3D plant phenotyping using deep learning

Gert Kootstra

Wageningen University and Research

14 Dec 2022

Digital Twin Conference 2022

WAGENINGEN
UNIVERSITY & RESEARCH

Outline of this tutorial

- 10:00-10:40 Intro to 3D plant phenotyping using deep learning
- 10:40-12:00 Hands-on session
  - Working with point clouds
  - Using deep learning to segment a point cloud

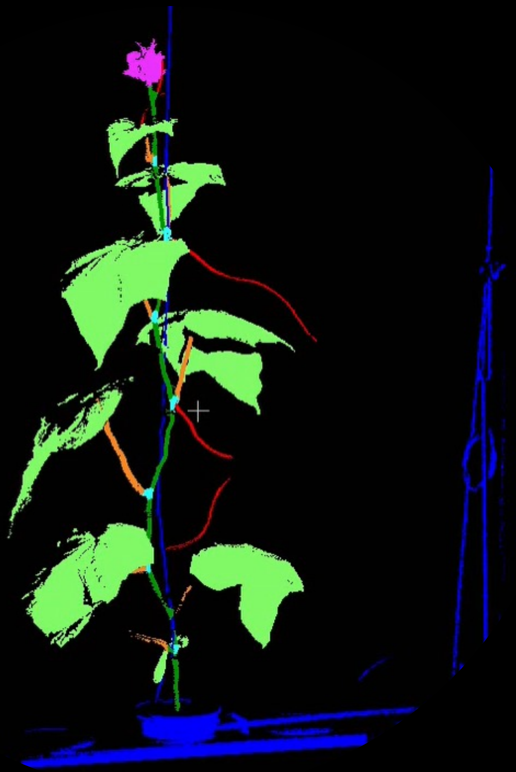
WAGENINGEN UNIVERSITY & RESEARCH

Outline of the introduction

- Introduction to 3D plant phenotyping
- 3D data acquisition
- Data processing
- Challenges
- Conclusion

Introduction to 3D plant phenotyping

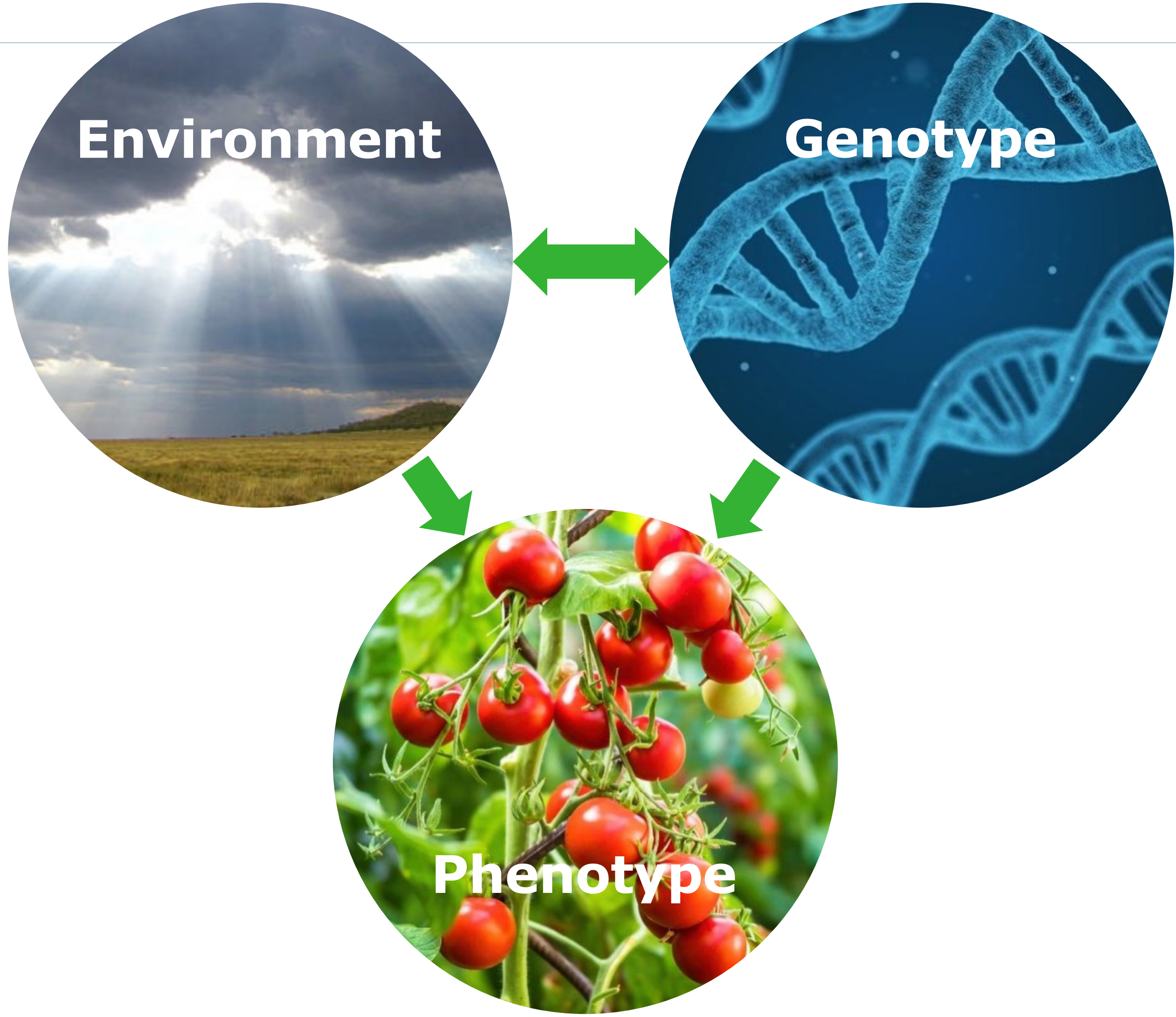
- Objectives
- Why 3D?



WAGENINGEN
UNIVERSITY & RESEARCH

Plant development

- Plant development is a result of the interaction of the genotype (DNA) and the environment
- G + E = P

