Enhancing Plant Disease Classification through Unsupervised Domain Adaptation (UDA)

MSc Research Project
MSCAI1B

Raj Nadigottu
Student ID: x23319658

School of Computing
National College of Ireland

Supervisor: Lavish Thomas

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Raj Nadigottu		
Student ID:	23319658		
Programme:	MSCAI1B	**Year:**	2024-25
Module:	Practicum		
Supervisor:	Lavish Thomas		
Submission Due Date:	11/08/2025		
Project Title:	**Enhancing Plant Disease Classification through Unsupervised Domain Adaptation (UDA)**		
Word Count:	5971 words	**Page Count**	11

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Raj Nadigottu

Date: 11/08/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Plant Disease Classification through Unsupervised Domain Adaptation (UDA)

Raj Nadigottu(x23319658)

Abstract

Classification models of plant diseases have been proven to show significant accuracy when trained and tested within a well-controlled laboratory environment but show poor performance when tested on real farming environments, where there is a domain mismatch between clean lab and field images. In this paper, we examine how Unsupervised Domain Adaptation (UDA) methods may be used to overcome this gap in order to achieve generalization with no need of labelled target-domain data. With the PlantVillage dataset, as the source domain and PlantDoc as the target domain, the proposed pipeline combines three components: CycleGAN used to do the visual translation of the domain, SimCLR that learns feature representations in a self-supervised fashion, and pseudo-labelling, which uses a back-and-forth arrangement to incorporate the target-domain samples with high confidence.

MobileNetV2 classifier has been trained on the original and adapted datasets, and then tested it on a separate set of PlantDoc images. Some of the UDA approaches produced better results than the baseline, but the improvements were uneven. In the final evaluation, the model achieved 23.08% accuracy and a macro F1 score of 0.1628. These outcomes show that, while UDA offers potential benefits for agricultural image classification, it also has clear limitations. Factors such as uneven class distribution, the complexity of the datasets, and the presence of noisy labels appeared to play a role in restricting performance.

1. Introduction

Plant diseases still pose a significant challenge to food production around the world, and there is a risk of food insecurity. It is critical that the disease is detected early and accurately to curb its spread to facilitate intervention early enough. When applied in a controlled laboratory environment where images are captured with uniform lighting and background, deep learning models especially convolutional neural networks (CNNs) have been found to perform very well. But, when transferred to practical applications, their output is frequently reduced as there is a dominant domain shift between training and deployment data brought about by the difference in image quality and lighting, background clutter, environment, etc.

Research Question:

How can unsupervised domain adaptation (UDA) techniques improve AI-based plant disease classification, enabling models to generalize effectively across diverse environmental conditions without requiring labeled real-world data?

The field of Unsupervised Domain Adaptation (UDA) has shown great potential in being able to minimize this domain shift. The goal of UDA methods is to make models trained on a labelled source dataset effective in a target, data-poor domain by modifying it (usually to match source distributions, or shift image styles). This is especially applicable in the agricultural setting where high-quality labelled datasets tend to be provided by controlled settings (e.g, PlantVillage), whereas real-world datasets (e.g., PlantDoc) tend to be small, noisy, and unlabelled.

This research investigates a complete UDA pipeline for plant disease classification that combines three key components:

1. **CycleGAN** for translating source images into a style closer to the target domain, reducing visual discrepancies.

2. **SimCLR** for contrastive learning to improve feature representation from unlabelled target domain images.

3. **Pseudo-labelling** to iteratively include high-confidence target domain predictions in training.

The aim is to check whether mixing these methods can aid classification of real-world plant disease images with better performance than that of a baseline model trained only in the source domain. Although some performance improvement is commonly replicated in the literature due to domain adaptation, it is also understood that performance gains are still small, or sometimes, even no improvement is reported in the literature, considering the nature of the dataset, imbalance of classes, and the power of visual complexity.

This work approaches the problem from both an improvement-oriented and diagnostic perspective:

- To assess whether the proposed UDA pipeline improves classification accuracy and F1 score on the target dataset.

- To analyse the conditions under which adaptation methods succeed or fail, contributing insights into the challenges of applying UDA for agricultural image classification.

Report Structure:
In Section 2, the literature review will be presented, in which the proposed work will be placed in the context of the already existing knowledge on the detection of plant diseases and domain adaptation. The methodology of the conducted research is described in Section 3, against the background of datasets, architectures of neural models, and experimental designs. Sections 4&5 give an explanation of the design requirements and the construction of the suggested UDA pipeline. The results and evaluation are provided in Section 6, and it analyse the quantitative measures along with the qualitative observations. Section 7 is the conclusion and future work, giving an overview of the main findings and determining possible directions forward and advancements as a result of the study.

2. Literature Review

2.1 Introduction to Domain Shift in Plant Disease Classification

Classification models of plant diseases tend to be very sensitive to similarity in domains of training and testing data. In more controlled conditions, i.e., the ones depicted in datasets like PlantVillage, it is unlikely to find any variation in the light, fixed backgrounds, and low noise (Mohanty et al., 2016). This makes the accuracy of classification high at the training and validation stages. However, in practice, using these models in practice in agricultural environments where there can be varying lighting conditions and also where the background is more complex and leaves may be partly occluded, their quality can be substantially reduced. It is referred to as domain shift (Wang and Deng, 2018; Kouw and Loog, 2019).

