improving the model's strength in detecting disease patterns irrespective of the orientation of the leaf. Furthermore, random cropping in the range of 0% to 20% shifts the attention to different areas of the leaf. It can simulate various distances and framings, improving adaptability and preventing overfitting to particular patterns. A random color shift from -15 to +15 degrees slightly changes the hue. This mimics natural light and sensor color variations, which train the model on health markers rather than color consistency.

Additionally, 7% of the images are converted to grayscale so that the model learns from the structural features and not the color to detect diseases better, as most of these diseases might not depend on specific color markers. Lastly, all images are resized to 640x640 pixels because it is important to maintain uniformity across the dataset for efficiency in processing, while optimal focus is to be learned by the model classifying diseases; hence, the dimensions are consistent without losing critical detail. This heavy augmentation strategy will make the model invariant to rotation, framing, lighting, and color variations, enhancing performance under diverse conditions in real-life applications.

3.3.4 Limitations

This dataset comes from a relatively small area and, therefore, might not represent the broader diversity of Ash Gourd plants from various climates and environmental conditions. Symptoms and leaf appearance can vary significantly due to regional factors like humidity, soil composition, and weather conditions. Consequently, this geographical limitation may mean that the dataset does not fully capture disease manifestations or health indicators relevant to other regions, affecting its applicability for studies or applications involving Ash Gourd plants elsewhere.

The dataset's overall sample size is also quite limited, reducing its diversity and comprehensiveness regarding disease symptoms and plant health indicators within each category: healthy, aphid, downy mildew, leaf curl, and leaf miner.

3.4 Experimental Deep Learning Models

This research developed a customized yolov11 model to detect Ash Gourd Leaf Diseases.

3.4.1. VGG16- A Deep Learning Architecture

The VGG16 architecture is a widely used convolutional neural network designed for image classification and feature extraction tasks. The model consists of sequential convolutional and pooling layers, enabling it to capture hierarchical feature representations.

In VGG16, the input image is passed through a series of convolutional layers, organized into five blocks. Each block contains multiple convolutional layers with small 3×3 filters, followed by a max-pooling layer for downsampling. This design ensures the model captures detailed spatial features at different levels of abstraction.

Block 1 processes the input with two convolutional layers, producing feature maps of size $(224, 224, 64)$, followed by max pooling to reduce the spatial dimensions to $(112, 112, 64)$.

Block 2 expands the feature space to 128 channels through two convolutional layers, with max pooling reducing the spatial size to $(56, 56, 128)$.

Block 3 introduces three convolutional layers, increasing the feature depth to 256, followed by max pooling, resulting in output dimensions $(28, 28, 256)$.

The extracted features from the final pooling layer are flattened and passed through three dense (fully connected) layers for classification. These dense layers combine the hierarchical features learned in the convolutional blocks to make predictions, with the last layer outputting probabilities for the target classes.

VGG16 is notable for its simplicity and efficiency in capturing visual features, making it a backbone for many advanced architectures like object detectors and segmentation

networks. Its strengths lie in its depth and use of small filters, which ensure detailed feature learning while maintaining computational efficiency.

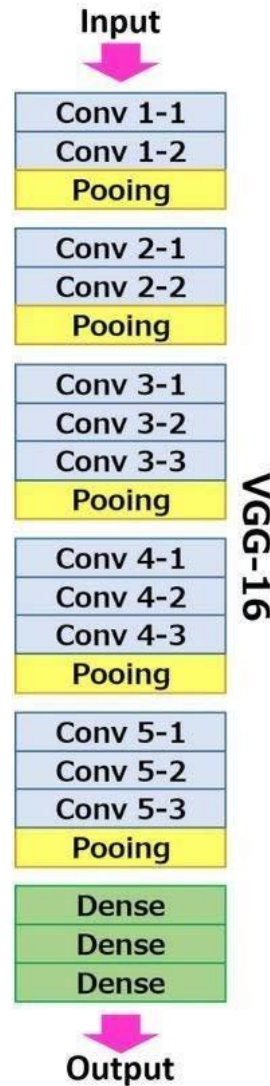


Fig. 3.5 VGG16 architecture

3.4.2 Efficient-Swin model

The Efficient-Swin model is a hybrid architecture that combines the feature extraction efficiency of convolutional neural networks (CNNs) with the global attention capabilities of transformers. It leverages EfficientNet-based MBConv layers for localized feature extraction and Swin Transformer blocks to capture hierarchical and long-range dependencies. The model uses patch merging to reduce spatial dimensions while

increasing feature depth, enabling multi-scale feature learning. This hierarchical structure integrates fine-grained details with broader contextual information, making it highly effective for complex image analysis tasks. With its efficient design and powerful attention mechanisms, the Efficient-Swin model is well-suited for applications requiring precision and scalability, such as image classification, object detection, and medical imaging.

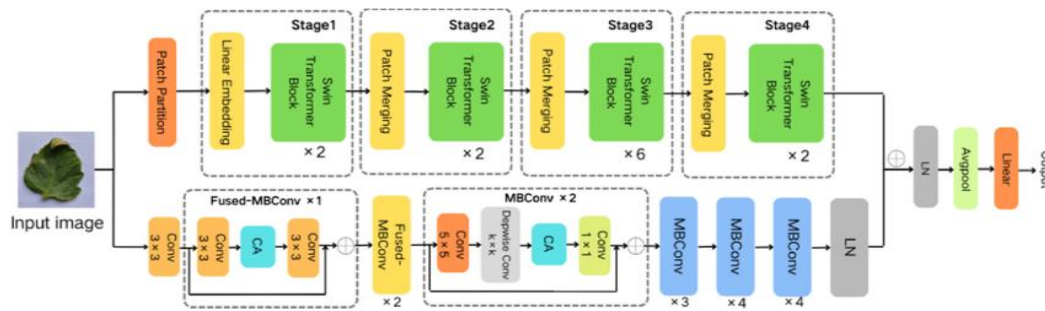


Fig. 3.6 Architecture of Efficient-Swin

3.4.3 YOLOv5, YOLOv6, YOLOv8, YOLOv9, YOLOv10

The YOLO (You Only Look Once) series has continually evolved, introducing significant advancements in object detection. YOLOv5 marked a key milestone with its lightweight structure, ease of deployment, and improved speed and accuracy, making it suitable for real-time applications. YOLOv6 further optimized performance by incorporating advanced anchor-free detection techniques and improved feature pyramids, enabling better generalization across diverse datasets. YOLOv8 introduced attention mechanisms and dynamic convolutional layers, enhancing feature extraction and improving detection precision for complex patterns. YOLOv9 built upon these advancements by integrating multi-scale feature fusion and transformer-based modules, addressing challenges in detecting small and overlapping objects. Most recently, YOLOv10 introduced innovative architecture designs, such as hybrid attention and adaptive feature refinement, offering state-of-the-art accuracy while maintaining computational efficiency. Together, these iterations demonstrate the continuous

innovation in the YOLO family, driving breakthroughs in real-time object detection and classification.

3.4.4. YOLOv11L (Large)

YOLOv11L (Large) is a robust and high-performing variant of the YOLOv11 family, optimized for environments where computational resources are abundant and maximum accuracy is essential. This model features a more complex backbone, neck, and head architecture, enabling it to extract rich and detailed feature representations. YOLOv11L incorporates advanced techniques like attention mechanisms, deeper layers, and enhanced spatial pyramid pooling for improved feature fusion and object localization. Designed for applications requiring high precision, such as autonomous driving, medical imaging, and video surveillance, YOLOv11L achieves superior detection performance while maintaining scalability for large-scale datasets and complex detection scenarios.

3.5 YOLO-Ash Customized Architecture

3.5.1. Network Structure

There is no doubt that YOLO-based models have set a benchmark in detection accuracy and real-time performance, establishing themselves as the gold standard for object detection across diverse applications. However, the YOLOv11 architecture, though designed for real-time object detection, faces significant limitations when applied to the detection of leaf diseases, which demand a deeper understanding of intricate patterns and subtle variations. Leaf structures are inherently complex, characterized by fine-grained textures, varying shapes, and diverse color profiles, all of which present unique challenges for conventional object detection models. Moreover, the presence of overlapping leaves, shadows, and uneven lighting in the dataset often obscure critical disease symptoms, such as slight discolorations or minor structural anomalies. These subtle features are easily overlooked by the standard feature extraction methods in YOLOv11, which are optimized for general-purpose detection rather than domain-specific tasks like leaf analysis. As a result, the default architecture struggle to prioritize

disease-relevant features, leading to suboptimal performance in accurately identifying and localizing diseased regions.