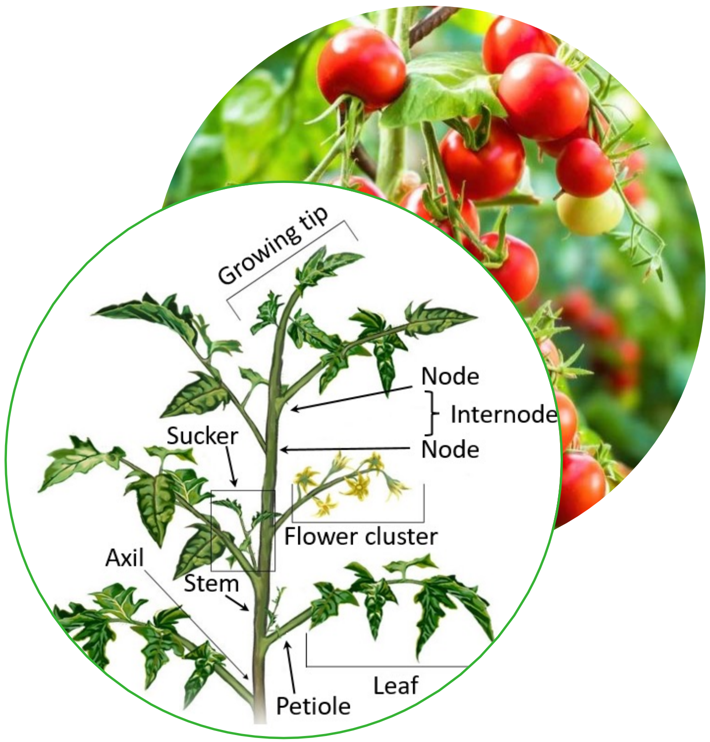


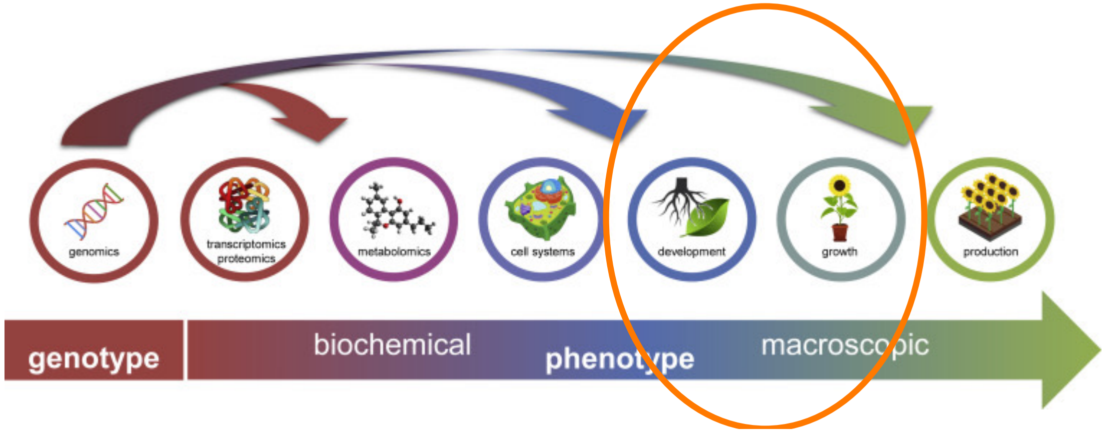
Level of automation

- In controlled environments (greenhouses), we have a lot of knowledge about the environment
- Genetic information can be gather efficiently with next-generation-genotyping tools
- But the phenotype at plant level is still mainly measured by hand

Digital plant phenotyping

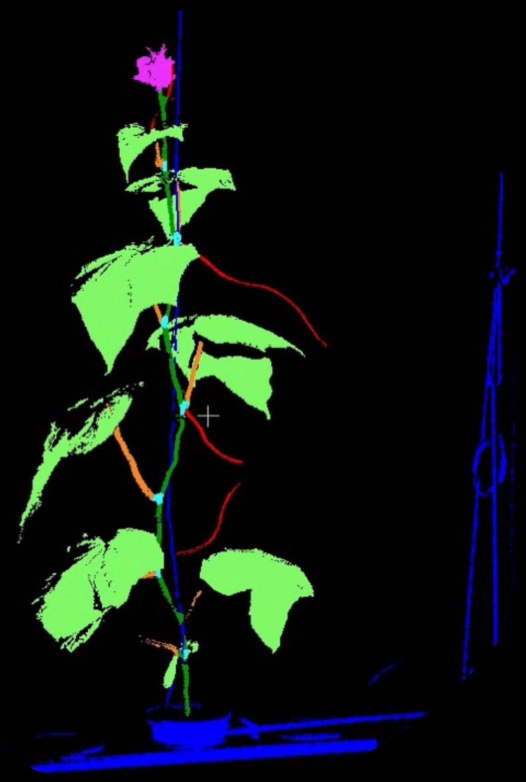
- Automatic assessment of plant traits
- Better quality of measurements
    - Objective measurements
- Better quantity of measurements
    - Moreplants





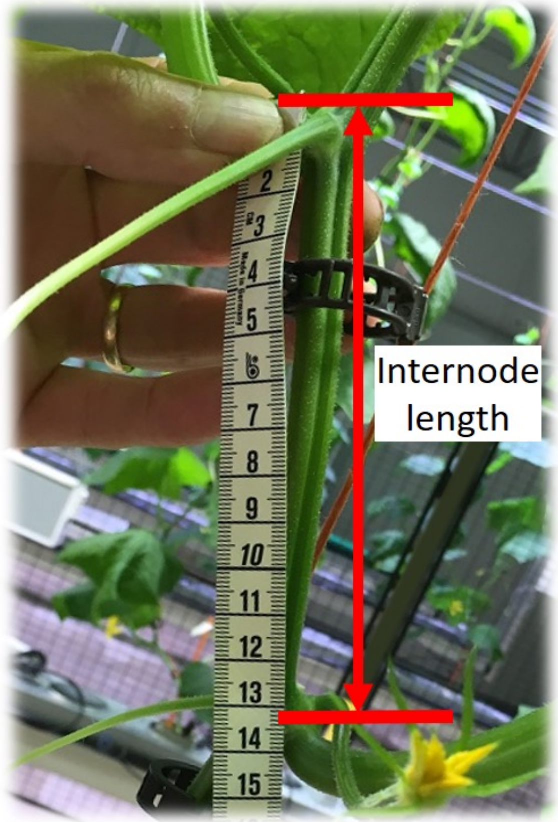
Introduction to 3D plant phenotyping

- Why 3D?



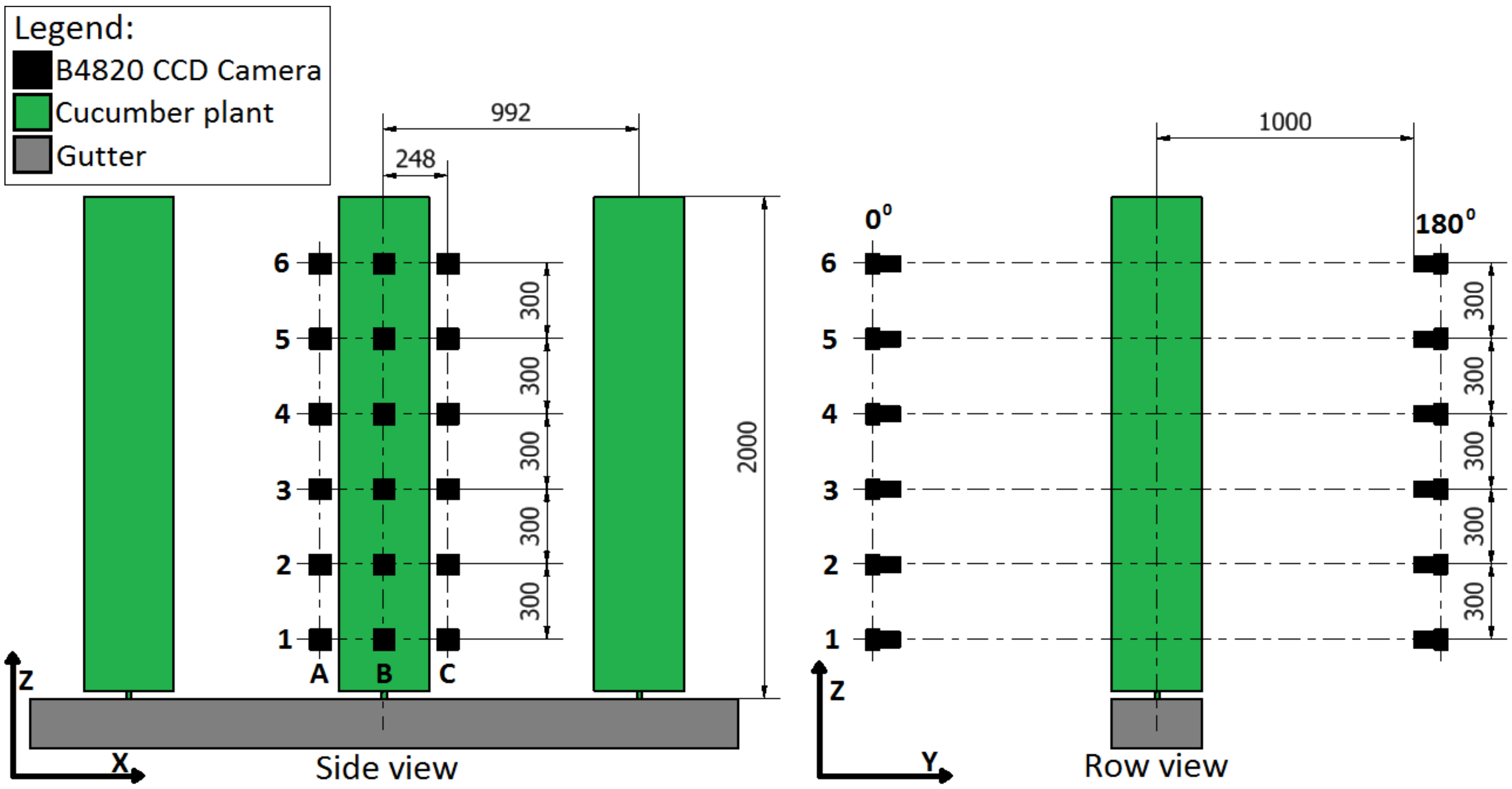
Why 3D? A motivating 2D example

- Internode length and development
  - Of interest for breeders
  - Light interception
  - Density of the crop
  - Breeding for robotics