In plant pathology, the object of domain shift is on differences in the quality of such individual images, background complexity, type of crop, disease symptoms, and environmental conditions between the source domain (lab-based datasets) and target domain (field-based datasets). Researchers have long pointed out that the performance degrades when translation is conducted between two datasets instantiated under different domains, and that the accuracy can decrease by over 30 percentage points in the worst case (such as in Fang et al., 2023; Kayal et al., 2019).

Technologically, domain shift is a key concern as far as agricultural is concerned because it is what makes AI-driven tools of disease detection useful to farmers. Uneven play in the field conditions restricts the usage of such tools despite fair play in the research study settings. This has prompted the utilisation of Unsupervised Domain Adaptation (UDA) methods, which focus on domain adaptation, or the effort to move a model (learned on a labelled source dataset) to work on an unlabelled target dataset (Wilson and Cook, 2020).

2.2 Plant Disease Detection and the Role of Deep Learning

In plant disease detection, deep learning has changed the approach of detection and automated diagnosis of disease based on visual diagnostic foundations through convolutional neural networks (CNNs) (Mohanty et al., 2016; Too et al., 2019). Examples of those models include MobileNetV2 and EfficientNet because they have been able to learn complex visual features of large datasets and are accurate to near-human levels under controlled environments.

Nevertheless, such models are very sensitive to issues to do with training data quality and variety. They are also frequently overly optimistic to lab-style conditions when trained on datasets such as PlantVillage where there is consistent lighting, and the positioning of the leaves is clear. This causes their accuracy to decline drastically on realistic examples including the PlantDoc, where there is variability due to the environmental factors. The strategies involved in dealing with this limitation should include domain generalization such as UDA to improve the robustness of the models.

2.3 Challenges in Cross-Domain Plant Disease Classification

Plant disease cross-domain classification is also challenged by a significant obstacle: the domain shift that exists between the data domains that were gathered on controlled lab measurements versus those measured in real-world settings (Fang et al., 2023). The changes in lightning, noise, the direction of the leaf, and the picture resolution make some difference in the representation of features. This has led to deep learning models that cannot classify or give errors on noisy images taken in a field.

The discrepancies in the classes also complicate the predicament because rare diseases can be under-represented within the specified territory. In addition, noise caused by the environment, including overlapping of leaves, shadows, and soil textures, decreases the separability of the disease features in the latent space. These issues require adaptation methods that would enable such alignment of feature distributions between sources and targets without making use of significant labelled data belonging to the target domain.

2.4 Unsupervised Domain Adaptation (UDA) in Computer Vision

Unsupervised Domain Adaptation (UDA) seeks to enhance the performance of the model on the target domain without employing labelled samples associated with the target domain (Kouw & Loog, 2019). Approaches to

UDA in computer vision aim to make the model produce learned representations that behave comparably across the source and the target domains, by aligning the feature distributions of the source and target domains. Adversarial learning, feature alignment, and style transfer are the most popular of these, commonly introduced into an existing deep learning pipeline.

One of the most important benefits of UDA is that it has the potential to use a massive amount of unlabelled data that is frequently more easily accessible in practice. This is especially useful in agriculture, where obtaining and annotating the images of various plant diseases is expensive and time-consuming in the field. UDA provides a feasible route to operationalizing models trained using curated data in complex real-world areas because it can fill the source-target gap.

2.5 CycleGAN for Domain Translation

Cycle-Consistent Generative Adversarial Network (CycleGAN) provides image-to-image translation. It does this in such a way that image-to-image translation is done without paired training data examples (Zhu et al., 2017). This model employs two generators and two discriminators to train both forward and backward mappings with consistency in cycles where an image is translated to a different domain and back, and the translated image should look close to the original.

CycleGAN can also be applied in agriculture to convert high-quality pictures of plants in the lab into ones that have a similarity to those in the field, which can be used to train the model of features that are pertinent to real-world scenarios. This has the effect of minimising the visual discrepancy between the source and target datasets, which is a standard issue of performance degradation in the setting of domain shift. When included this translated image data in the training pipeline, models will see a greater number of variations in lighting, backgrounds, and leaf texture and will be more robust in actual deployment situations.

2.6 Self-Supervised Learning with SimCLR

SimCLR (Simple Framework for Contrastive Learning of Visual Representations) is one of the self-supervised learning methods which use the concept of contrastive loss to learn useful feature embedding with no labelled data (Chen et al., 2020). It is applicable by creating several augmented images of the same picture and making the model learn to represent them nearby within the embedding space, while pushing apart representations of different images.

SimCLR is able to learn domain invariant features in the real world setting on unlabelled plant images using a backbone encoder like ResNet-50. This is very useful especially in agricultural data sets where there is limited labelled field images. SimCLR also supports in disentangling representation in photos and in those with a CycleGAN translation in our pipeline to allow the classifier to generalize to diverse environments.

2.7 Pseudo-Labelling for Semi-Supervised Learning

Pseudo-labelling is a semi-supervised method in which the model's high-confidence output on unlabelled examples is taken as a proxy label, and the further training is done (Lee, 2013). The technique is helpful in domains such as agriculture, where annotations are expensive and labour-intensive, because samples specific to the target domain may be used without having to annotate them manually.