To address the challenges of leaf disease detection, we introduce a refined architecture derived from YOLOv11 tailored to enhance leaf disease detection by addressing the complexities of subtle patterns and overlapping structures in leaf datasets. The network structure is illustrated in Figure 3.7 outlines the core components of the YOLOv11 architecture and their customized functions. The backbone, responsible for feature extraction, remains efficient and lightweight, while the neck section incorporates the Convolutional Block Attention Module (CBAM). CBAM dynamically adjusts spatial and channel feature importance, ensuring disease-relevant patterns are prioritized. Additionally, the Spatial Pyramid Pooling-Fast (SPPF) module has been redesigned to include parallel max-pooling layers with kernel sizes of 5×5 , 3×3 , and 1×1 , followed by a dilated convolution layer. This enhanced structure significantly improves the integration of multi-scale features, critical for detecting fine-grained and overlapping disease regions.

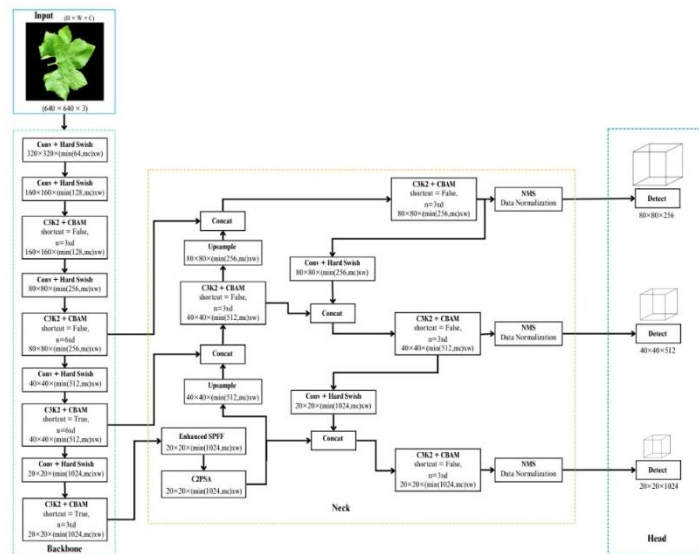


Fig. 3.7 YOLO-Ash Customized Architecture

The architecture further incorporates the Hard-Swish activation function in the neck, replacing Swish to improve computational efficiency without compromising feature

representation. Finally, the head section is augmented with a Non-Maximum Suppression (NMS) layer, which eliminates redundant bounding boxes and ensures that only the most confident detections are retained. This integration optimizes detection precision and performance for densely packed or overlapping leaves. By leveraging these targeted modifications, the architecture effectively captures complex leaf features, seamlessly integrates multi-scale information, and enhances detection accuracy and efficiency, demonstrating its suitability for robust and scalable leaf disease detection task ability to detect complex leaf features with improved accuracy, efficiency, and robustness.

3.5.2. Enhanced Feature Extraction with Attention

In deep learning models, particularly those designed for complex tasks such as leaf disease detection, it is essential to focus on disease-relevant features while suppressing irrelevant information. However, standard YOLOv11 models lack mechanisms to emphasize important spatial and channel features, often leading to noise propagation and suboptimal performance. To overcome this limitation, we integrated the Convolutional Block Attention Module (CBAM) into the neck section of YOLOv11. CBAM introduces a sequential application of channel and spatial attention mechanisms, enhancing the model's ability to prioritize subtle patterns, such as texture irregularities or discolorations, while reducing irrelevant feature redundancies [11,12]. The architecture of CBAM is depicted in Figure 3.8.

The core idea of CBAM is to refine the input feature map F by sequentially applying channel attention and spatial attention to adaptively recalibrate feature weights. Let the input feature map F be defined as $\in R^{C \times H \times W}$, where C represents the number of channels, and H and W are the spatial dimensions of the feature map. The output of CBAM, denoted as F_{CBAM} , is computed as follows:

$$F_{CBAM} = F_{spatial} \cdot F_{channel} \cdot F,$$

where $F_{channel}$ and $F_{spatial}$ represent the channel attention and spatial attention weights, respectively, and (\cdot) dot denotes element-wise multiplication.[13]

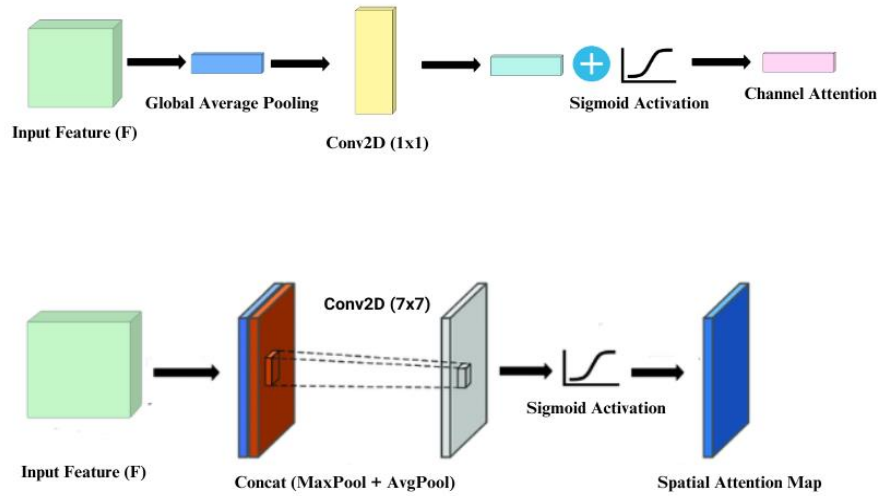


Fig 3.8 Convolutional Block Attention Module.

Channel Attention

The customized channel attention mechanism in our architecture focuses on enhancing the most significant channels of the feature map while suppressing the less informative ones. Unlike the traditional approach, this customization simplifies the attention process by exclusively utilizing global average pooling to aggregate information across spatial dimensions [14]. For a given input feature map F of size $C \times H \times W$, the channel-wise descriptors are computed as:

$$f_{\text{avg}}(c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j),$$

where c is the channel index, and H and W represent the height and width of the feature map, respectively. Instead of combining global max pooling and average pooling, only the average pooling descriptor f_{avg} is used in this customization to reduce computational complexity while maintaining performance [15]

The channel descriptors f_{avg} are then passed through a single 1×1 convolution layer instead of a Multi-Layer Perceptron (MLP), which reduces parameters and accelerates computation. The weights for channel attention are computed as:

$$M_{\text{channel}} = \sigma(\text{Conv}_{1 \times 1}(f_{\text{avg}})),$$

where σ represents the sigmoid activation function, and $M_{\text{channel}} \in R^C$ is the attention map for the channels. This approach simplifies the implementation by avoiding complex bottleneck structures while focusing on efficiency [13,14].

Finally, the refined feature map after applying channel attention is computed as:

$$F_{\text{channel}} = M_{\text{channel}} \cdot F,$$

where (\cdot) denotes element-wise multiplication between the attention weights and the input feature map. This refinement selectively emphasizes the most relevant channels in the feature map, improving the model's ability to detect nuanced features in leaf structures [15].

Spatial Attention

Spatial attention complements channel attention by identifying the most relevant spatial regions within the feature map. The spatial attention mechanism aggregates F_{channel} along the channel axis using global average pooling and global max pooling, resulting in two descriptors:

$$f_{\text{avg}}^{\text{spatial}} = \frac{1}{C} \sum_{k=1}^C F_{\text{channel}}(k, :, :), \quad f_{\text{max}}^{\text{spatial}} = \max_k F_{\text{channel}}(k, :, :).$$

These descriptors $f_{\text{avg}}^{\text{spatial}}$ and $f_{\text{max}}^{\text{spatial}}$, are concatenated and passed through a convolutional layer with a kernel size of 7×7 , followed by a sigmoid activation, to compute the spatial attention weights:

$$M_{\text{spatial}} = \sigma(\text{Conv}_{2D}([f_{\text{avg}}^{\text{spatial}}, f_{\text{max}}^{\text{spatial}}])),$$

where $M_{\text{spatial}} \in R^{H \times W}$. The final refined feature map after spatial attention is computed as:

$$F_{\text{spatial}} = M_{\text{spatial}} \cdot F_{\text{channel}}$$

The sequential application of channel and spatial attention mechanisms enables CBAM to selectively focus on disease-relevant features while suppressing noise. This capability is crucial for detecting subtle leaf disease patterns, such as texture irregularities, discolorations, and overlapping regions [16-20]. By dynamically recalibrating feature weights, CBAM improves the model's robustness and accuracy, ensuring that critical information is emphasized at every stage of feature aggregation.