Boogaard, F. P., Rongen, K. S., & Kootstra, G. (2020). Robust node detection and tracking in fruit-vegetable crops using deep learning and multi-view imaging. *Biosystems Engineering*,192, 117–132. https://doi.org/10.1016/j.biosystemseng.2020.01.023

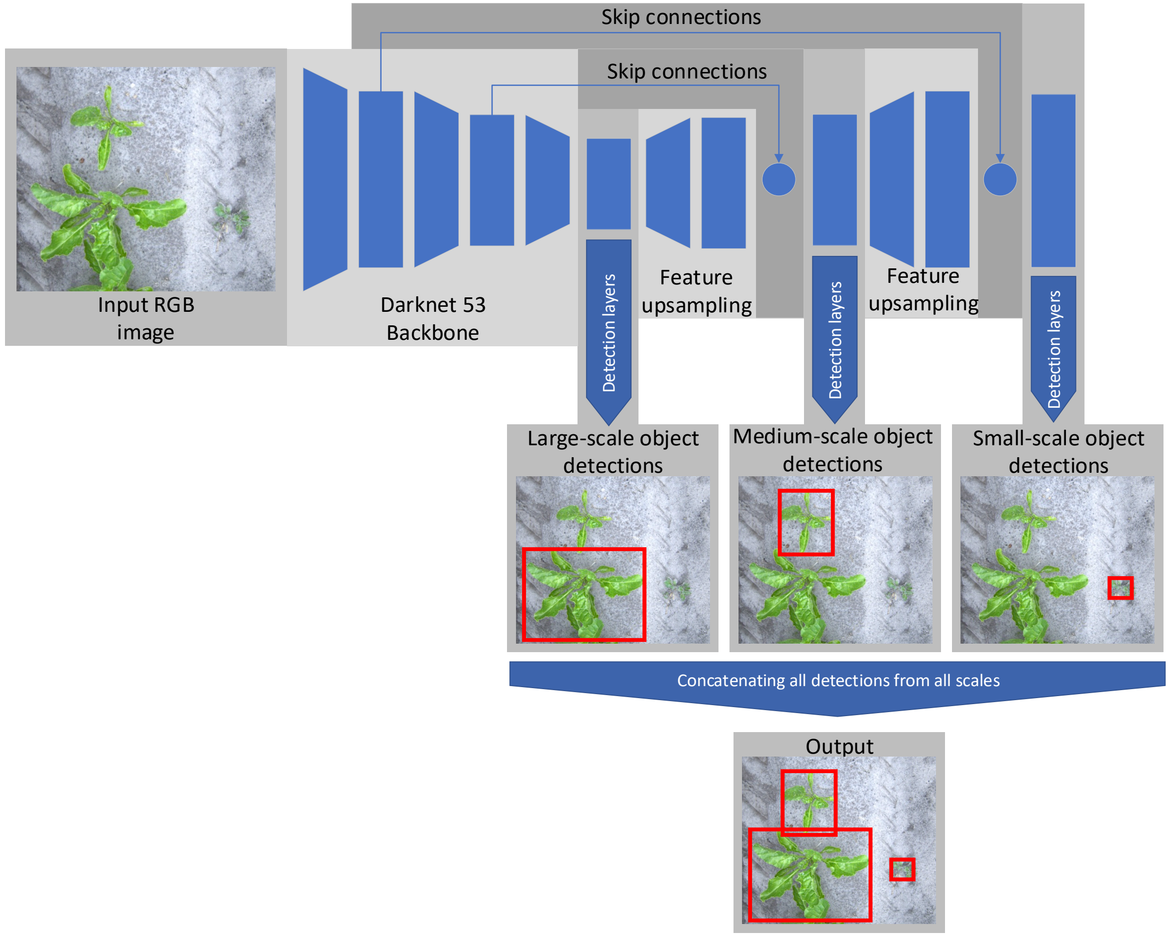
A robotic multi-view camera setup



A robotic multi-view camera setup

Node detection using deep learning

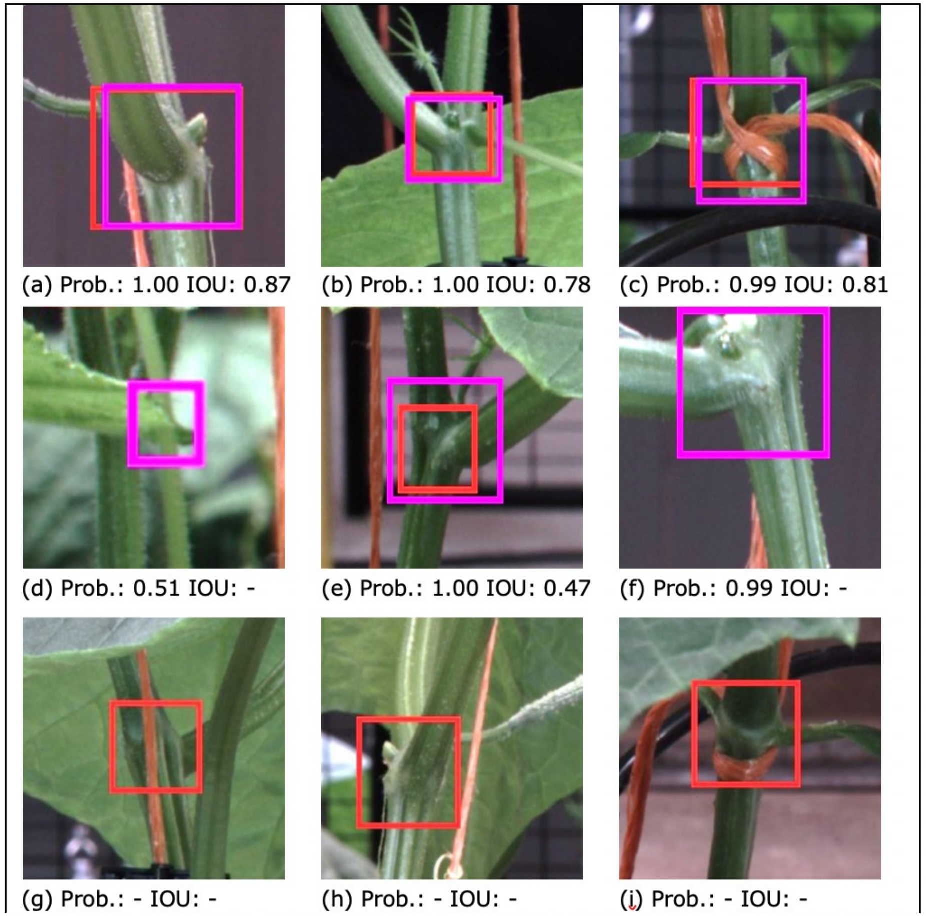
- We used the YOLOv3 neural network for object detection
- Image with in total 10.314 nodes visible
- 4-folds Cross validation with 75-25% split