Pseudo-labelling was introduced in our project when PlantDoc images were pseudo-labelled using the classifier trained on PlantVillage data as well as the data that was translated using CycleGAN. Any prediction below a given probability value was discarded in order to eliminate noise by retaining only those with a high probability. The pseudo-labelled images were subsequently added to the training set and enhanced the exposure of real-world variances, as well as strengthening inter-domain feature alignment.

2.8 Related Works and Research Gaps

Recent studies have addressed domain adaptation in the context of plant disease detection, where significant improvements over the basic model are found. As indicated by Fang et al. (2023), CycleGAN boosted the classification process when used with adversarial domain adaptation in the real world. CycleGAN was also used by Jiang et al. (2021) to translate PlantVillage images into field-like images and obtain a better match with target-domain features. Other contrastive learning methods like SimCLR have been successful even in the low-label setting and presented robust representations (Chen et al., 2020).

However, most works have been dedicated to only one adaptation strategy, and little work has been done regarding the combination of visual translation, self-supervised learning, and semi-supervised labels in a common pipeline. There is also a gap in evaluating these techniques on highly imbalanced datasets like PlantDoc, whereby the minority classes are underrepresented. Our study addresses these gaps by merging CycleGAN, SimCLR, and pseudo-labelling to tackle multi-level domain discrepancies in plant disease classification.

3. Research Methodology

3.1 Overview of Research Approach

This study involves an experimental study that seeks to explore the use of unsupervised domain adaptation (UDA) methods in solving the issue of domain shift in plant disease classification. The main aim of consideration is to understand whether the combination of the image-level adaptation and feature-level adaptation can result in enhanced model generalization of the controlled laboratory images to the reality pictures in actual agricultural environments.

It was started with the creation of a baseline MobileNetV2 classifier trained and evaluated using the PlantVillage dataset, also known as the source domain, and the PlantDoc dataset as the target domain. The large decline in performance during this test was an indicator of the extent of the domain gap. To minimise this discrepancy, the research used 3 UDA components: CycleGAN to perform bi-directional translation between lab-style and field-style images, SimCLR to train domain-agnostic features through self-supervised learning, and pseudo-labelling to consider high-confidence predictions of target images that are not labelled. The components were incorporated in a sequential pipeline, and each of the stages was to increasingly enhance adaptation.

The assessment plan involved the combination of scoring, based on quantification, like accuracy, precision, recall, macro F1-score, and a qualitative performance assessment, which includes t-SNE visualizations. Such a compounded strategy provided that the gains would be evaluated not merely based on the numbers of their performance, but also in light of the correspondence of feature spaces between the two worlds.

It was the deliberate choice to get CycleGAN, SimCLR, and pseudo-labelling in the given order to reflect the character of the gap between the domains. CycleGAN was initially used to minimize pixel-level variations between source and target domain, which offers visually adapted images. Respectively, SimCLR came as a feature-level approach that made sure representations were not affected by environmental fluctuations due to transcendence. Lastly, pseudo-labelling was proposed to recapitulate on use of authentic target-domain samples in training so that the classifier is provided with a direct experience on field-data. Such sequencing made each stage subsequently rely on the previous one: the initial stage of visual adaptation was followed by the second stage of representation learning and the final piece is the integration into the target domain.

3.2 Dataset Description

In this study, the domain shift challenge was reproduced using two publicly available data sets representing two different disciplines. The source domain was the PlantVillage dataset, which includes more than 54,000 labelled images of healthy and diseased leaves of a plant taken under controlled settings. These images have homogeneous backgrounds, consistent illumination, and low noise levels, which means that they can be used well in model training at the beginning but are less reliable of the conditions in the field. Figure 3.1 illustrates typical PlantVillage samples, showing uniform backgrounds and lighting that simplify classification but do not reflect real farming conditions.

Figure 3.1 : PlantVillage images

The target domain was the PlantDoc dataset, which contained around 2,600 images that were gathered in real-world farming conditions. Such images are extremely diverse and include uneven lighting, a complicated background, occlusion, and different image resolutions. It has a total of 23 crop-disease classes, although this data set is highly imbalanced with respect to classes, which compounds the classification problem. Figure 3.2 highlights PlantDoc samples with varied angles, shadows, and cluttered backgrounds, underlining the domain gap.

Figure 3.2: PlantDoc images

The two datasets had to be aligned so that compatibility could be assured, and hence, class alignment was done so that only the common classes in the two domains are retained. All of the images are resized to standard input resolution, normalized and, for CycleGAN training, were scaled pixel values to [-1, 1]. This preprocessing allowed the datasets to be consistent and ready to undergo domain adaptation experiments.

3.3 Baseline Model Development

A baseline classifier created on MobileNetV2 architecture was required to define a benchmark of performance. This network was chosen because of its efficiency in computations and classification, which is balanced and makes it ideal when solving the tasks of image recognition on resource-constrained devices. The model was initialized with ImageNet-pretrained weights to leverage transfer learning benefits.

Training was only done on PlantVillage dataset, where images were resized to the necessary set of inputs, and normalized. The methods of data augmentation (random rotations and horizontal flips) were used to enhance the generalisation in the source domain. The training model was based on categorical cross-entropy loss and Adam optimizer on multiple epochs.