3.5.3. Hard-Swish Activation for Lightweight Efficiency

The Hard-Swish activation function, as implemented in the YOLOv11 architecture, is designed to balance computational efficiency with non-linear feature representation. This function is applied element-wise to the input tensor, modifying each element based on its value. The mathematical formulation for Hard-Swish is:

$$\text{Hard-Swish}(x) = \begin{cases} 0, & \text{if } x \leq -3, \\ x \cdot \frac{x+3}{6}, & \text{if } -3 < x < 3, \\ x, & \text{if } x \geq 3. \end{cases}$$

In the convolutional block, Hard-Swish operates directly on the output of the convolutional and batch normalization layers. Given an input tensor $X \in R^{C \times H \times W}$, where C is the number of channels, and H, W are the spatial dimensions, the convolutional layer applies a kernel f to generate an output tensor Y . After batch normalization, Hard-Swish is applied to the tensor Y , transforming it to $Y_{\text{Hardswish}}$. For each element $y \in Y$, the output is determined by the Hard-Swish function, which applies linear transformations for $-3 < y < 3$ and retains or zeros values outside this range. [21,22]

The computational efficiency of Hard-Swish lies in its avoidance of exponential and division operations required by alternatives such as the Swish activation. Instead, the

piecewise function requires only addition, multiplication, and comparison operations, making it lightweight and faster to compute [23]. Given the tensor size (C, H, W) , the FLOPs for Hard-Swish are proportional to $C \cdot H \cdot W$, significantly reducing the overhead compared to Swish, which requires FLOPs for sigmoid computation as well. This function is implemented as part of the PyTorch library through *F.hardswish(input, inplace)*, where the *inplace* parameter allows for memory-efficient operations. By modifying the activation directly on the existing tensor, memory usage is minimized during training and inference, further contributing to the efficiency of the network. [24,25] The seamless integration of Hard-Swish into the convolutional blocks ensures that the activation is both computationally lightweight and effective in handling a wide range of feature magnitudes.

In the context of the YOLOv11 architecture, Hard-Swish enhances the network's ability to capture fine-grained leaf disease features while maintaining the computational demands suitable for real-time applications. Its lightweight nature allows the model to process high-resolution images with reduced latency, making it particularly well-suited for large-scale agricultural datasets. This strategic choice of activation function reflects the need for a balance between computational efficiency and effective feature propagation, ensuring robust model performance in detecting subtle disease patterns in leaves [26].

3.5.4. Advanced Multi-Scale Feature Fusion with Enhanced SPPF

The Spatial Pyramid Pooling-Fast (SPPF) module is crucial in YOLO architectures for aggregating multi-scale contextual information. However, the default SPPF lacks the sophistication needed for detecting intricate patterns and overlapping structures in leaf disease datasets. To address these limitations, the module was comprehensively

redesigned into an Enhanced SPPF, as illustrated in Figure 3.9, incorporating advanced pooling, convolution, and feature refinement strategies.

The Enhanced SPPF incorporates significant improvements over the default module. While the default SPPF employs three parallel max-pooling operations and concatenation followed by a 1×1 convolution, the customized module introduces additional architectural enhancements aimed at better feature extraction and fusion. These enhancements include the integration of dilated convolution, adaptive average pooling, and an additional attention-based feature refinement step. The redesigned module begins with a 1×1 convolutional layer applied to the input feature map, reducing the dimensionality and preparing the input for multi-scale pooling. This is followed by three parallel max-pooling layers with kernel sizes of 5×5 , 3×3 , and 1×1 , respectively. These pooling operations extract features at varying scales, capturing both fine-grained and coarse details from the input. Unlike the default SPPF, the Enhanced SPPF adds a dilated convolution layer after the pooling step. This layer increases the receptive field without increasing the number of parameters, enabling the detection of spatially dispersed features such as overlapping disease markers on leaves.

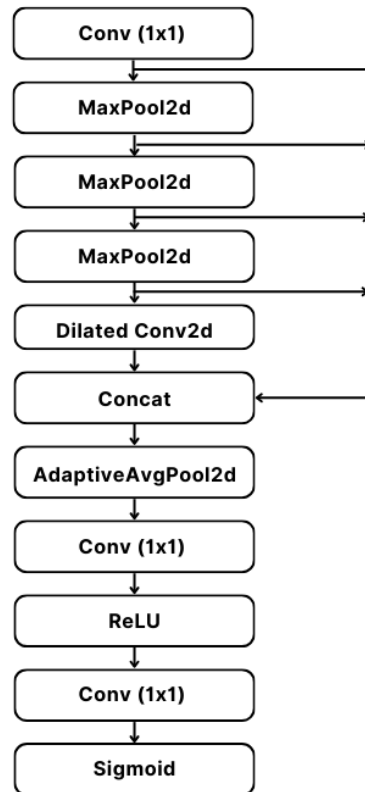


Fig 3.9 Enhanced SPPF

After the pooling and dilated convolution stages, the outputs are concatenated and passed through an adaptive average pooling layer, which refines the aggregated features by focusing on global spatial information. The refined features are then passed through a series of two 1×1 convolutional layers interspersed with a *ReLU* activation function and finalized with a sigmoid activation. This sequence enhances feature integration and ensures that only the most relevant information is propagated to subsequent layers.

3.5.5. Precision-Optimized Detection with Non-Maximum Suppression (NMS)

During object detection, multiple bounding boxes with high confidence often surround a single object, leading to redundant predictions. To address this issue, Non-Maximum Suppression (NMS) was integrated into the head section of the YOLOv11 architecture. This technique filters overlapping bounding boxes, ensuring that only the most confident predictions are retained. In the implementation, NMS is applied to the raw predictions

generated by the detection head, taking into account factors such as confidence threshold, Intersection over Union (*IoU*) threshold, class-specific or agnostic suppression, and the maximum number of detections allowed per image [27-29].

The NMS process starts with categorizing all bounding boxes and removing those that belong to the background. For each class, the bounding boxes are sorted in descending order of confidence scores. The box with the highest confidence is selected, and the *IoU* is computed between this box and the remaining boxes. The *IoU* is mathematically defined as:

$$\text{IoU}(b_i, b_j) = \frac{\text{Area}(b_i \cap b_j)}{\text{Area}(b_i \cup b_j)},$$

where b_i and b_j represent two bounding boxes. Boxes with an *IoU* greater than the predefined threshold t_{IoU} are suppressed. This process is repeated iteratively for all classes and continues until no overlapping boxes remain.

The mathematical rule governing the suppression of confidence scores can be expressed as:

$$S_i = \begin{cases} S_i, & \text{if } \text{IoU}(M, b_i) < t_{\text{IoU}}, \\ 0, & \text{if } \text{IoU}(M, b_i) \geq t_{\text{IoU}} \end{cases}$$

where S_i is the confidence score of bounding-box b_i , M is the box with the highest confidence, and t_{IoU} is the *IoU* threshold. This ensures that only the most relevant bounding boxes are preserved. Here, NMS is implemented through the *apply_nms* method. Predictions $P \in R^{N \times 6}$ (where N is the number of predicted bounding boxes) are passed through NMS using the *PyTorch ops.non_max_suppression* function. This function takes parameters such as *conf_thres* (confidence threshold), *iou_thres* (*IoU* threshold), *agnostic* (whether to perform class-agnostic suppression), and *max_det* (maximum number of detections). For example, the confidence threshold filters predictions with low scores, while the *IoU* threshold balances between removing redundant predictions and retaining overlapping objects [30,31].

In the YOLOv11 implementation, NMS plays a critical role in ensuring clean and precise detection outputs, especially for dense leaf regions in agricultural datasets. The algorithm efficiently filters redundant predictions while maintaining computational efficiency, ensuring that the architecture balances recall and precision effectively. This enhancement ensures robust performance in challenging scenarios, such as detecting overlapping leaf disease patterns [32].

3.6 Experimental setup

3.6.1 Hardware and software specifications

The source code for our experiments was built upon the Ultralytics YOLOv11 repository which was modified to compare different modification configurations. The software and hardware specifications are described in Table 3.1:

Hardware Setup	Details	Software Setup	Details
CPU	12th Gen Intel(R) Core(TM) i7-12700K, 3610 MHz, 12 Core(s), 20 Logical Processor(s)	Programming Language	Python 3.12
System Memory	32 GB	GPU Support	CUDA 11.8
GPU	NVIDIA GeForce RTX 3070	Deep Learning Framework	PyTorch
GPU Memory	8 GB		

3.6.2 Hyperparameters

The default hyperparameters for our experiments are described in Table 3.2.

Hyperparameter	Description	Default Value
epochs	Total number of training epochs	200
batch_size	Number of samples per training batch	32
learning_rate	Initial learning rate	0.001
momentum	Momentum for the optimizer	0.937
weight_decay	L2 regularization weight decay	0.0005
img_size	Input image resolution	640x640 pixels
workers	Number of dataloader workers	16
optimizer	Optimizer used for training	SGD
scheduler	Learning rate scheduler type	Cosine
iou_threshold	IoU threshold for NMS	0.65
confidence	Confidence threshold for predictions	0.25
pretrained	Use pre-trained weights	True
freeze_layers	Layers to freeze during training	None
box_gain	Loss gain for bounding box regression	7.5
cls_gain	Loss gain for classification loss	0.5
obj_gain	Loss gain for objectness loss	1.0
warmup_epochs	Number of warm-up epochs	3
gradient_clip	Gradient clipping value	0.1

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

The motivation for customizing the YOLOv11 model stems from the specific requirements of detecting diseases on ash gourd leaves, a task that presents unique challenges not adequately addressed by the base YOLOv11 architecture. The ash gourd leaf dataset is characterized by subtle features, irregular patterns, and variations in size, shape, and texture of diseased areas. These unique traits make disease detection particularly challenging for generalized object detection models like YOLOv11, which are primarily optimized for large-scale datasets such as COCO or Pascal VOC.