Step 1: node detection

- Detection of nodes using YOLOv3

CV-run (-)	Validation precision (-)	Validation recall (-)	Validation F1 score (-)
1	0.93	0.91	0.92
2	0.95	0.92	0.94
3	0.96	0.93	0.94
4	0.95	0.93	0.94
Average	0.95	0.92	**0.94**



Step 2: multi-view node clustering

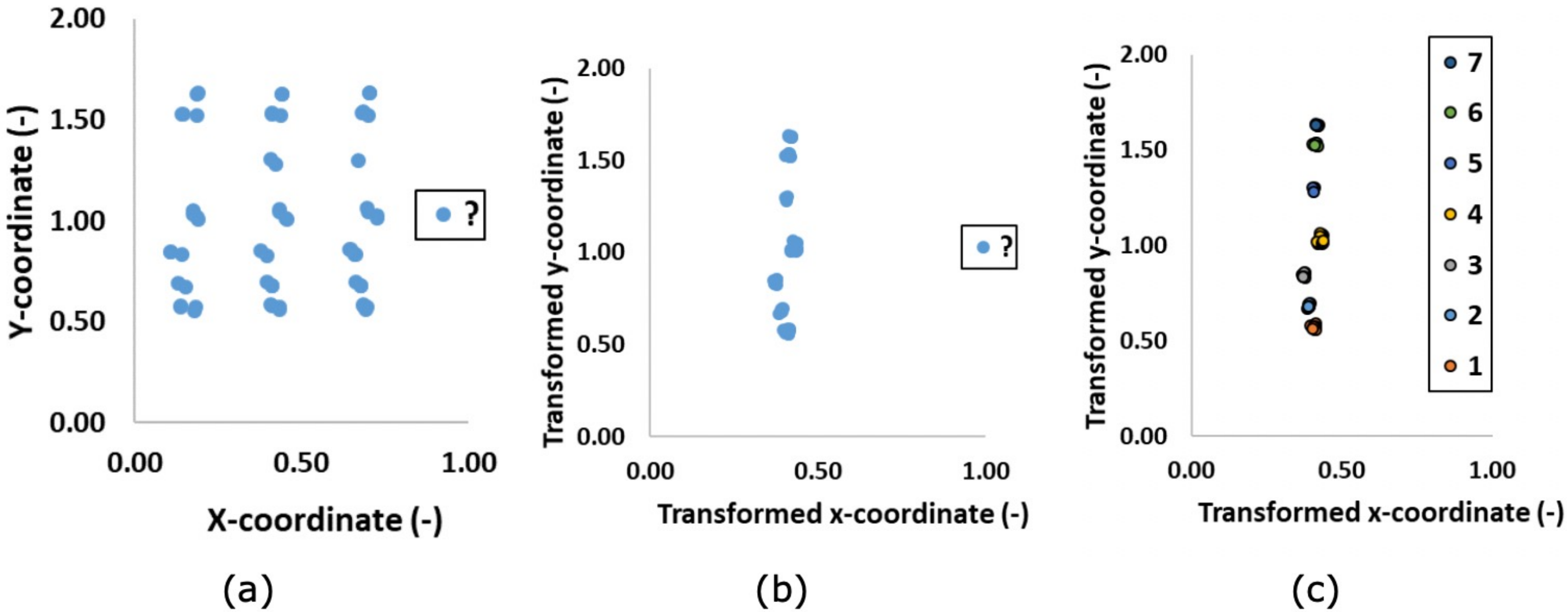
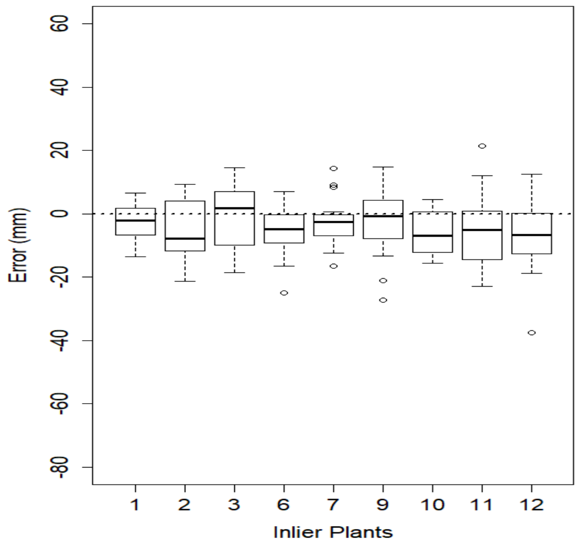


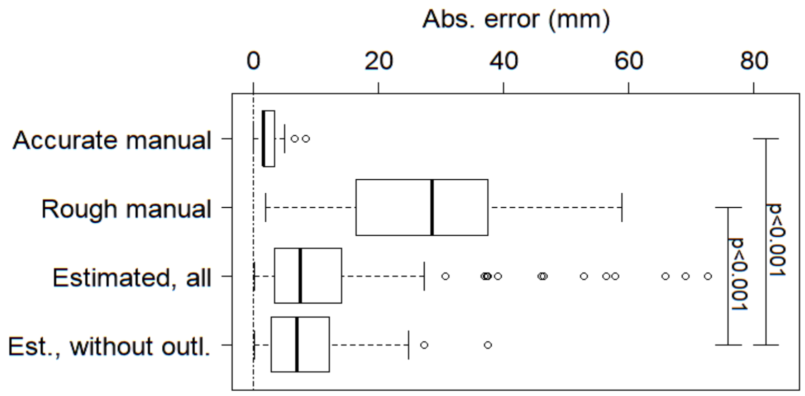
Figure 5 – a) Node coordinates for one plant detected in multiple viewpoints, b) Detected nodes mapped onto the reference coordinate frame and c) clustered node detections in the appropriate order.

WAGENINGEN UNIVERSITY & RESEARCH

Affinity propagation clustering (Frey & Dueck, 2007)

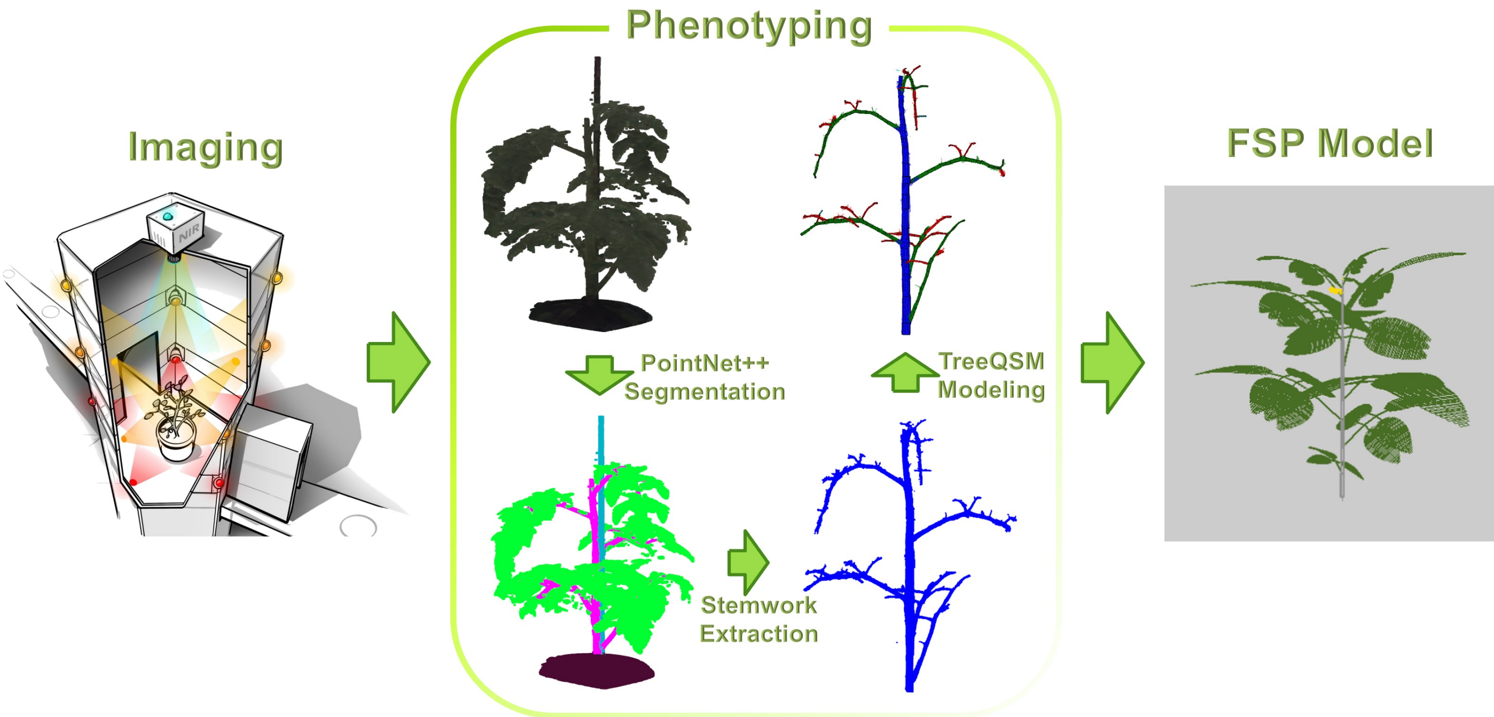
Step 3: Internode-length estimation results