Comparison of the baseline model on the PlantDoc dataset without any types of domain adaptation resulted in a significant drop in accuracy as compared to the validation results on PlantVillage. It proved that there is a noticeable domain gap between the images generated in controlled conditions in the lab and photos taken in the field and, as such, necessitates the use of domain adaptation methods that are further investigated in the following stages of the study.

3.4 CycleGAN for Domain Adaptation

CycleGAN was used as an unpaired image-to-image method to deal with the visual disparities between laboratory-grade PlantVillage pictures and PlantDoc field pics. This method allows re-styling the images in the source domain to a style similar to the target domain without pairing training samples.

The CycleGAN model has two generators and two discriminators trained in one shot, where the source and target domains learn how to map one to another. A cycle-consistency loss kept an image translation to the target and back to the source domain is same as the original image. Averaged sets of images of the two datasets were standardized, and trained during several epochs, until the translations are visually compelling. Figure 3.3 presents CycleGAN-translated images that preserve disease symptoms while adopting the visual complexity of field conditions.

The plantDoc-style images that are generated were included in the training set as well as PlantVillage images. This step sought to decrease a pixel-level domain mismatch by exposing the classifier to instances of both a laboratory and field-style as a means of preparing it to better generalize to real-world agricultural imagery.

Figure 3.3: CycleGAN translated PlantVillage images

3.5 Feature Learning with SimCLR

In the translation of the images using CycleGAN technique, although it minimizes pixel differences, there may be residual domain gaps at the feature representation level. To resolve this, SimCLR self-supervised learning framework was employed, to learn domain-relevant features on CycleGAN translated PlantVillage images resembling the visual style of the PlantDoc data set. This had the effect of learning features without the need to have the actual real-world PlantDoc data with labels, and learning features that had characteristics like that of the target domain.

SimCLR has a convolutional encoder, a projection head, and a comparable loss function that train to maximize similarity among disparate augmentation-views of the same picture and distinguish other pictures. To simulate variability of field conditions, CycleGAN-translated images were heavily augmented through random crops, colour jittering, and Gaussian blur to create a library of heavily augmented pre-trained images to replicate the patterns of variability present in field conditions. The encoder able to generate resilient feature embeddings after training that were less susceptible to the differences in the environment.

The trained encoder was subsequently combined into the classification pipeline enabling the final model to capitalize on these trained features in addition to original and CycleGAN-translated images. This was done in order to enhance the generalization capabilities of the model to the real plant disease detection.

3.6 Pseudo-Labelling and Iterative Refinement

Although using an image translation algorithm and feature learning provides a good solution, it is also difficult to attain an ideal alignment between the source and target domains. Pseudo-labelling was proposed as a semi-supervised method to extend this closure even more to use real-world PlantDoc images in training without manual annotation. A classifier trained on labelled PlantVillage and CycleGAN-translated information, and

whose predictions could be used to label the unlabelled PlantDoc images, was utilized in this method. The predictions with confidence above a certain threshold were viewed as pseudo-labels and augmented the training set with target-domain samples.
The pseudo-labelled ones were then added to existing training data, and the model was again perfected in further iterations. Such a process presented the classifier with more varied and realistic training data and stimulated its flexibility towards target-domain features. Although pseudo-labelling has the potential to generate noisy labels, it represented an effective scheme to exploit large volumes of unlabelled data in the hope of getting better generalization into real-world scenarios.

3.7 Evaluation Metrics and Analysis

The work of the models was assessed both based on quantitative measures and qualitative visualizations. Accuracy, precision, recall, and F1-score were measured on a held-out test set of the PlantDoc dataset, which provides a direct measure of how well each approach performed to the domain shift. Other metrics used were macro-averaged to sustain the equal presentation of all classes, especially due to limited samples.
In addition to quantitative findings, t-SNE visualizations were created to find a comparison between feature embeddings of the default and adapted models. These plots were useful in gauging the level of similarity of the features of the source and target domains. It was evaluated in several setups such as: the baseline was only trained on PlantVillage data, the models with CycleGAN-translated images, the models with feature extraction using SimCLR, and the entire pipeline with pseudo-labelling. This enabled the comparison of the impact of each step of adaptation on performance to be compared easily.

3.8 Tools and Frameworks Used

The Python implementation was performed with TensorFlow and Keras as the main deep learning platforms to design, train, and test the model. MobileNetV2 was used as the baseline classifier because of its efficiency and the suitability for the transfer learning. The models, such as CycleGAN and SimCLR, were implemented in the TensorFlow ecosystem, and they took advantage of its data processing and augmentation utilities to process bulk data input images.
In visual inspection, t-SNE plots were created based on the learned feature embeddings by using scikit-learn to show how the clusters of feature embeddings were distributed. GoogleColab provided interactive environment where iteration was possible, and changes could be made in discreet steps and debugged. Such tools in combination allowed a painless integration of several domain adaptation methods into a cohesive pipeline, including data preparation, model calibration all the way to model assessment.

4. Design Specification

4.1 System Overview

In this project, it is proposed to optimize the plant disease classification in the real-world field conditions using the Unsupervised Domain Adaptation (UDA) approaches by means of a modular pipeline.
The system integrates the pixel, feature, and label-adaptation strategies to close the domain gap between the source images (high-quality plant images with labels (PlantVillage) and the field images (PlantedDoc).
Figure 4.1 demonstrates a top-level overview of the suggested architecture; the figure shows the flow of dataset inputs, adaptation phases to the eventual evaluation.