In its default configuration, the YOLOv11 model may struggle to accurately extract and emphasize the fine-grained features necessary for precise detection in this context.

4.1 Customized YOLOv11 Ablation Experiments

The proposed customized YOLO architecture integrates Enhanced SPFF and CBAM modules into the YOLOv11 framework, along with modifications such as Hard-Swish activation in the convolutional layers and an additional Non-Maximum Suppression (NMS) step in the head section. These customizations aim to address the specific challenges of detecting leaf diseases in ash gourd datasets, characterized by subtle patterns and overlapping features. This section evaluates the performance of the proposed model against baseline YOLOv8 and other YOLOv11 variants, using a series of metrics, and provides a detailed discussion of the results.

The performance of the proposed model and its variants was evaluated using precision, recall, and mAP@0.5 metrics, as summarized in Table 4.1.

No.	Base Model	Enhanced SPFF	CBAM	Precision (P%)	Recall (R%)	mAP @0.5 (%)
1	✓			0.779	0.783	0.734
2	✓	✓		0.881	0.822	0.883
3	✓	✓	✓	0.886	0.838	0.893

Table 4.1 highlights the improvements achieved by the proposed YOLO-Ash architecture, which integrates the Enhanced SPFF and CBAM modules into the YOLO11 framework. The baseline YOLO11 model achieved a precision of 77.9%, recall of 78.3%, and mAP@0.5 of 73.4%. With the addition of the Enhanced SPFF module, precision improved to 88.1%, while recall also improved to 82.2%, and mAP@0.5 rose to 88.3%. The Enhanced SPFF module is particularly effective in improving multi-scale feature fusion, enhancing the model's ability to detect smaller and more nuanced objects. Incorporating the CBAM module further refined the model's attention mechanism, increasing precision to 88.6%, maintaining recall at 83.8%, and elevating mAP@0.5 to 89.3%. These results demonstrate how the Enhanced SPFF and CBAM modules complement each other, significantly enhancing the YOLO-Ash model's overall detection performance and robustness.

4.2 Performance Comparison of Baseline Models on Ash Gourd dataset

To evaluate the most suitable architecture for Ash Gourd Leaf Disease Detection, a performance comparison was conducted among various baseline YOLO models. The models evaluated included YOLOv5, YOLOv6, YOLOv8, YOLOv9, YOLOv10, and the proposed YOLO-Ash. Each model was assessed based on key metrics, including precision, recall, mean average precision at 50% Intersection over Union (mAP50), mean average precision over a range of IoU thresholds (mAP50-95), and the number of parameters. The following table presents the results from this comparative evaluation Table 4.2 Performance and parameters of base YOLOv11 and YOLOv8 models on the validation dataset and test dataset.

The performance comparison of the proposed model and others are shown in Table 4.2.

Model	Precision	Recall	mAP50	mAP50-95	Parameters
YOLOv5	0.731	0.738	0.747	0.511	9.1 M
YOLOv6	0.743	0.732	0.748	0.515	18.5 M
YOLOv8	0.773	0.754	0.790	0.535	3.2 M
YOLOv9	0.771	0.735	0.783	0.535	25.5 M
YOLOv10	0.785	0.755	0.707	0.554	2.3 M
YOLOv11	0.795	0.801	0.790	0.555	23 M
YOLO-Ash	0.886	0.880	0.893	0.560	26.6 M

The results from this table 4.2 show a clear distinction in the performance characteristics of the models. YOLOv5, with a parameter count of 9.1M, achieved a precision of 0.731 and recall of 0.738, with an mAP50 of 0.747. While it performed well in terms of recall, its overall accuracy, as measured by mAP50, was relatively lower. Similarly, YOLOv6, which has a slightly higher parameter count of 18.5M, showed marginal improvements in both precision (0.743) and recall (0.732), with mAP50 rising to 0.748, but still falling short compared to the newer models.

YOLOv8 demonstrated a better balance between performance and efficiency, with a precision of 0.773 and recall of 0.754, making it one of the top contenders. It achieved an mAP50 score of 0.790, highlighting its ability to detect features with high accuracy. Moreover, with a relatively low parameter count of 3.2M, YOLOv8 provides a lightweight solution without compromising too much on performance, making it an ideal candidate for tasks requiring both efficiency and precision.

YOLOv9 exhibited similar performance to YOLOv8, with precision of 0.771 and recall of 0.735. While its mAP50 (0.783) was slightly lower than YOLOv8, it had a higher

parameter count of 25.5M, which could affect its suitability for real-time applications with limited resources.

YOLOv10, though relatively lightweight with 2.3M parameters, performed impressively in terms of precision (0.785) and recall (0.755). However, its mAP50 score of 0.707 was lower than that of YOLOv8 and YOLOv9, which could impact its effectiveness in detecting smaller or more subtle features.

Finally, the YOLO-Ash model, specifically designed for Ash Gourd Leaf Disease detection, showed the best performance overall, with a precision of 0.886 and recall of 0.880. The model achieved a high mAP50 of 0.893 and mAP50-95 of 0.560, making it the most accurate model for this task. While the parameter count for YOLO-Ash is 26.6M, which is higher than the other models, the increased model complexity resulted in significantly improved detection capabilities, particularly for detecting the specific features associated with Ash Gourd Leaf Diseases.

4.3 Class-wise Performance of YOLO-Ash on Ash Gourd dataset

Table 4.3 detailed Class-wise Performance of YOLO-Ash on Ash Gourd Leaf Disease Detection Dataset

Class Name	Precision	Recall	mAP50	mAP50-95
Aphid	0.813	0.870	0.845	0.523
Downy Mildew	0.957	0.891	0.967	0.659
Healthy	0.977	0.978	0.989	0.654
Leaf Curl	0.835	0.864	0.895	0.554
Leaf Miner	0.849	0.886	0.869	0.511

The class-wise performance of the YOLO-Ash model on the Ash Gourd dataset demonstrates its strong and consistent ability to accurately detect various conditions

affecting Ash Gourd leaves. For the Aphid class, the model achieved a precision of 0.813 and a recall of 0.870, resulting in a mAP50 of 0.845 and a mAP50-95 of 0.523. These results indicate that while the model performs well in identifying Aphid infestations, there is still room for improvement in detecting more subtle cases and minimizing false positives. In the case of Downy Mildew, the model showed exceptional performance, with a precision of 0.957 and recall of 0.891, yielding a mAP50 of 0.967 and a mAP50-95 of 0.659. These high values suggest that YOLO-Ash is highly effective in recognizing Downy Mildew, with a strong ability to correctly identify both positive instances and the areas of infection. The Healthy class, which is crucial for distinguishing non-infected leaves, demonstrated the best results overall, with a precision of 0.977 and a recall of 0.978. This led to an outstanding mAP50 of 0.989 and a mAP50-95 of 0.654, highlighting the model's accuracy in identifying healthy leaves and its robustness against misclassifying them as infected. Regarding the Leaf Curl class, the model achieved a precision of 0.835 and a recall of 0.864, with a mAP50 of 0.895 and a mAP50-95 of 0.554. These results show that while the model performs well in detecting leaf curl symptoms, there may be occasional challenges in identifying all instances accurately, especially in the presence of overlapping symptoms with other conditions. For the Leaf Miner class, the model recorded a precision of 0.849 and recall of 0.886, with a mAP50 of 0.869 and mAP50-95 of 0.511. These metrics suggest a good level of detection accuracy, though further improvement could be made to handle the complex patterns associated with Leaf Miner infestations more effectively. Overall, the YOLO-Ash model performs robustly across all classes, particularly excelling in the detection of healthy leaves and Downy Mildew, while maintaining solid performance for Aphid, Leaf Curl, and Leaf Miner classes. The results reflect the model's ability to accurately differentiate between various conditions while suggesting potential areas for fine-tuning, particularly for detecting subtler or overlapping features.