- Error of estimations are generally in range of [-10mm, 10mm]
- This is not as accurate as accurate human measurements
- But more accurate than rough manual estimation
- But…
- … estimation are quite off for non-vertically growing plants
- We need 3D plant phenotyping

Pipeline from 3D imaging data to a digital twin

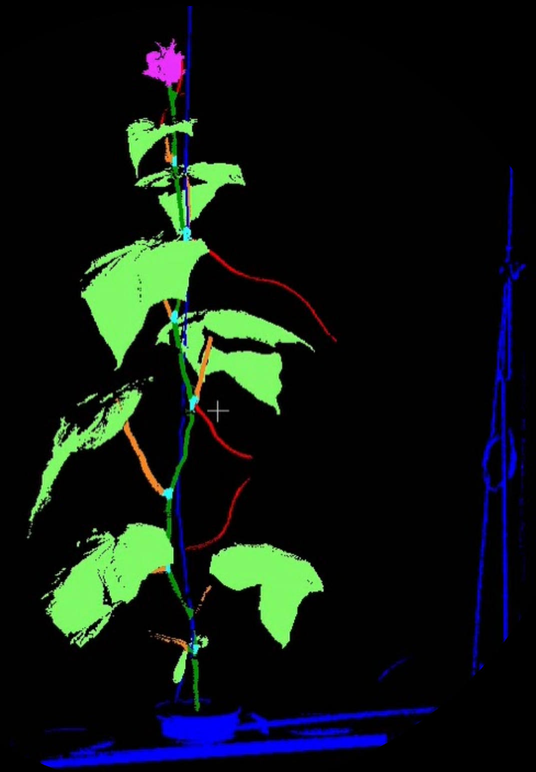


3D data acquisition → Point cloud segmentation → Reconstruction of plant parts → Phenotypic traits

WAGENINGEN UNIVERSITY & RESEARCH

3D data acquisition

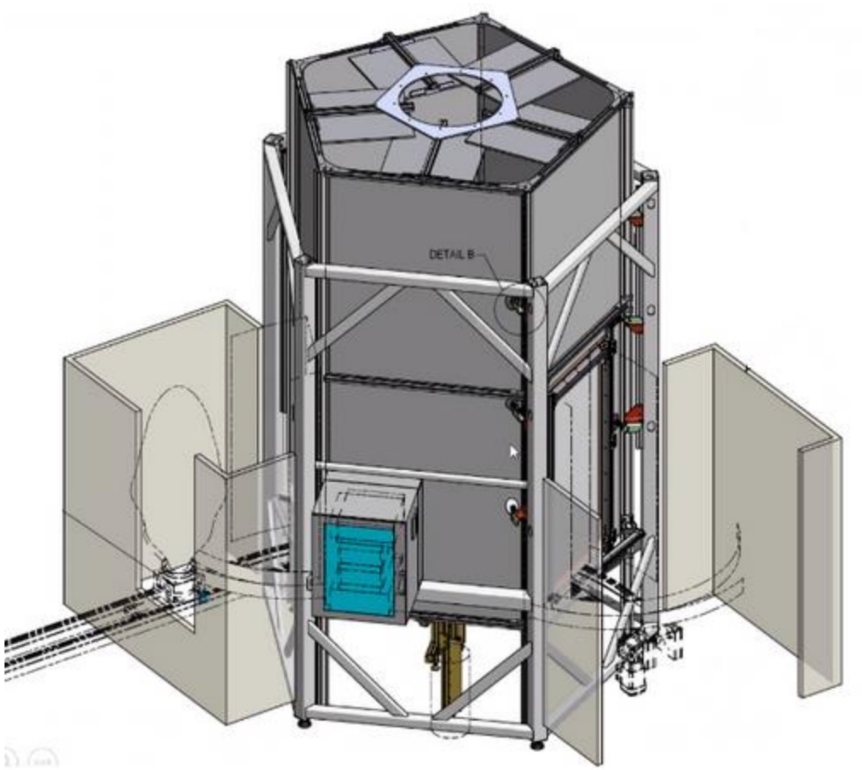
- Different sensors and sensing technologies
- Multi-view approaches

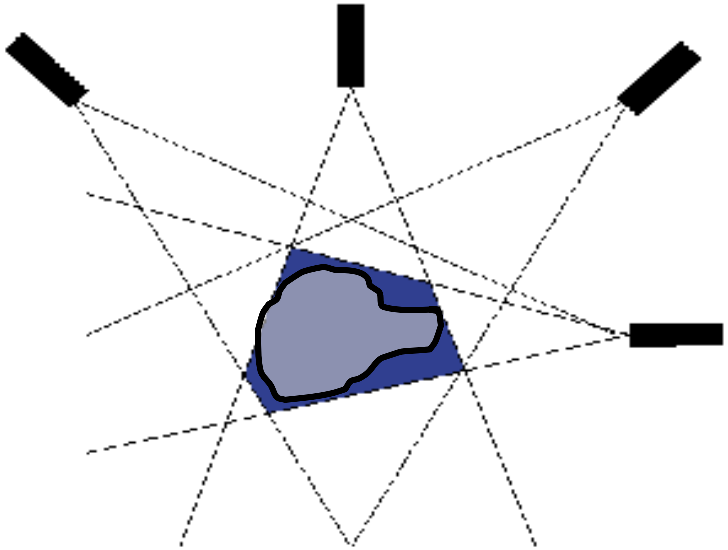


Different 3D sensors

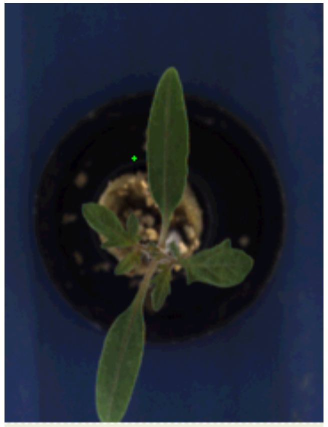
- Shape from silhouette
- Laser triangulation
- Structure from Motion and Multi-View Stereo
- Terrestial laser scanners