4.2 High-Level Architecture

The architecture has been regulated in neat steps. The pipeline is started by data preprocessing, during which datasets are synchronized by the class and standardized. The CycleGAN module carries out image-to-image translation where the source-domain images are translated to have a visual similarity to the target domain. SimCLR then does feature-level adaptation, i.e., learning robust, domain-invariant feature representations using unlabelled translated images. The pseudo-labelling phase labels target-domain samples with confidence and uses them as part of training. The last classifier is learned with a mixture of real, CycleGAN translated (and pseudo-labelled) data, and then tested on a held-out target domain (dataset).

4.3 Detailed Workflow

The system follows these sequential steps:

1. **Data Alignment and Preprocessing** – Map common classes between PlantVillage and PlantDoc, resize images, and normalize pixel values to ensure consistency across datasets.

2. **CycleGAN Translation** – Train CycleGAN to translate PlantVillage images into a PlantDoc-like style, generating adapted training data that visually matches the target domain.

3. **SimCLR Pretraining** – Apply SimCLR to the CycleGAN-translated PlantVillage images to learn feature representations that are more robust to domain-specific variations, using contrastive learning without requiring labeled target-domain data.

4. **Pseudo-Labelling** – Use a trained classifier to predict labels for target-domain (PlantDoc) samples, retaining only high-confidence predictions to reduce noise.

5. **Final Classifier Training** – Train the classifier using a combination of original PlantVillage images, CycleGAN-translated images, and pseudo-labelled PlantDoc samples for improved domain generalisation.

6. **Evaluation** – Assess performance using accuracy, precision, recall, F1-score, and t-SNE feature visualisations to compare feature space alignment between source and target domains.

4.4 Technology Stack and Tools

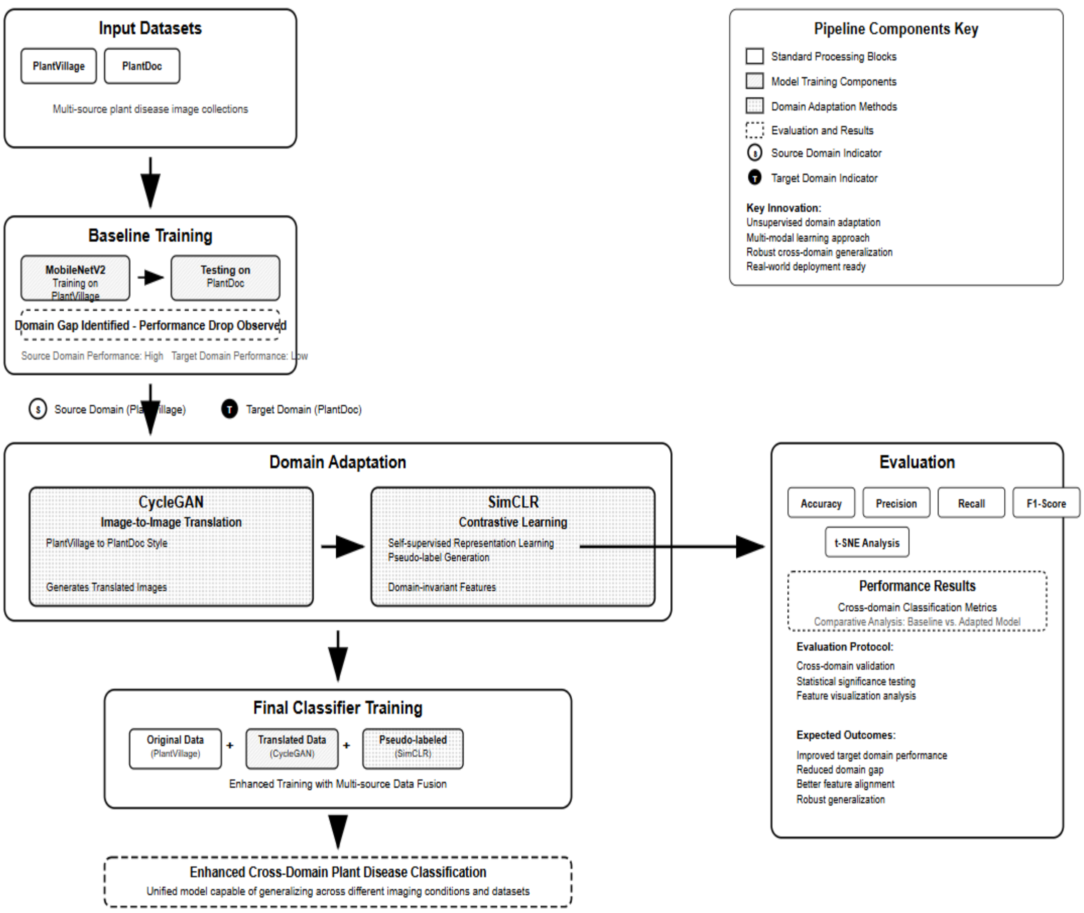
Pipeline was applied to Python with the utilisation of TensorFlow/Keras (deep learning models, MobileNetV2, CycleGAN, SimCLR) and scikit-learn (data preprocessing and t-SNE visualisation). The experiments took place in GoogleColab.