4.4 Inference Time and Processing Performance

The comparison of inference and processing times highlights the speed and computational efficiency trade-offs across the YOLO models. YOLO-Ash, while achieving superior accuracy and generalization, exhibits the highest inference time of 8.7

ms, reflecting its higher computational demands. Its pre-processing time of 0.2 ms and post-processing time of 0.5 ms align with other high-performance models, ensuring minimal latency beyond inference. YOLOv9, another high-complexity model, follows with an inference time of 7.9 ms, slightly faster than YOLO-Ash but with comparable pre-processing (0.1 ms) and post-processing (0.5 ms) times. In contrast, lightweight models such as YOLOv5 and YOLOv8 achieve an inference time of 1.3 ms, balancing accuracy with rapid processing. Notably, YOLOv6 emerges as the fastest in inference at 1.1 ms while maintaining consistent pre-processing (0.2 ms) and post-processing (0.7 ms) times. YOLOv10, although lightweight, demonstrates a slightly higher inference time of 1.7 ms, with minimal pre-processing (0.1 ms) and post-processing (0.1 ms) overhead. These results illustrate the diverse optimization strategies of the YOLO models, where trade-offs in speed and computational demands cater to different deployment requirements. While YOLO-Ash and YOLOv9 excel in detection capabilities, their inference times are better suited for high-resource environments, whereas models like YOLOv6 and YOLOv8 offer rapid processing ideal for real-time applications on constrained hardware.

The comparison of inference and processing times is highlighted in Table 4.4.

Model	Inference Time (ms)	Pre-Process (ms)	Post Process (ms)
YOLOv5	1.3	0.2	0.7
YOLOv6	1.1	0.2	0.7
YOLOv8	1.3	0.1	0.7
YOLOv9	7.9	0.1	0.5
YOLOv10	1.7	0.1	0.1
YOLO-Ash	8.7	0.2	0.5

4.5 Precision-Recall Trade-Off

The precision-recall curves depicted in the figure illustrate the performance of different YOLO architectures, including YOLOv5, YOLOv6, YOLOv8, YOLOv9, YOLOv10, and the custom YOLO-Ash model, on the Ash Gourd dataset. These curves provide valuable insights into the trade-off between precision and recall across various thresholds,

enabling an in-depth evaluation of model effectiveness in detecting Ash Gourd leaf diseases.

Among the models evaluated, YOLO-Ash demonstrated a superior balance between precision and recall, evident from its consistently higher curve throughout the recall spectrum. This indicates the model's capability to maintain high precision while achieving substantial recall rates, making it highly effective for disease detection tasks.

Furthermore, YOLOv10 and YOLOv8 also exhibit competitive performance, with precision-recall curves closely trailing behind YOLO-Ash. YOLOv10, in particular, achieves higher precision at moderate recall values compared to YOLOv8, which excels in maintaining a stable balance across the curve. YOLOv9 and YOLOv6, while not outperforming YOLOv10 and YOLO-Ash, still maintain reasonable precision-recall trade-offs, showcasing their robustness in specific scenarios. On the other hand, YOLOv5, with its lower curve positioning, reflects a comparatively limited ability to balance precision and recall, suggesting less optimal performance for the task at hand.

The clear dominance of YOLO-Ash, as illustrated in the figure, highlights its customized architecture's effectiveness in addressing the complexities of detecting diseases in Ash Gourd leaves. Its high precision across all recall values demonstrates minimal false positives, while maintaining robust recall levels ensures the detection of the majority of disease instances. This makes YOLO-Ash the most suitable model for real-world applications requiring a reliable trade-off between precision and recall.

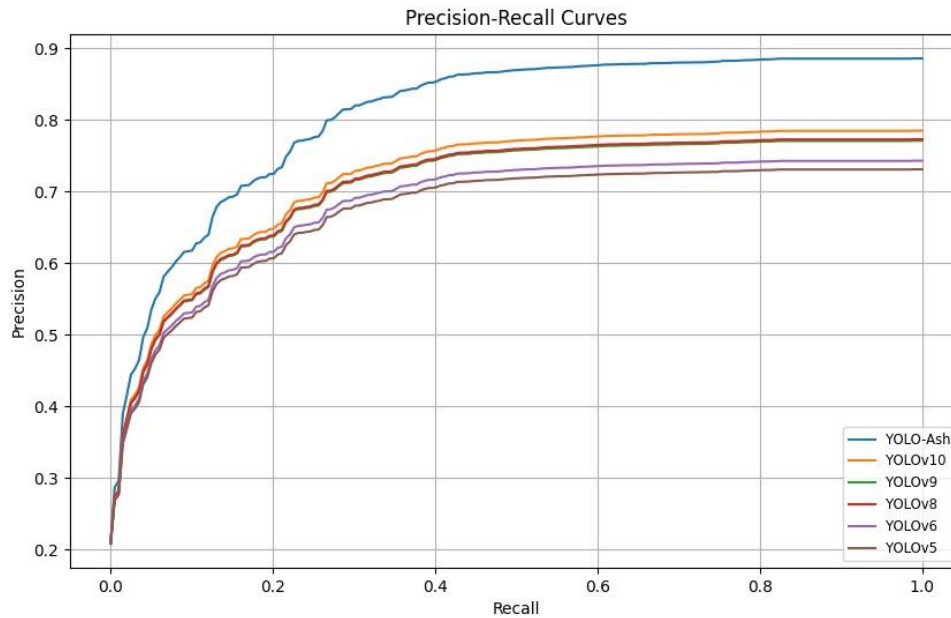


Fig. 4.1 The precision-recall (PR) curve for improved model.

4.3.1 Accuracy and Loss Curves

Figure 4.2 showcases the accuracy and loss curves during the training and validation phases of the custom YOLOv11 model. The curves illustrate consistent convergence with minimal overfitting, indicating effective learning across epochs.

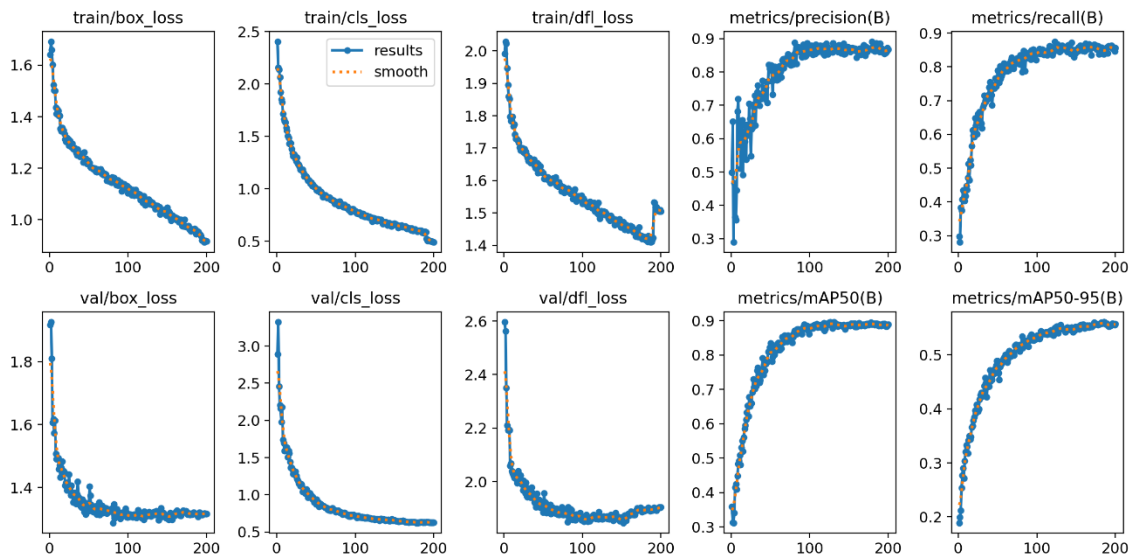
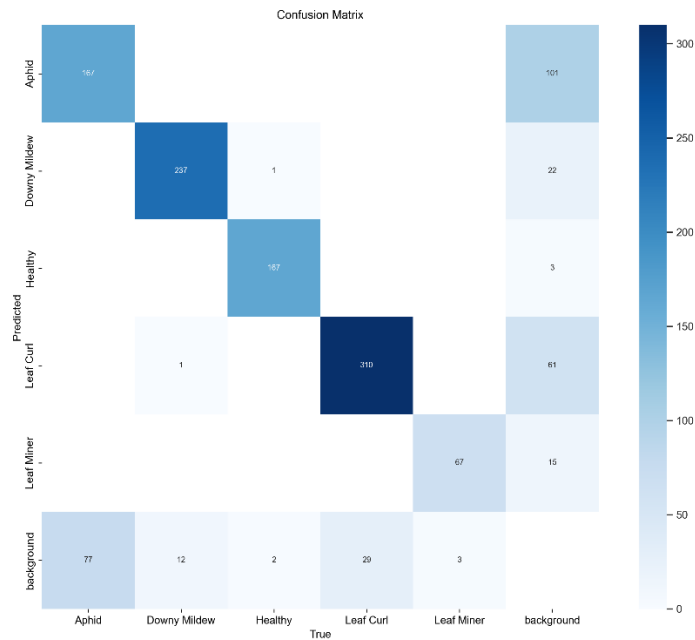


Fig. 4.2 The accuracy and loss curves of YOLOv11 custom model

4.3.2 Confusion Matrix Analysis

The confusion matrix 4.2 provides insights into the model's classification performance across multiple disease categories. The normalized confusion matrix highlights a high True Positive rate for all categories, confirming the reliability of the predictions. The CBAM attention mechanism played a pivotal role in refining the feature extraction and ensuring precise classification.