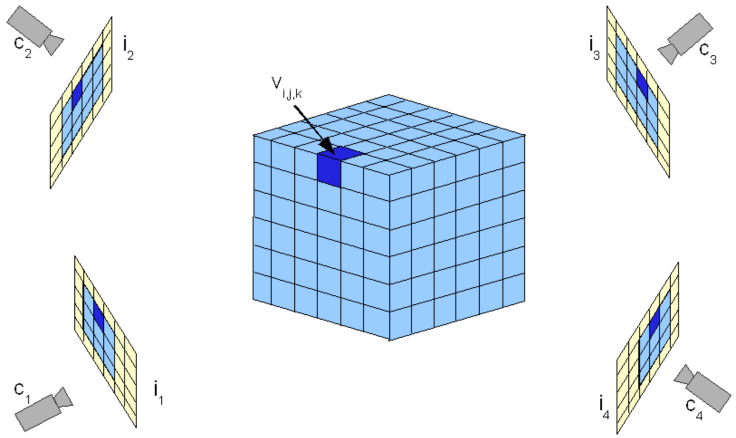
Shape from silhouette: MARVIN





Shape from silhouette





Laser triangulation

- PlantEye F500 of Phenospex
- High-resolution RGBI point cloud
- Partially complete point cloud

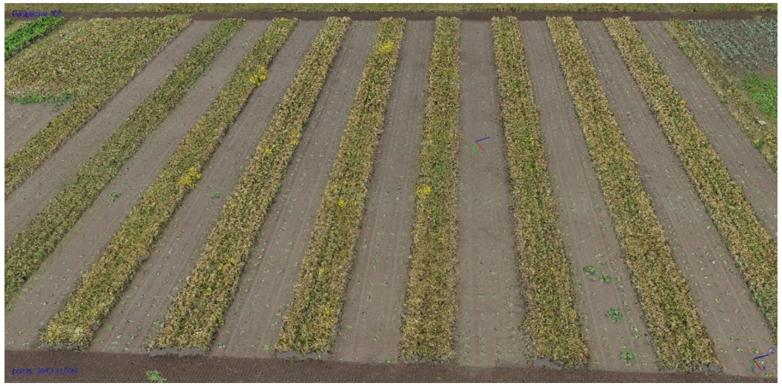
WAGENINGEN
UNIVERSITY & RESEARCH

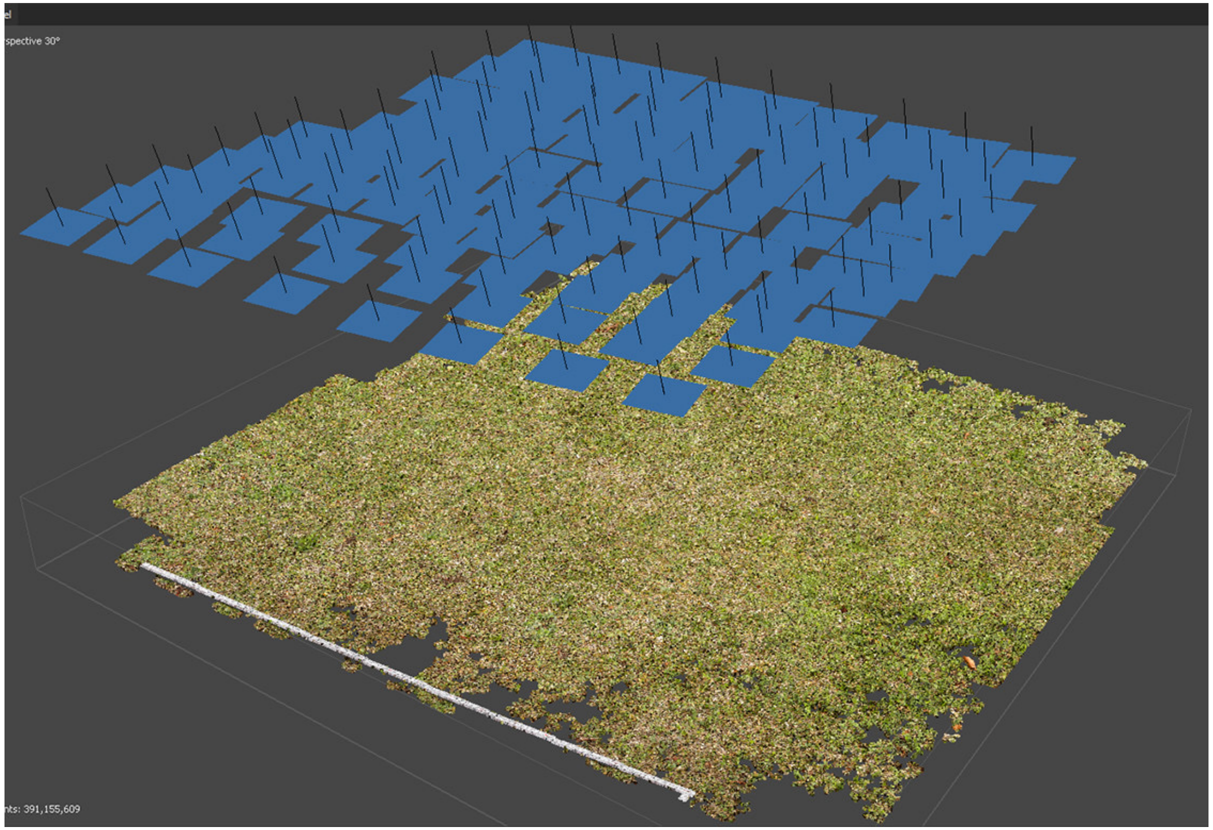
RIJK ZWAAN

Structure from motion and multi-view stereo

300 high-res photos

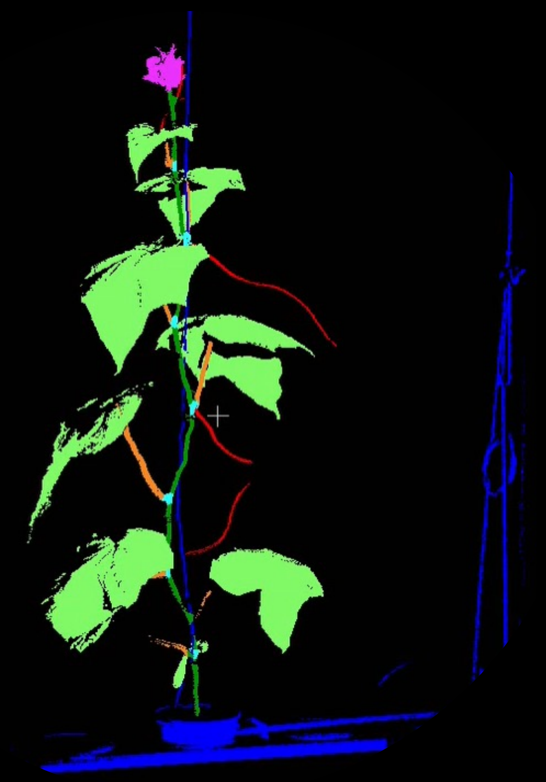
Structure from motion and multi-view stereo





