

Problem definition

What is Phenotyping? (Netherlands Plant Eco-phenotyping Centre)

- "Root phenotyping is the study of observable root traits, such as structure, growth, and environmental interaction."
- "In NPEC's case, it involves tracking traits like root length, branching using image analysis"
- "At NPEC, seeds of *Arabidopsis thaliana* are sown in Petri dishes, imaged daily as roots grow, and exposed to microbes or treatments."
- "These images are used to quantify key traits and analyze plant-microbe interactions."

Insights Phenotyping Provides:

- "Phenotyping reveals how roots grow and adapt under different conditions, such as limited water or varying microbial environments."
- "It supports breeding for resilient crops, enabling sustainable and improved agricultural practices."

Challenges in Phenotyping:

1. **Complex Root Systems:**
 - "Overlapping and crisscrossing roots make manual segmentation time-consuming and error-prone."
2. **Precision in Inoculation:**
 - "Microbial treatments require pinpoint accuracy to ensure consistent experimental outcomes."
3. **Scalability:**
 - "Processing 10,000+ seedlings manually is impractical and creates a bottleneck in data analysis."
4. **Time and Resources:**
 - "Current manual methods are slow and require significant human labour, limiting research efficiency."

In conclusion, automation is needed...

Overview of the proposed solution:

Solution Components:

- Computer vision: root analysis → segmentation, length, tips.
- Robotics: inoculation → precision, consistency.

Benefits to NPEC:

- Streamlines root phenotyping → less manual effort.
- Scalable → processes thousands of seedlings daily.
- Reliable data → root traits, growth, plant-microbe studies.
- Accelerates research → genotypes, stress resistance, nutrient uptake.

Business Impact:

- Faster experiments, larger-scale studies.
- Supports sustainable agriculture → resilient crops.

Results & Evaluation:

Results (Computer Vision):

- U-Net → root detection accuracy with 82.18 F1 score on a validation set.
- 12.47 (% error score) sMAPE → high accuracy in root length detection.
- Tip detection → essential for inoculation pipeline input.
- Root segmentation → identifies roots (1–5) in complex Petri dish images.
- **Reinforcement Learning:**
 - working on RL PPO model for robotic inoculation.
 - RL pipeline will integrate with CV for complete automation.
- **Business Value:**
 - Accurate root analysis → scalable phenotyping for 10,000+ seedlings.
 - Consistent and reliable data → accelerates genotype identification.
 - Directly supports plant-microbe studies and sustainable agriculture goals.
- **Transition to Next Slide:**
 - "Next, I'll walk you through the development and iterations that led to these results."

Error Analysis and Iterations of the CV pipeline:

Iteration 1:

- **Initial Steps (Constants):**
 - Image preprocessing: crop, pad, patchify, and unpatchify.
 - U-Net model trained on 2023 dataset (11k images split into train/val).
- **Challenges/Error Analysis:**
 - Poor segmentation results → **F1 score of 76.88 on validation set.**
 - Missing large portions of roots → insufficient for downstream tasks (e.g., root segmentation and analysis).
 - Could not attempt individual root segmentation.
- **Impact on the Project:**
 - Identified key issues with training data or model setup.
 - Provided a baseline for improvement → next iteration needed better performance.
- **Transition to Next Slide:**
 - *"This iteration helped identify the limitations of the initial dataset and model. In the next iteration, I implemented changes to improve segmentation accuracy."*

Iteration 2:

Improvements from Iteration 1:

- Combined datasets → larger, more diverse training data.
- F1 score improved from 76.88 to **82.18**, enabling clear masks.
- Introduced **morphological operations** to reduce noise and close gaps.

Root Segmentation Workflow:

- Divided masks into **5 vertical sections** → bounding boxes identify roots.
- Skeletonised roots and calculated lengths using **Dijkstra's algorithm**.
- Assigned IDs to roots → accounted for missing or non-contiguous plants.

Challenges and Learnings:

- Vertical sectioning caused some roots to be missegmented (cut off).
- Small plants not identified → revealed the need for more adaptable segmentation methods.

Business Value:

- Enabled detailed root phenotyping by accurately identifying and measuring roots.
- Provided data critical for analysing root traits and advancing genotype research.

Transition to Next Slide:

- "This iteration brought significant improvements, but there were still challenges to address. In the next iteration, I focused on refining segmentation further to capture smaller and more complex root systems."

Iteration 3:

Improvements from Iteration 2:

- Retained inference model → consistent, high-quality masks.
- **Dynamic sectioning** → adaptive segmentation captured edge cases more effectively.
- Enhanced ability to retain roots in challenging scenarios (e.g., plants at the edges of the Petri dish).

Workflow Consistency:

- Skeletonized roots → calculated lengths using Dijkstra's algorithm.
- Assigned root IDs based on dynamic sections.

Challenges and Learnings:

- Very small plants still missed → significant relative error in sMAPE due to their small size.
- Improvement in sMAPE (12.47) showed progress but highlighted the need for better detection of tiny plants.

Assumptions:

5 plants max per dish.

Plants seeded in **5 sections**.

Seeds placed in **top quarter** of the dish.

Assumptions guided **segmentation process**.

Limitations:

Model not perfect but accurate.

Overlapping roots → misinterpretations.

Tiny roots mistaken for scratches/noise.

Further Steps:

Most cases identified **correctly**.

Inference Model:

- Improve segmentation accuracy → tiny roots and overlapping areas.
- Address the current model's edge-case failures.

Preprocessing:

- Better noise filtering to detect all roots, even very small ones.
- Adapt to challenging plant positions for better segmentation.

Reinforcement Learning Integration:

- PPO hyperparameter tuning → optimizing n_steps.
- Use CV root tip detection as the pipette goal.
- Automate inoculation fully through RL-CV integration.

Pipeline Refinement:

- Focus on accurate **root development (length) analysis**.
- Ensure robust and scalable solutions for phenotyping.

Future Applications:

- Apply pipeline to other phenotyping tasks.
- Enable large-scale experiments to advance genotype research.

The End...