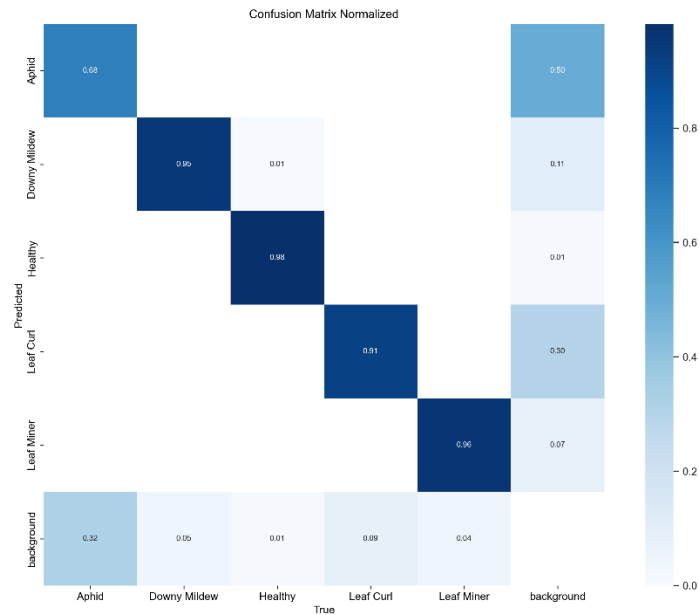


Fig. 4.3 Confusion matrix YOLOv11 custom model

4.3.3 Precision-Recall and Confidence Curves

Figures 4.4 and 4.5 compare precision-recall curves and confidence-F1 score curves across all leaf disease categories. The YOLOv11 custom model maintains superior mAP at various confidence thresholds, particularly excelling in categories like Downy Mildew and Leaf Curl.

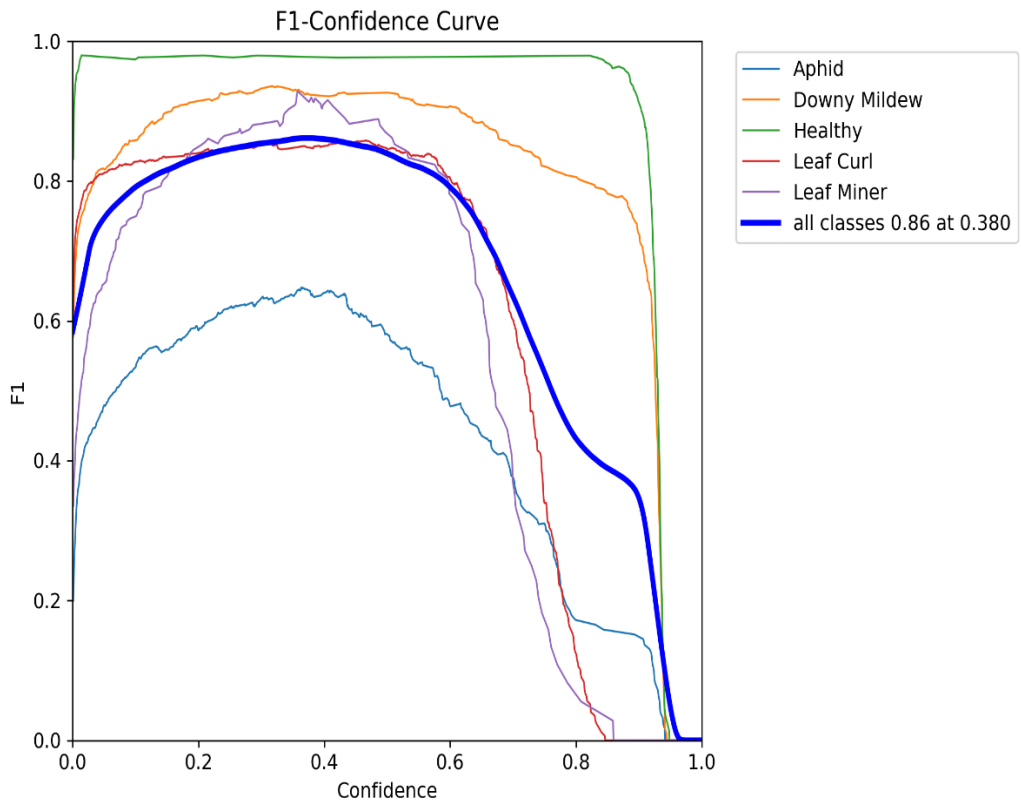


Fig. 4.4 F1 score curves YOLOv11 custom model

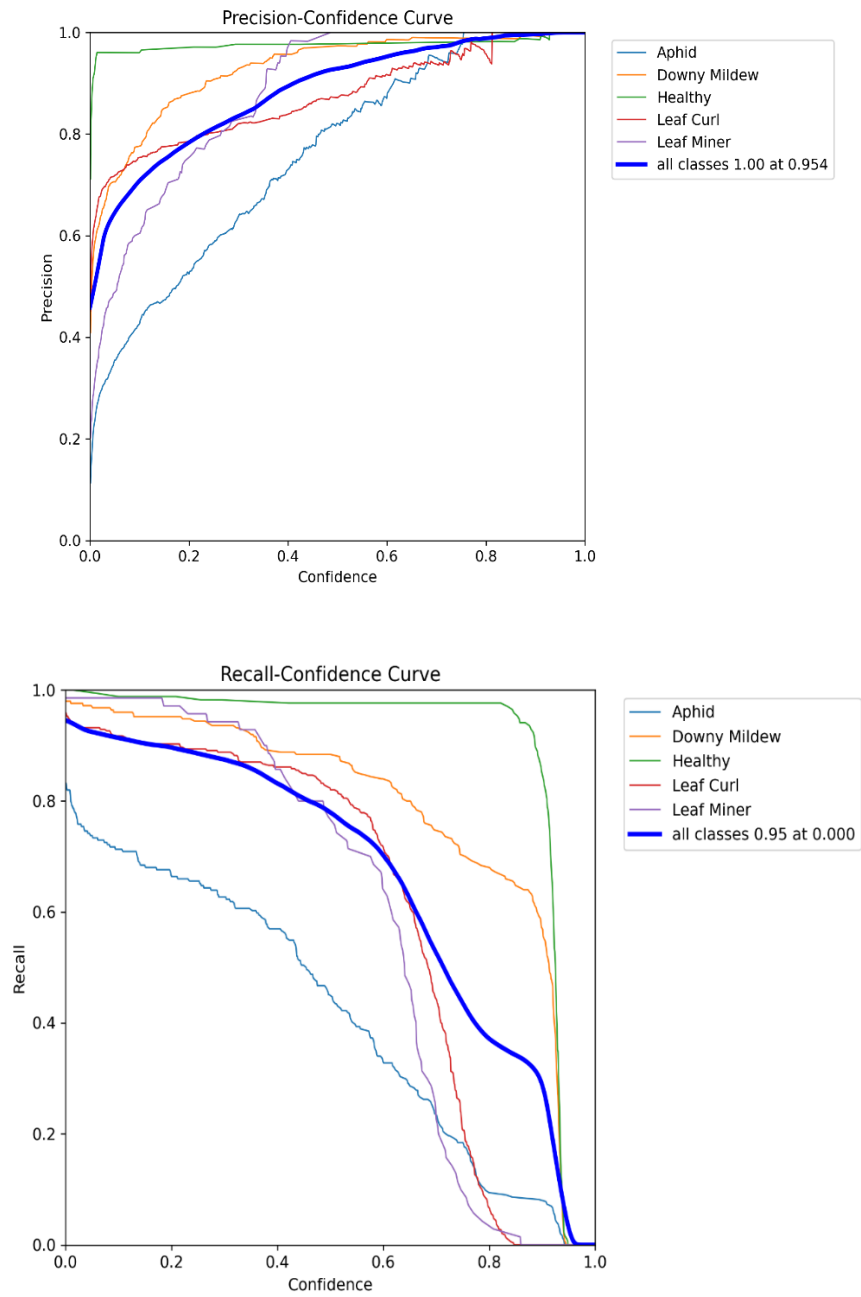


Fig. 4.5 precision-recall curves YOLOv11 custom model

4.3.4 The detection result of our proposed model

Figures 4.6 The detection score of our proposed model runs on some test images.