3D data processing

- Multi-view 2D deep learning
- 3D deep learning: spatial vs spectral data

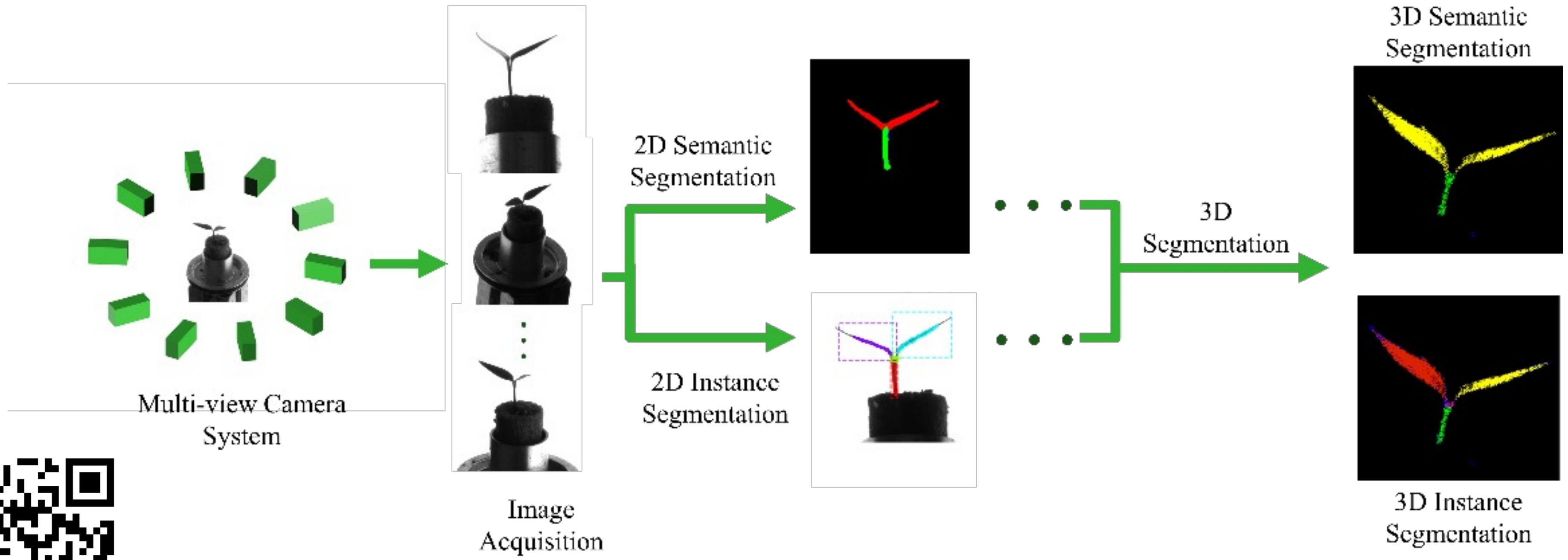


WAGENINGEN
UNIVERSITY & RESEARCH

3D data processing

- Multi-view 2D deep learning

A multi-view deep-learning approach

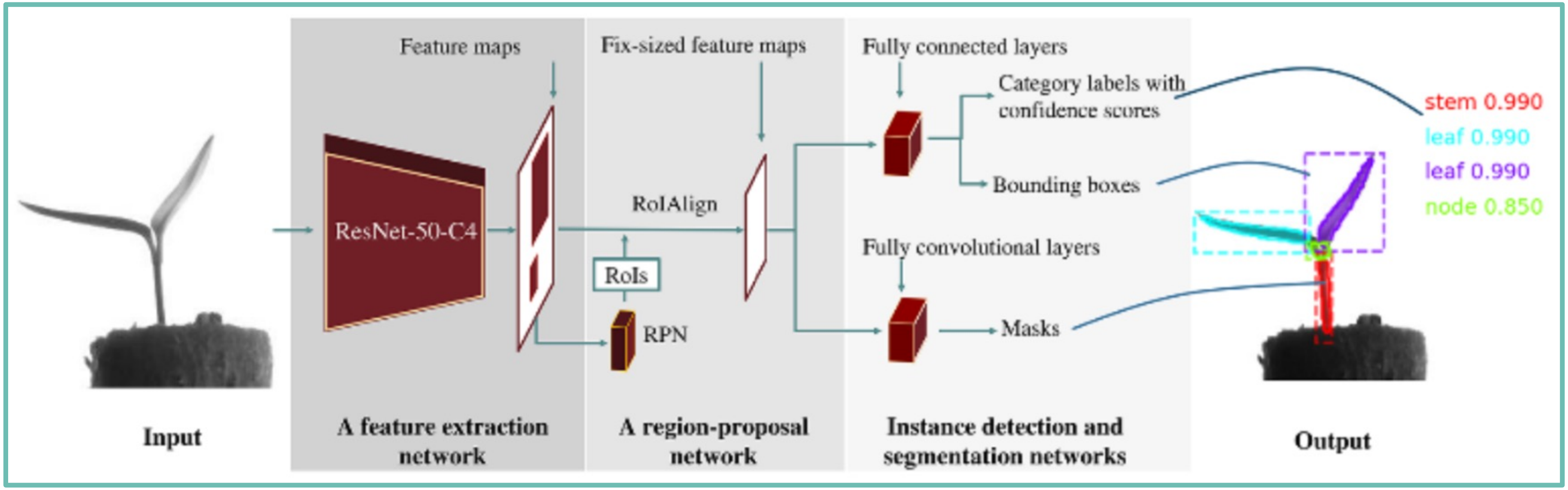


Shi, W., van de Zedde, R., Jiang, H., & Kootstra, G. (2019). Plant-part segmentation using deep learning and multi-view vision. *Biosystems Engineering*,187, 81–95

Instance segmentation

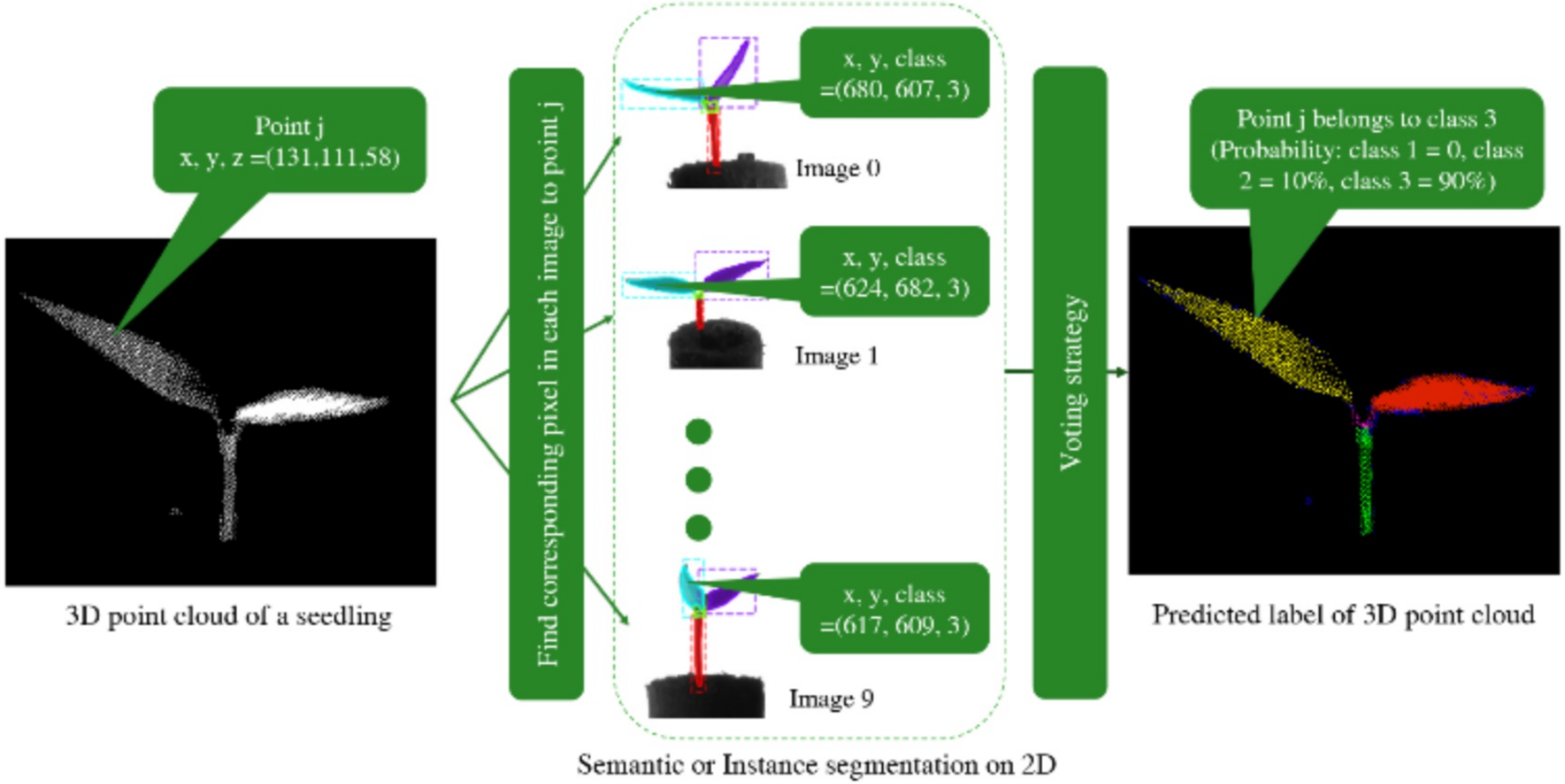
- Mask R-CNN

Ren *et al.* (2017)