4.5 Design Decisions and Justifications

- **CycleGAN** was chosen for image-level adaptation as it allows unpaired translation between domains without requiring aligned image pairs.
- **SimCLR** was selected for feature-level adaptation, enabling robust feature learning from unlabelled images.
- **Pseudo-labelling** was incorporated to expand the labelled target dataset without manual annotation.
- **MobileNetV2** was adopted as the classifier due to its lightweight design, making it suitable for deployment in low-resource agricultural environments.

Figure 4.1 provides an overview of the full pipeline, showing how image-level, feature-level, and label-level adaptation components connect sequentially.

Plant Disease Classification Pipeline with Domain Adaptation



5. Implementation / Solution Development

The ultimate approach combined the three adaptation methods, CycleGAN, SimCLR, and pseudo-labelling-based methods-into a comprehensive pipeline. Preparation of datasets was crucial in the preparation of uniformity on data format as well as subsets to be applied at each phase, and accuracy in terms of classes. Training of an existing CycleGAN model (balanced subsets of PlantVillage (source) and PlantDoc (target) images) was used as an initial step in the process. This created pictures that were visually modified PlantVillage images that maintained some significant features of the diseases that they represented and that were formatted to represent a target domain.

In the next stage, SimCLR was applied to the PlantVillage images which were translated using CycleGAN, rather than to the raw PlantDoc dataset. This approach allowed the model to develop feature representations that could handle changes in lighting, background noise, and other conditions found in real-world settings, while still retaining the disease characteristics present in the original source images. The encoder produced during this step was then used to initialise the classifier. with the goal of boosting its ability to adapt and perform well on the target domain.

The trained classifier was then employed in the pseudo-labelling step, which provides labelling of the unlabelled PlantDoc images. Predictions with confidence levels above a chosen level were only retained so that the new samples added are trustworthy. The original PlantVillage images, together with the translated dataset created based on CycleGAN, and the pseudo-labelled images, were combined to produce a larger training set.

In the last stage, a MobileNetV2-based classifier was trained on this merged dataset. The performance over a held-out PlantDoc test set was measured by standard measures (accuracy, precision, recall, F1-score), and was visualised through t-SNE to demonstrate the feature correspondence between the source and target domains.

The systematic combination enabled evaluation of all the adaptation techniques/methods separately, giving evidence on how each of them contributes and limits them in enhancing the performance of cross-domain plant disease classification.

6. Results and Evaluation

6.1 Overview of Evaluation Approach

The assessment was done in a gradual manner that is to quantify the effects of varying Unsupervised Domain Adaptation (UDA) approaches on the classification of the disease in plants. The domain shift was measured by using the baseline model that was trained on only PlantVillage and applied to PlantDoc. CycleGAN-based image translation, SimCLR feature learning, and pseudo-labelling were added in successive phases to evaluate their independent effects and see how they can be combined to impact model performance.

The classification accuracy, macro-averaged precision, recall, and F1 score were used as the performance measures. Macro metrics were selected as the weight of each of the classes should be equal, whether it is imbalanced or not. Only its PlantDoc dataset was utilized as a test dataset; this way, target-domain images were not utilised to supervise training.

6.2 Baseline Model Performance

The baseline classifier with MobileNetV2 recorded impressive accuracy in the training and validation stages in PlantVillage but badly failed in testing with PlantDoc, showing how extreme the domain shift is.

As indicated in Table 6.1, the model attained an accuracy of 97.57% and 96.31% on the training and validation set, respectively, at an Epoch 8, but the test accuracy on PlantDoc decreased to 27.81%. This eminent performance drop proves that models based on clean and controlled images of PlantVillage have little to no generalizations to PlantDoc field images, which is also in line with similar past works on domain shift in plant disease datasets (Fang et al., 2023; Mohanty et al., 2016).

Table 6.1 – Baseline Model Performance

Metric	Training (Epoch 8)	Validation (Epoch 8)	PlantDoc Test
Accuracy	97.57%	96.31%	27.81%
Loss	0.064	0.1175	5.9193

6.3 Phase 3 - CycleGAN Domain Adaptation

This phase introduced CycleGAN-generated PlantDoc-style images into training. The aim is to bridge the visual gap between lab and field images without using labelled target data.

Training accuracy was obtained as 89.47 percent, as seen in Table 6.2, with validation accuracy 27 percent, but test accuracy decreased a little to 26.05 percent. Some classes performed better, e.g., Corn - Northern Leaf Blight and Squash - Powdery mildew, although a large number of classes either showed a poor prediction or completely went unpredicted. This is consistent with the shortcomings noted in Jiang et al. (2021), in which class-specific image-to-image translation boosted performance but failed to raise the general accuracy.

Table 6.2 – Phase 3 Model Performance

Metric	Training (Epoch 10)	Validation (Epoch 10)	PlantDoc Test
Accuracy	89.47%	27.66%	26.05%
Loss	0.3135	5.7957	–

6.4 Phase 4b – SimCLR Feature Learning with CycleGAN-Augmented Data

During this step, SimCLR pre-trained the feature extractor on CycleGAN translated images in an effort to learn domain-invariant representations before the training of a classifier.