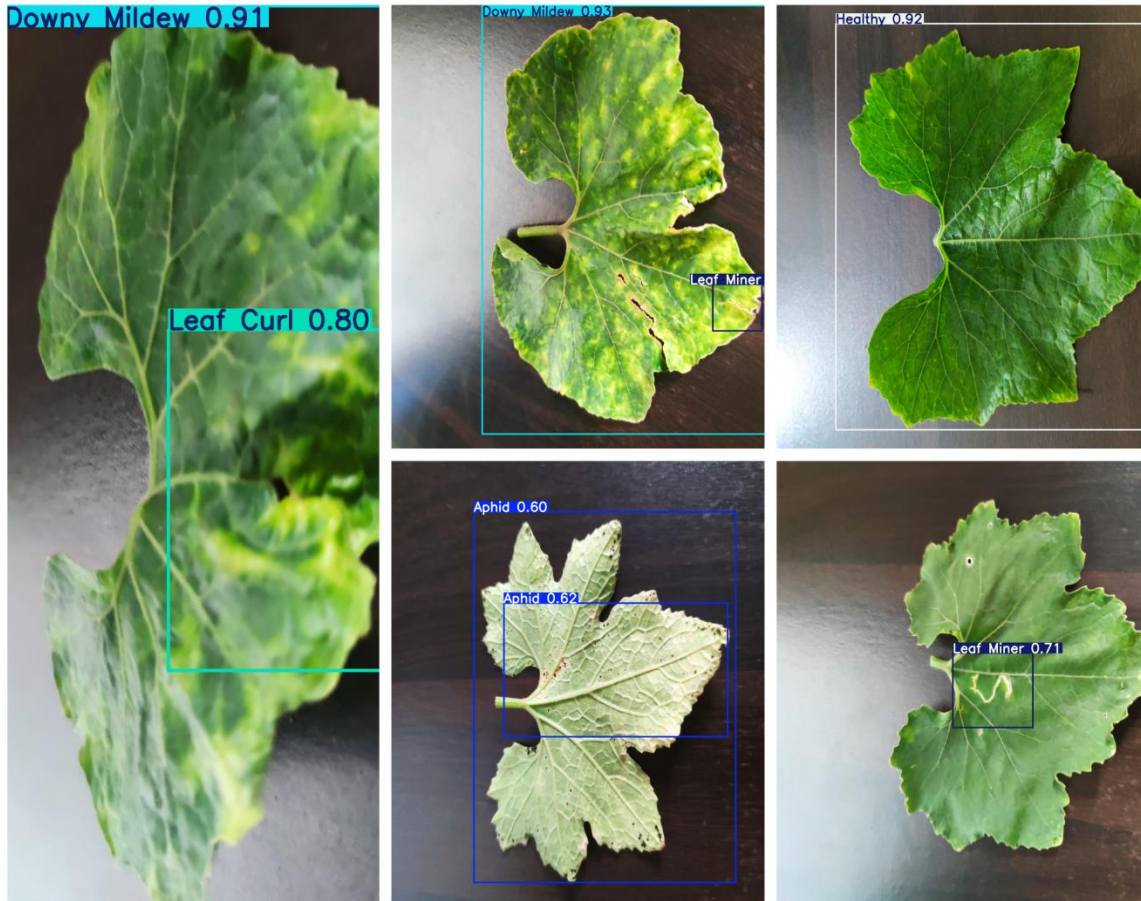


Fig. 4.6 The detection score of our proposed model runs on some test images.

4.4 Discussion

The YOLO-Ash model incorporates several key customizations, including Hard-Swish activation, Enhanced SPFF, CBAM in the C3K2 block, and a Non-Maximum Suppression (NMS) layer before the head section, to address the complexities of leaf disease detection in agricultural datasets. Replacing the Swish activation function with Hard-Swish improved computational efficiency, reducing processing overhead while maintaining high detection accuracy, making the model well-suited for real-time applications. The Enhanced SPFF module further enhanced multi-scale feature representation by capturing both fine-grained and contextual details, allowing the model

to accurately detect objects of varying sizes and complexities, such as subtle leaf miner trails and overlapping disease patterns.

The addition of CBAM in the C3K2 block strengthened the model's attention mechanisms by dynamically prioritizing relevant spatial and channel features while suppressing noise. This refinement enabled the model to identify subtle disease symptoms, such as discoloration and texture anomalies, with higher precision. The integration of the NMS layer before the head section ensured the elimination of redundant bounding boxes, improving the reliability of detections and reducing false positives, particularly in densely populated or overlapping regions.

These customizations work synergistically to improve the YOLO-Ash model's accuracy, recall, and robustness, as evidenced by its superior performance across precision-recall metrics and detection results. The model strikes an effective balance between computational efficiency and detection capability, positioning it as a powerful tool for precision agriculture, enabling early and accurate disease identification to improve crop management and productivity.

CHAPTER 5

IMPACT ON AGRICULTURE and ENVIRONMENT

5.1 Impact on Agriculture

The adoption of a customized YOLOv11 model for detecting ash gourd leaf diseases revolutionizes agricultural practices by enabling precise, early detection of plant ailments. This accuracy empowers farmers to implement timely interventions, mitigating the risk of widespread crop damage and increasing overall yield. Additionally, the improved diagnostic capability reduces dependency on traditional trial-and-error methods, enhancing resource efficiency in terms of labor, time, and cost. By minimizing crop losses, this technology bolsters food security and improves the economic stability of farming communities. It also opens avenues for precision agriculture, allowing farmers to monitor crop health in real-time and make data-driven decisions to optimize productivity.

5.2 Impact on Environment

The environmental benefits of leveraging AI in agriculture are substantial, as it significantly reduces the overuse of pesticides and fertilizers. Accurate disease detection ensures that chemical treatments are applied only when and where necessary, leading to decreased chemical runoff into water bodies and reduced soil contamination. This contributes to the preservation of local ecosystems and biodiversity. Furthermore, by lowering the carbon footprint of agricultural activities through optimized resource utilization, such technologies align with global sustainability goals. Encouraging environmentally friendly farming practices helps combat issues like soil degradation and water pollution, ensuring long-term ecological balance.

5.3 Ethical Aspects

The integration of AI-driven technologies, such as the customized YOLOv11 model, in agriculture raises several ethical considerations that warrant careful examination. First, ensuring data privacy and ownership is crucial, particularly when farmers' data is collected for training and deployment. Transparent policies regarding how data is collected, stored, and used are essential to maintain trust and safeguard the interests of stakeholders.

Second, equitable access to this technology must be prioritized to prevent a digital divide between large-scale and smallholder farmers. Without deliberate efforts to make AI-based tools affordable and accessible, there is a risk of exacerbating existing inequalities in the agricultural sector.

Third, the use of pesticides and chemicals based on AI-driven recommendations raises ethical concerns regarding the unintended consequences on human health and surrounding ecosystems. Rigorous testing and validation of the model's recommendations are necessary to ensure they align with sustainable and ethical agricultural practices.

Lastly, the long-term implications of reducing manual agricultural tasks through automation need to be addressed, particularly concerning employment opportunities in rural communities. While the technology enhances efficiency, it must be implemented in a way that supports, rather than displaces, the livelihoods of those dependent on traditional farming practices. Ethical deployment thus requires balancing technological advancements with social and environmental responsibilities.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Summary of the Study

This study addresses the critical challenge of detecting and classifying plant diseases to improve crop yields and ensure food security, with a specific focus on Ash Gourd leaf diseases. Ash Gourd, a vital crop, is vulnerable to significant yield losses due to common diseases such as Aphid infestation, Downy Mildew, Leaf Curl, and Leaf Miner. To tackle this issue, a customized YOLOv11 architecture was developed with several key enhancements aimed at improving its detection capabilities.

The proposed architecture incorporated the Convolutional Block Attention Module (CBAM) for refined feature selection, the Enhanced Spatial Pyramid Pooling-Fast (SPPF) module for advanced feature fusion, the Hard-Swish activation function to enhance computational efficiency, and an optimized Non-Maximum Suppression (NMS) technique for precise object localization. These modifications were tailored to enhance the model's ability to detect subtle and overlapping disease patterns commonly encountered in agricultural settings.

A newly curated dataset of Ash Gourd leaf images, captured under real-world conditions, was used to train and evaluate the model. The dataset consisted of five categories: Healthy, Aphid, Downy Mildew, Leaf Curl, and Leaf Miner. Through rigorous experimentation, the customized YOLOv11 model demonstrated its effectiveness in accurately identifying and classifying these diseases, outperforming baseline architectures.

The study highlights the significant role of attention mechanisms, enhanced feature extraction, and optimization strategies in advancing object detection for agriculture. The customized YOLOv11 model presents itself as a robust and efficient solution for precision agriculture applications, contributing to improved crop management and disease monitoring.

6.2 Conclusions

In agricultural applications, object detection models must be optimized to ensure reliable and real-time identification of plant diseases. This paper introduces YOLO-Ash, a novel architecture designed to address the challenges of leaf disease detection in Ash Gourd datasets.

