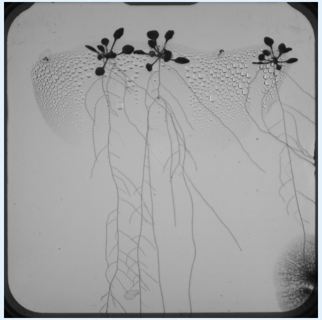
Automating Root Phenotyping for NPEC

Michał Bątkowski

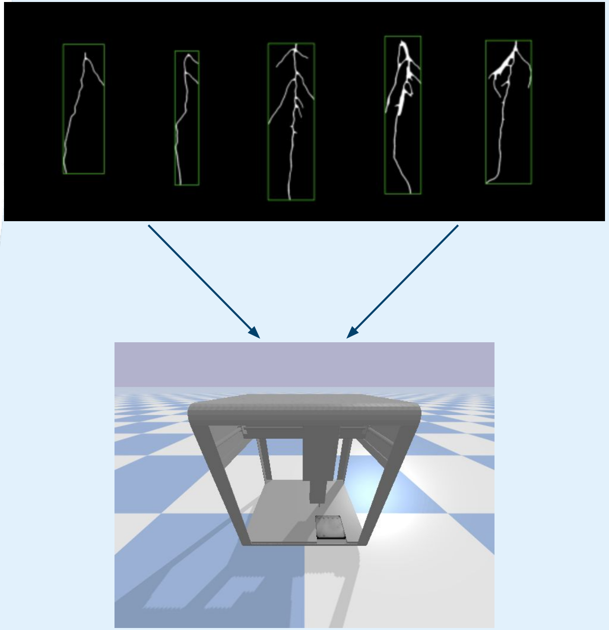
1. Problem definition

"Why Automate Root Phenotyping?"



- **Phenotyping Overview:**
  - Study of plant traits (growth, nutrient absorption, stress resistance).
  - Critical for sustainable agriculture and developing resilient crops.
  - Root phenotyping reveals traits like growth, nutrient uptake, and stress resistance.
- **Challenges Faced by NPEC:**
  - Overlapping roots complicate manual segmentation.
  - Robotic inoculation requires high precision.
  - Manual methods are slow, error-prone, and unscalable.

2. Overview of the proposed solution



Our Solution:

- **Computer Vision:**
  - Segments roots, measures their length, and detects root tips for detailed analysis.
- **Robotics:**
  - Trained via reinforcement learning to inoculate roots with precision.

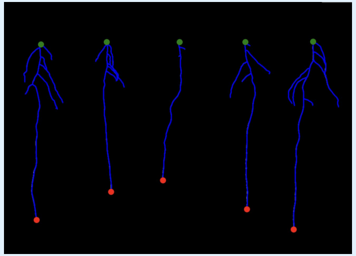
How It Helps NPEC:

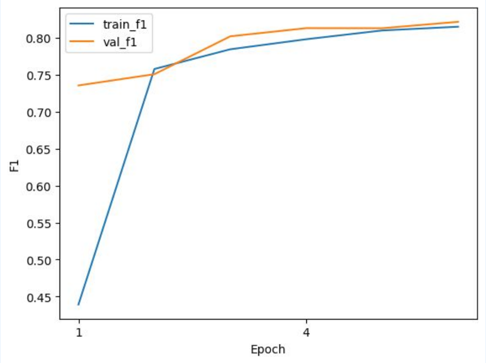
- Streamlines root phenotyping and reduces manual effort.
- Processes 10,000+ seedlings with consistent accuracy.
- Accelerates analysis of root traits for identifying resilient genotypes.
- Enables scalable and reliable plant-microbe interaction studies.

3. Results & Evaluation

Computer Vision Results:

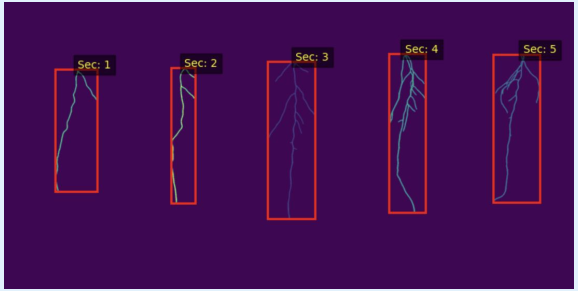
- **U-Net Model:** Achieved **82.18 F1 score** on the validation set for root detection.
- Achieved **12.47 sMAPE** in root length detection.
- Accurately detects **root tips**, providing input for inoculation.
- Identifies and segments **individual roots** (1–5) on Petri dish images.





Business Value:

- Enables precise, automated root analysis for phenotyping.
- Provides reliable data for plant trait studies and experiments.
- Forms the foundation for integrating robotics with RL to scale inoculation



4. Error Analysis and Iterations of the CV pipeline

Iteration 1

Iteration 1: Initial Model

- Dataset: 2023 only.
- Poor masks → large portions missing.
- F1 score: **76.88**.

Iteration 2

Iteration 2: Enhanced Segmentation

- Combined datasets (2023 + 2024).
- Enhanced Morphological operations to reduce noise.
- F1 score: **82.16** → sMAPE: **15.58**.

Dynamic sectioning to address edge cases.

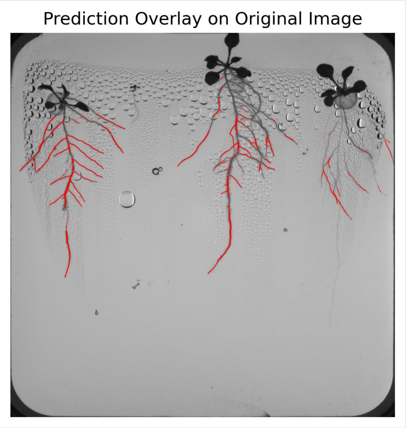
Improved root segmentation and analysis.