The PlantDoc test accuracy was reduced as recorded in Table 6.3 to 21.00 percent with a macro F1 score of 0.10. On the one hand, not all classes like Corn Northern Leaf Blight lost considerable recall (0.97), but overall, the conclusion could be made that training SimCLR only on translated images had low diversity in features, which had lower generalization than the phase with the CycleGAN only.

Table 6.3 – Phase 4b Model Performance

Metric	PlantDoc Test
Accuracy	21.00%
Macro Precision	0.21
Macro Recall	0.14
Macro F1 Score	0.1

6.5 Final UDA – Combined CycleGAN, SimCLR, and Pseudo-Labelling

The last step combined all of the UDA elements, including original PlantVillage data, images translated by the CycleGAN, and top-confidence pseudo-labels of PlantDoc dataset.

According to Table 6.4, the accuracy of the test was 23.08%, and the macro F1 was equal to 0.16. Although the final model was efficient in the sense that it enhanced some of the minority classes, the accuracy was worse than the baseline. This is contrary to the earlier hypothesis that the performances would continue to be improved by including domain adaptation techniques. Nevertheless, this result concurs with that reported by Arazo et al. (2020), who mentioned that pseudo-labelling may induce noise when domain mismatch is significant, which may decrease performance in complicated target domains.

Table 6.4 – Final UDA Model Performance

Metric	PlantDoc Test
Accuracy	23.08%
Macro Precision	0.3
Macro Recall	0.19
Macro F1 Score	0.16

Table 6.5 – Results Summary Table

Phase / Model	Training Accuracy	Validation Accuracy	PlantDoc Test Accuracy	Macro Precision	Macro Recall	Macro F1 Score
Baseline (MobileNetV2)	97.57%	96.31%	27.81%	–	–	–
Phase 3 – CycleGAN	89.47%	27.66%	26.05%	0.28	0.22	0.19
Phase 4b – SimCLR + CycleGAN	–	–	21.00%	0.21	0.14	0.10
Final UDA (All components)	–	–	23.08%	0.30	0.19	0.16

Table 6.5 presents the findings in all the phases in order to compare them. The baseline did well on validation, and not so well on domain shift. Nevertheless, CycleGAN had some small gains on individual classes, but the overall level of accuracy was poor. It is even worse that within the case of SimCLR, which utilised data that was translated, the performance failed to generalise, resulting in further decline. Combining CycleGAN, SimCLR and pseudo-labelling into the final UDA pipeline also increased minority classifying but failed to beat the base line. This survey shows the imbalanced advantages of each adaptation stage, and the challenge of making the huge domain gaps.

6.6 Analysis and Implications

In academic settings, such findings undermine a popular assumption (Wang and Deng, 2018; Kouw and Loog, 2019) that the more UDA components are added, the better. The study shows that when the source-target gap is large and target data is highly variable, domain adaptation gains may be inconsistent or even negative.
Going over the findings as a practitioner, it is essential to note that implementing such models in the real agricultural setting necessitates not only adaptation strategies but also the need to emphasize the curation of the dataset, refined noise handling of pseudo-labels, and maybe even multi-modal input (Ferentinos, 2018).

7. Conclusion and Future Work

7.1 Conclusion

In this study, we set out to see whether Unsupervised Domain Adaptation (UDA) could help a plant disease classifier trained on a lab dataset work better on real-world images. The lab dataset was PlantVillage, and the real-world set was PlantDoc. To try and bridge the gap between them, three methods were tested in sequence: using CycleGAN to create PlantDoc-style images from PlantVillage photos, applying SimCLR to learn stronger feature representations, and adding pseudo-labels from the target domain into training.

The original MobileNetV2 baseline was only trained on PlantVillage, and was able to reach a very high in-domain accuracy (96.31% validation accuracy), before falling to a very poor accuracy (27.81%) on Plantdoc, confirming higher domain shift. The CycleGAN model added PlantDoc-style images to the training, and the result was just a slight decrease in accuracy (26.05 percent). SimCLR phase, where domain-invariant features were learnt, was followed by further reduction (by 21.00%), probably caused by the lack of diversity in translated image training dataset.

The last combined UDA strategy had an accuracy of 23.08 percent, with improvements in a few minority classes, but it is lower than the performance of baseline model.
Such results go against the first hypothesis that incremental accuracy would be achieved by sequentially adding UDA techniques. Rather, the findings point out that when source–target domain gap is large and target-domain data is highly variable, domain adaptation gains may be inconsistent. It is consistent with what recent studies have been observing (Arazo et al., 2020; Jiang et al., 2021) in that adaptation strategies could add noise or bias, particularly where pseudo-labels are not perfect and when synthetic pictures do not show complete variability of the target domain.

Academically, the work makes it clear that there is a need to critically assess methods of adaptation in a real deployment environment instead of expecting theoretical advances to have a direct transfer to practice. For practical use in agriculture, it shows that AI models need access to rich, varied, and carefully collected field datasets if they are to perform reliably outside the lab.

7.2 Future Work

Building on these findings, future research could explore several promising directions:

1. **Enhanced Feature Extractors**
   Use a deeper or higher-resolution backbone, e.g., EfficientNet or ConvNeXt, as both the baseline and adapted model, which could improve representation ability on par with more overfit ones.