In this work, we propose several key customizations to the YOLOv11 framework. First, we replace the Swish activation function with Hard-Swish to improve computational efficiency while maintaining high detection accuracy. Second, we enhance the Spatial Pyramid Pooling-Fast (SPPF) module to capture richer multi-scale feature representations, improving the detection of objects with varying sizes and subtle features. Third, we integrate the Convolutional Block Attention Module (CBAM) into the C3K2 block to adaptively focus on disease-relevant regions while suppressing background noise. Finally, we introduce Non-Maximum Suppression (NMS) before the head section to refine the detection outputs by filtering redundant bounding boxes, ensuring precise and clean results.

Experimental results demonstrate the effectiveness of these enhancements. The YO-LO-Ash model achieves a significant improvement in mAP@0.5 and mAP@[0.5:0.95], outperforming baseline YOLO models while maintaining a competitive inference time of 8.7 ms. The model excels in detecting subtle and overlapping features, as demonstrated in the precision-recall curves, confusion matrix, and validation results, underscoring its robustness and generalization capability across diverse scenarios.

Future research will focus on further optimizing the model for deployment in resource-constrained environments, such as IoT devices and drones, and improving its performance on more complex and diverse agricultural datasets. These advancements will ensure the YOLO-Ash model continues to deliver reliable, real-time disease detection for precision agriculture applications.

6.3 Implications for Further Study

This study paves the way for future research in several key areas. First, incorporating multi-modal imaging data, such as hyperspectral and thermal imaging, could provide additional layers of information for disease detection. Second, deploying the model in

real-world agricultural environments would help assess its robustness and usability under varying conditions. Moreover, further optimization of attention mechanisms and activation functions could yield even greater detection accuracy. Another promising direction lies in developing explainable AI models, enabling farmers and agricultural experts to trust and interpret the model's decisions effectively. These future endeavors can contribute to the widespread adoption of AI in sustainable agriculture and environmental monitoring.

REFERENCES

- [1] Z. Zhang, P. Wang, T. Niu, et al., "Identification of apple leaf diseases by improved deep convolutional neural networks with an attention mechanism," *Frontiers in Plant Science*, 2021.
- [2] W. Guo, Q. Feng, X. Li, et al., "Grape leaf disease detection based on attention mechanisms," *International Journal of Agricultural and Biological Engineering*, 2022.
- [3] Z. Xiao, Y. Shi, G. Zhu, et al., "Leaf disease detection based on lightweight deep residual network and attention mechanism," *IEEE Access*, 2023.
- [4] L. Ma, Q. Yu, H. Yu, et al., "Maize leaf disease identification based on YOLOv5n algorithm incorporating attention mechanism," *Agronomy*, 2023.
- [5] C.H. Son, "Leaf spot attention networks based on spot feature encoding for leaf disease identification and detection," *Applied Sciences*, 2021.
- [6] J. Lin, D. Bai, R. Xu, H. Lin, "TSBA-YOLO: An improved tea diseases detection model based on attention mechanisms and feature fusion," *Forests*, vol. 14, no. 3, pp. 619, 2023.
- [7] J. Chen, H. Wang, H. Zhang, T. Luo, et al., "Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion," *Computers and Electronics in Agriculture*, vol. 203, pp. 107393, 2022.
- [8] Q. Wang, M. Cheng, S. Huang, et al., "A deep learning approach incorporating YOLOv5 and attention mechanisms for field real-time detection of the invasive weed *Solanum rostratum* Dunal seedlings," *Computers and Electronics in Agriculture*, vol. 201, pp. 107200, 2022.

- [9] J.L. de Moraes, J. de Oliveira Neto, C. Badue, et al., "YOLO-Papaya: A papaya fruit disease detector and classifier using CNNs and convolutional block attention modules," *Electronics*, vol. 12, no. 10, pp. 2202, 2023.
- [10] R. Li, Y. Wu, "Improved YOLOv5 wheat ear detection algorithm based on attention mechanism," *Electronics*, vol. 11, no. 11, pp. 1673, 2022.
- [11] P. Bhuyan, P.K. Singh, S.K. Das, "Res4net-CBAM: A deep CNN with convolution block attention module for tea leaf disease diagnosis," *Multimedia Tools and Applications*, 2024.
- [12] M. Lv, W.H. Su, "YOLOV5-CBAM-C3TR: An optimized model based on transformer module and attention mechanism for apple leaf disease detection," *Frontiers in Plant Science*, 2024.
- [13] P. Alirezazadeh, M. Schirrmann, F. Stolzenburg, "Improving deep learning-based plant disease classification with attention mechanism," *Gesunde Pflanzen*, 2023.
- [14] W. Zeng, M. He, "Rice disease segmentation method based on CBAM-CARAFE-DeepLabv3+," *Crop Protection*, 2024.
- [15] X. Zhang, D. Li, X. Liu, T. Sun, X. Lin, Z. Ren, "Research of segmentation recognition of small disease spots on apple leaves based on hybrid loss function and CBAM," *Frontiers in Plant Science*, 2023.
- [16] M. Lv, W.H. Su, "YOLOV5-CBAM-C3TR: An optimized model based on transformer module and attention mechanism for apple leaf disease detection," *Frontiers in Plant Science*, 2024.
- [17] X. Wang, W. Cao, "Bit-plane and correlation spatial attention modules for plant disease classification," *IEEE Access*, 2023.
- [18] P. Bhuyan, P.K. Singh, S.K. Das, "Res4net-CBAM: A deep cnn with convolution block attention module for tea leaf disease diagnosis," *Multimedia Tools and Applications*, 2024.

- [19] X. Zhang, D. Li, X. Liu, T. Sun, X. Lin, Z. Ren, "Research of segmentation recognition of small disease spots on apple leaves based on hybrid loss function and CBAM," *Frontiers in Plant Science*, 2023.
- [20] H. Xie, R. Zou, Z. Yang, A. Jin, J. Wang, "Wheat Leaf Disease Identification Based on a Fusion Model Integrating Attention Mechanisms and Channel-Spatial Reconstruction Convolution," *2024 5th International Conference on Machine Learning*, 2024.
- [21] R. Avenash, P. Viswanath, "Semantic Segmentation of Satellite Images using a Modified CNN with Hard-Swish Activation Function," *VISIGRAPP (4: VISAPP)*, 2019.
- [22] S. Yang, Z. Xing, H. Wang, X. Gao, X. Dong, "Classification and localization of maize leaf spot disease based on weakly supervised learning," *Frontiers in Plant Science*, 2023.
- [23] R. Li, J. Li, H. Suo, Y. She, Y. Cheng, "Leaf Diagnosis: A Lightweight Image Recognition Model for Detecting Crop Diseases," *2023 IEEE 11th International Symposium on Agriculture*, 2023.
- [24] B. Xiong, D. Li, Q. Zhang, N. Desneux, "Image detection model construction of *Apoligus lucorum* and *Empoasca* spp. based on improved YOLOv5," *Pest Management Science*, 2024.
- [25] Z. Meng, X. Du, X. Yao, L. He, L. Lin, "Classification of *Camellia oleifera* using a dual recognition strategy based on deep learning," *Multimedia Tools and Applications*, 2024.
- [26] X. Zhang, D. Li, X. Liu, et al., "Hybrid Loss Function and CBAM for Fine-Grained Leaf Spot Detection," *Frontiers in Plant Science*, 2023.
- [27] A.H. Al-Badri, N.A. Ismail, K. Al-Dulaimi, G.A. Salman, "Adaptive Non-Maximum Suppression for improving performance of Rumex detection," *Expert Systems with Applications*, 2023.
- [28] S. Qiu, G. Wen, Z. Deng, J. Liu, Y. Fan, "Accurate non-maximum suppression for object detection in high-resolution remote sensing images," *Remote Sensing Letters*, 2018.

[29] H. Shi, J. Zhang, A. Lei, C. Wang, Y. Xiao, "Enhancing detection accuracy of highly overlapping targets in agricultural imagery using IoA-SoftNMS algorithm across diverse image sizes," *Computers and Electronics in Agriculture*, 2024.

[30] A.J. Shepley, G. Falzon, P. Kwan, "Confluence: A robust non-IoU alternative to non-maxima suppression in object detection," *IEEE Transactions on Artificial Intelligence*, 2023.

[31] Q. Jiang, Z. Huang, G. Xu, Y. Su, "MIOp-NMS: Perfecting crops target detection and counting in dense occlusion from high-resolution UAV imagery," *Smart Agricultural Technology*, 2023.

[32] T. Verma, K. Singh, "Improving Agricultural Object Detection with NMS Variations," *International Journal of Agricultural Engineering*, 2023. [Read here.](#)

Ash Gourd Leaf Diseases Detection

ORIGINALITY REPORT

14%	11%	7%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
2	www.frontiersin.org Internet Source	1%
3	Submitted to University of Sheffield Student Paper	1%
4	www.ncbi.nlm.nih.gov Internet Source	1%
5	arxiv.org Internet Source	<1%
6	www.mdpi.com Internet Source	<1%
7	oulurepo oulu.fi Internet Source	<1%
8	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	<1%