PROJECT PLANT DISEASE CLASSIFICATION

PROJECT TIER 1

RODRIGO SIERRA

1378 COMPUTER VISION

TEACHER: PATRICIA MCMANUS

OCTOBER 30 2025

Image source: https://gardenerspath.com/how-to/disease-and-pests/septoria-leaf-spot-tomatoes/

Image source: https://www.iowapublicradio.org/gardening/tomato-diseases-in-iowa/

Image source: https://www.starnewsonline.com/story/news/2021/03/10/burgaws-thomas-produce-farm-famous-tomatoes-close/6946296002/

WHAT IS THE PROBLEM?

What real-world problem are you solving?

Who cares about this problem?

Why is it important?

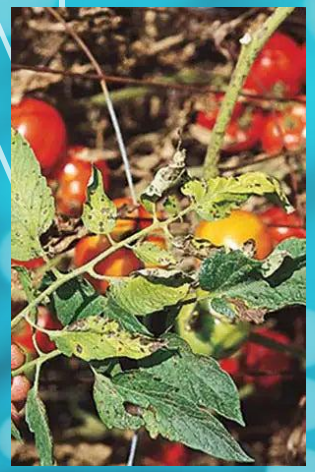


Image source: https://www.gurneys.com/pages/tomato-diseases-and-treatment?srsltid=AfmBOoqaZuStZXv5Myhn9-spD1pGL0tw3brCynRtcGi13ivfG3uJ7mp2//

Image source: https://www.greenlife.co.ke/tomato-farming-for-beginners-planting-growing-and-harvesting//

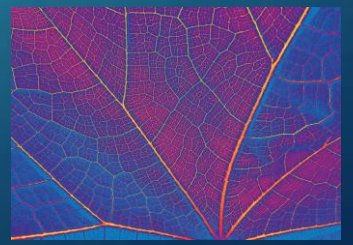
MY SOLUTION

What will my system do?

How will it solve the problem?

TECHNICAL APPROACH?

- CV Technique: Classification.
- Model: ResNet50.
- Framework: PyTorch
- Why this approach?
  - + High accuracy.
  - + Facilitate.



DATA PLAN

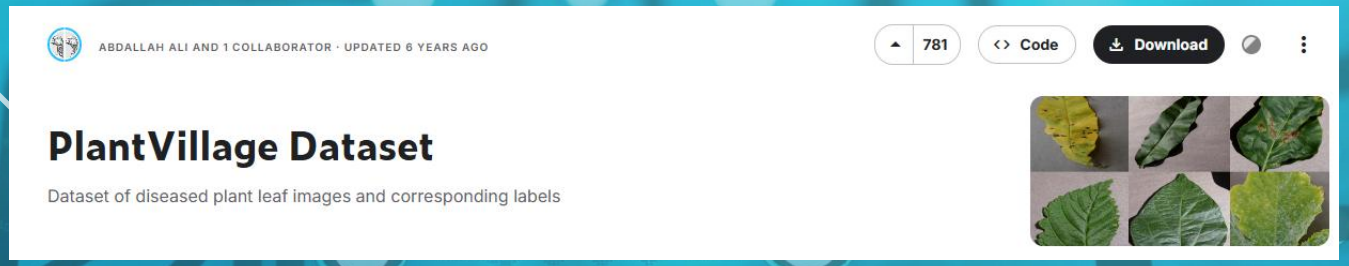


Image source: https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset

- Source: PlanVillage Dataset.
- Size: 54.000+ images.
- Labels: 38 different classes.
- Link: https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset

SYSTEM DIAGRAM

Image pre-processing

Training and validation dataset

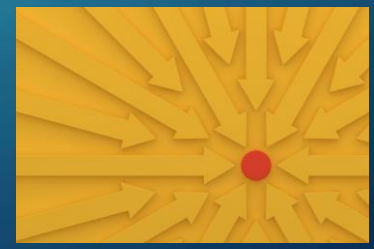
Training deep learning

Performance assessment

Plant leaf image classification

SUCCESS METRICS

- Primary Metric: Accuracy
- Target: Greater than 90% accuracy.
- Secondary Metrics: F1-score.

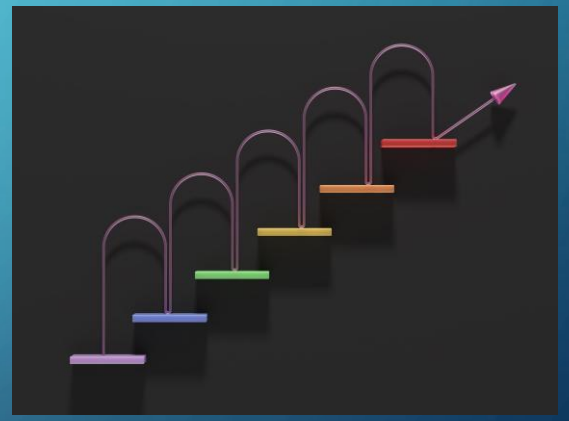


WEEK-BY-WEEK PLAN

Week	Tasks	Milestone
10 (Oct 30)	Get dataset, set up environment	Dataset ready
11 (Nov 06)	Train/fine-tune model	Model working
	Load Pre-trained model (ResNet)	
	Organize dataset	
	Fine-tune model	
12 (Nov 13)	Test and improve	Good accuracy
	Evaluate performance	
	Improve performance	
	(if required)	
13 (Nov 20)	Create demo, make video	Demo ready
	Make demo notebook	
	Record demo video	
	Update README.md	
14 (Nov 27)	Final testing, documentation	Everything done
	Clean up code	
	Make presentations slides	
	Final testing	
15 (Dec 04)	Presentation	Presentation done

CHALLENGES & BACKUP PLANS

Risk	Probability	Mitigation
Overfitting	Medium	Implement strong data augmentation.
Class Imbalance	Medium	Use of Weighted Cross-Entropy Loss.
GPU	Medium	Use Google Colab Pro
Interruption Training Time	Medium	Use Google Colab Pro



RESOURCES NEEDED

- Kaggle: Source of dataset.
- Compute: Google Colab or Google Colab Pro.
- Cost: 0 – 10 USD.
- API: No.
- Github: Repository.