2. **Hybrid Adaptation Strategies**
   Blend the processes of visual-level adaptation (e.g., CycleGAN) with adversarial domain adaptation techniques like Domain-Adversarial Neural Networks (DANN) to indeed better match feature distributions.

3. **Improved Pseudo-Label Filtering**
   Apply confidence thresholding, ensemble-based voting, or label smoothing to reduce the impact of incorrect pseudo-labels.

4. **Multi-Modal Data Integration**
   Incorporate environmental and contextual data such as soil moisture, temperature, and geolocation alongside visual data to improve robustness in field conditions.

5. **Dataset Expansion and Quality Control**
   Curate larger, more balanced target-domain datasets that capture seasonal, geographic, and crop variety differences to improve adaptation outcomes.

6. **Deployment-Oriented Optimization**
   Convert final models to lightweight formats (e.g., TensorFlow Lite) and evaluate performance on mobile devices for real-time agricultural use cases.

7.3 Stakeholder Relevance

- **Researchers**: Offers empirical results that popular UDA pipelines do not necessarily lead to better performance in situations of large domain shifts, motivating a better specific assessment of adaptation techniques.

- **Agricultural Companies and Startups**: Outlines front-line risks of directly applying laboratory-trained plant disease models in the field, vindicating the need to invest in field data and model retraining.

- **Policy Makers and Agricultural Institutions**: Stresses the requirement of open and good-quality quality and diverse agricultural data to enable AI in the agricultural sector.

- **Farmers and Agricultural Extension Services**: Emphasizes that the available AI tools might not be enough in real-life situations and should be implemented with the assistance of expert confirmation.

References:

1. **Wang, M. and Deng, W.**, 2018. Deep visual domain adaptation: A survey. *Neurocomputing*, 312, pp.135–153.

2. **Kouw, W.M. and Loog, M.**, 2019. A review of domain adaptation without target labels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3), pp.766–785.

3. **Mohanty, S.P., Hughes, D.P. and Salathé, M.**, 2016. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, p.1419.

4. **Too, E.C., Yujian, L., Njuki, S. and Yingchun, L.**, 2019. A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, pp.272–279.

5. **Kayal, P., Sharma, S., Yadav, M., Arya, D., Dutta, S. and Chakravarty, D.**, 2019. PlantDoc: A dataset for visual plant disease detection. *arXiv preprint* arXiv:2001.01407.

6. **Fang, Z., Cui, F., Zhang, Y., Xie, J., Dong, X. and Lin, K.**, 2023. From laboratory to field: Unsupervised domain adaptation for plant disease recognition in the wild. *Plant Phenomics*, 2023, pp.1–15.

7. **Zhu, J.Y., Park, T., Isola, P. and Efros, A.A.**, 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. IEEE, pp.2223–2232.

8. **Jiang, B., He, Y., Deng, Y., Yang, G. and Lin, C.**, 2021. Unsupervised domain adaptation using CycleGAN for plant disease identification. *Computers and Electronics in Agriculture*, 181, p.105933.

9. **Chen, T., Kornblith, S., Norouzi, M. and Hinton, G.**, 2020. A simple framework for contrastive learning of visual representations. In: *Proceedings of the International Conference on Machine Learning (ICML)*. PMLR, pp.1597–1607.

10. **Lee, D.H.**, 2013. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In: *Workshop on Challenges in Representation Learning, ICML*.

11. **Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C.A., Cubuk, E.D., Kurakin, A. and Li, C.L.**, 2020. FixMatch: Simplifying semi-supervised learning with consistency and confidence. *Advances in Neural Information Processing Systems*, 33, pp.596–608.

12. **Tzeng, E., Hoffman, J., Saenko, K. and Darrell, T.**, 2017. Adversarial discriminative domain adaptation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.7167–7176.

13. **Wilson, G. and Cook, D.J.**, 2020. A survey of unsupervised deep domain adaptation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(5), pp.1–46.

14. **Ferentinos, K.P.**, 2018. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, pp.311–318.

15. **Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J. and Hughes, D.P.**, 2017. Deep learning for image-based cassava disease detection. *Frontiers in Plant Science*, 8, p.1852.

16. **Mwebaze, E., et al.**, 2019. iCassava 2019: Fine-grained visual categorization challenge. *arXiv preprint* arXiv:1908.02900.

17. **Benaim, S. and Wolf, L.**, 2017. One-sided unsupervised domain mapping. *Advances in Neural Information Processing Systems (NeurIPS)*, pp.752–762.

18. **Arazo, E., Ortego, D., Albert, P., O'Connor, N.E. and McGuinness, K.**, 2020. Pseudo-labeling and confirmation bias in deep semi-supervised learning. *Pattern Recognition*, 107, p.107501.

19. **Abadi, M., et al.**, 2016. TensorFlow: A system for large-scale machine learning. In: *OSDI*, pp.265–283.

20. **Chollet, F.**, 2015. Keras. *GitHub repository*. Available at: https://github.com/keras-team/keras

21. **Pedregosa, F., et al.**, 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, pp.2825–2830.

22. **Van der Maaten, L. and Hinton, G.**, 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, pp.2579–2605.

23. **Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.C.**, 2018. MobileNetV2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.4510–4520.