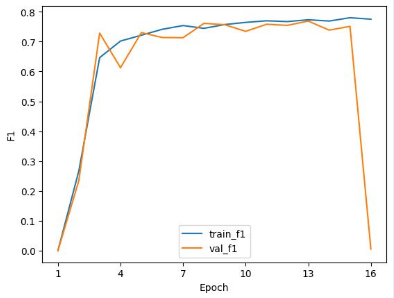
sMAPE: **12.47**.

Iteration 3

Iteration 1

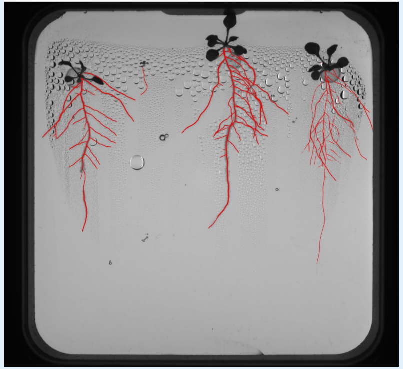
- **Objective:**
  - Train an initial model to segment roots from Petri dish images.
- **Steps Taken:**
  - Cropped and padded images for patching.
  - Ran inference with a U-Net model trained on the 2023 dataset (11k images).
  - Dataset split: 8769 training, 2243 validation images.
- **Challenges Faced:**
  - Achieved **76.88 F1 score** on the validation set → poor performance.
  - Large portions of roots missing in predicted masks.
  - Unable to segment individual roots due to low-quality predictions.
- **Insights Gained:**
  - Initial dataset/model was insufficient for accurate segmentation.
  - Highlighted the need for better preprocessing, training strategies, or dataset improvements.





Iteration 2

Prediction Overlay on Original Image



Objective:

- Improve root segmentation accuracy and enable root length analysis.

Steps Taken:

- Combined **2023 + 2024 datasets**, increasing data to **116k train / 22k val images**.
- Adjusted patch step size → created more training data.
- Trained U-Net model (batch size 32) → achieved **82.18 F1 score (validation)**.
- Preprocessed masks with morphological operations to remove noise and close gaps.
- Divided masks into **5 vertical sections** and applied bounding boxes for root segmentation.
- Calculated skeleton lengths using **Dijkstra's algorithm** and assigned unique root IDs.

Results:

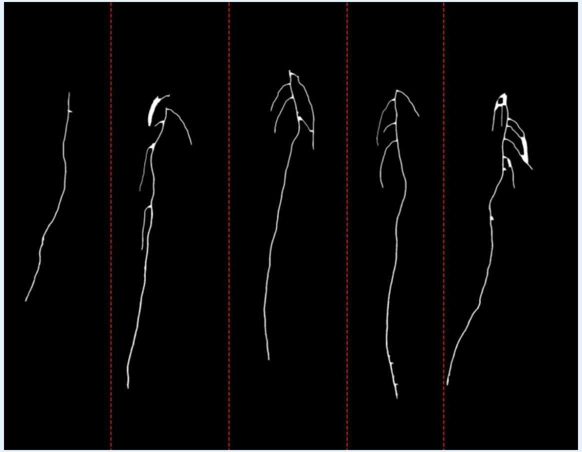
- Achieved **15.58 sMAPE** in root length detection.
- Successfully segmented and identified roots in most cases.

Challenges Identified:

- Vertical sectioning caused some roots to be cut off.
- Smaller plants were sometimes missed during segmentation.

Iteration 3

- **Objective:**
  - Refine root segmentation to handle edge cases and improve accuracy.
- **Steps Taken:**
  - Retained the same U-Net model with **82.18 F1 score** for inference.
  - Introduced **dynamic sectioning** → adaptive to root positions and sizes.
  - Retained components based on **y threshold** or **size threshold**.
  - Applied existing workflow (skeletonization, bounding boxes, Dijkstra's algorithm).
- **Results:**
  - Achieved **12.47 sMAPE** in root length detection.
  - Improved segmentation of edge cases (e.g., plants at unusual positions).
- **Challenges Identified:**
  - Very small plants still missed → large impact on percentage error (sMAPE).
  - Score improvement limited by inability to detect tiny root structures.



Assumptions

Number of Plants: Assumed a maximum of **5 plants per Petri dish.**

Seeding Placement:

- Assumed plants are seeded within **5 equal vertical sections.**
- Plants are placed in the **top quarter** of the dish.

Uniformity: Assumed seeds are distributed evenly across sections.

Limitations

Model Accuracy:

Inference model is **not perfect**, though generally accurate.

Tiny Roots vs. Noise:

Very small roots are sometimes misclassified as noise or scratches.

Overlapping Roots:

Misinterpretations occur when roots overlap or grow over the shoot.

Summary and Further steps

Improving the Inference Model:

Address edge cases (e.g., tiny roots, overlapping areas).

Enhanced Preprocessing:

Refine steps to identify all roots, including small plants.

Integration with RL:

Continue hyperparameter tuning for the PPO model.

Use CV root tip detection as input for inoculation goals.

Future Applications:

Apply pipeline to other phenotyping tasks.

Enable large-scale experiments to advance genotype research.

Thank you for your attention!

Breda University
OF APPLIED SCIENCES