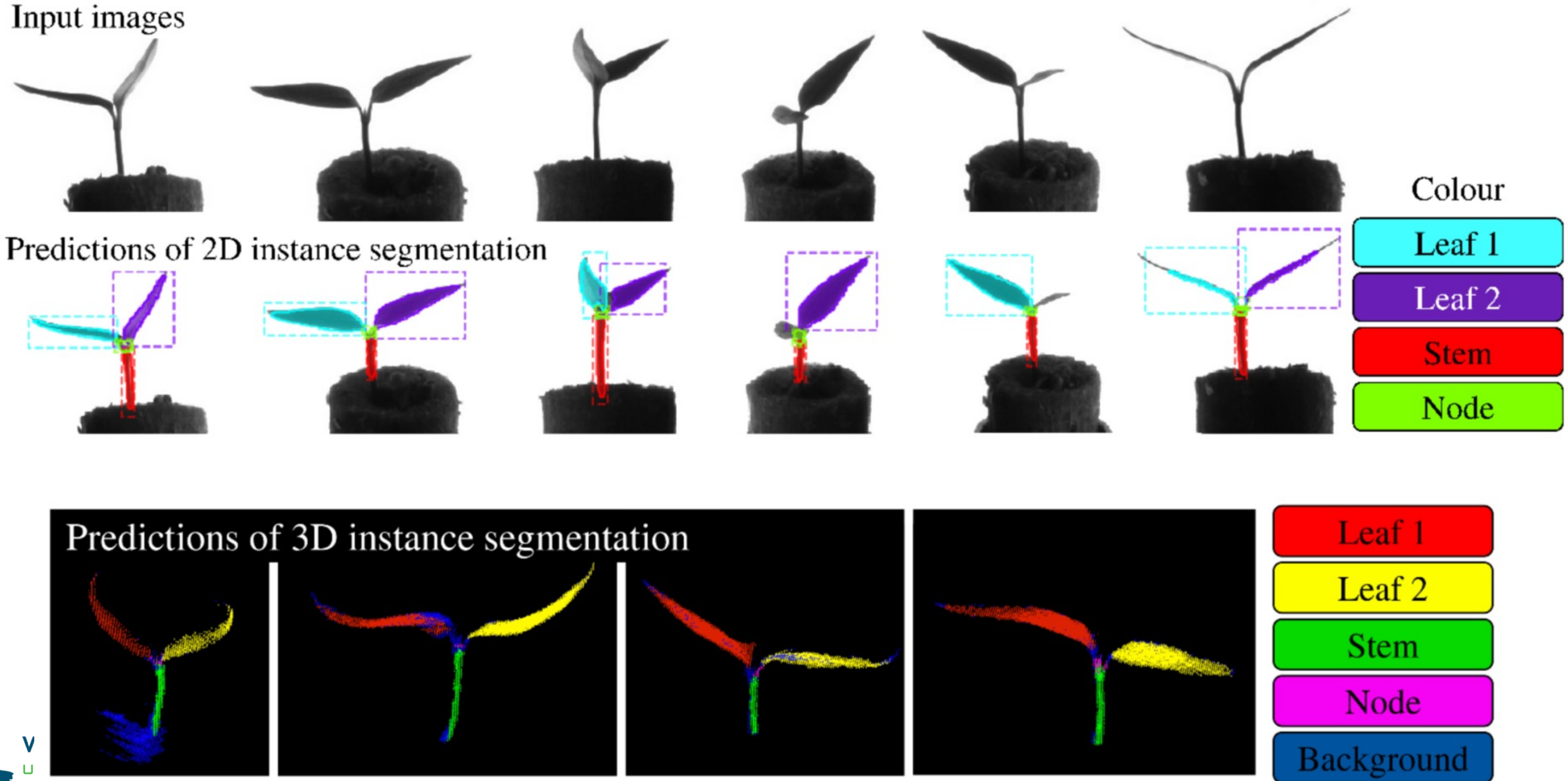
WAGENINGEN UNIVERSITY & RESEARCH

Multi-view voting



Shi, W., van de Zedde, R., Jiang, H., & Kootstra, G. (2019). Plant-part segmentation usingdeep learning and multi-view vision.Biosystems Engineering,187, 81–95

Results instance segmentation



Results instance segmentation

Pixel-wise instance segmentation

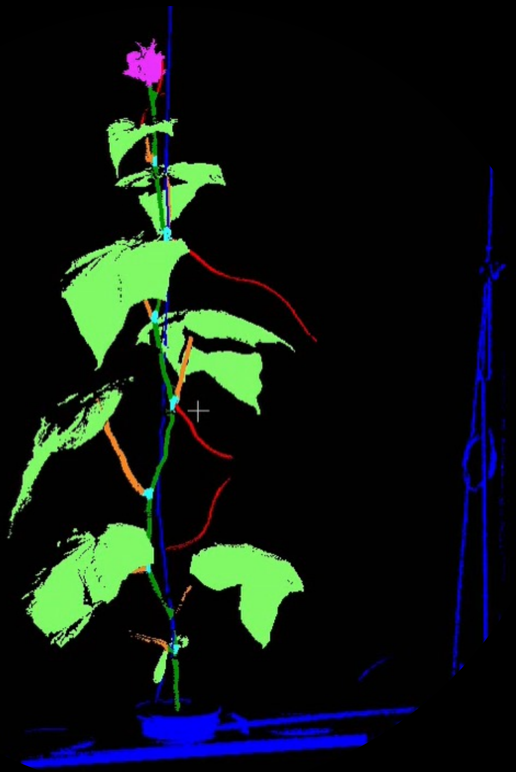
	Precision			Recall			F1-score		
	2D	3D	P-value	2D	3D	P-value	2D	3D	P-value
Stem	0.77 (0.16)	0.97 (0.04)	0.000*	0.65 (0.20)	0.79 (0.09)	0.000*	0.70 (0.15)	0.87 (0.06)	0.000*
Leaf 1	0.95 (0.10)	1.00 (0.00)	0.000*	0.66 (0.19)	0.92 (0.06)	0.000*	0.78 (0.14)	0.96 (0.03)	0.000*
Leaf 2	0.93 (0.13)	1.00 (0.00)	0.000*	0.67 (0.20)	0.89 (0.09)	0.000*	0.78 (0.16)	0.94 (0.05)	0.000*

Object-wise instance detection

		Precision			Recall			F1-score		
		2D	3D	P-value	2D	3D	P-value	2D	3D	P-value
Stem		0.68	1.00	-	0.67	1.00	-	0.68	1.00	-
Leaf		0.83	1.00	-	0.82	1.00	-	0.83	1.00	-
Node	1 mm	0.68	0.96	0.001*	0.65	0.96	0.003*	0.67	0.96	0.002*
Node	2 mm	0.97	1.00	0.177	0.92	1.00	0.126	0.94	1.00	0.141

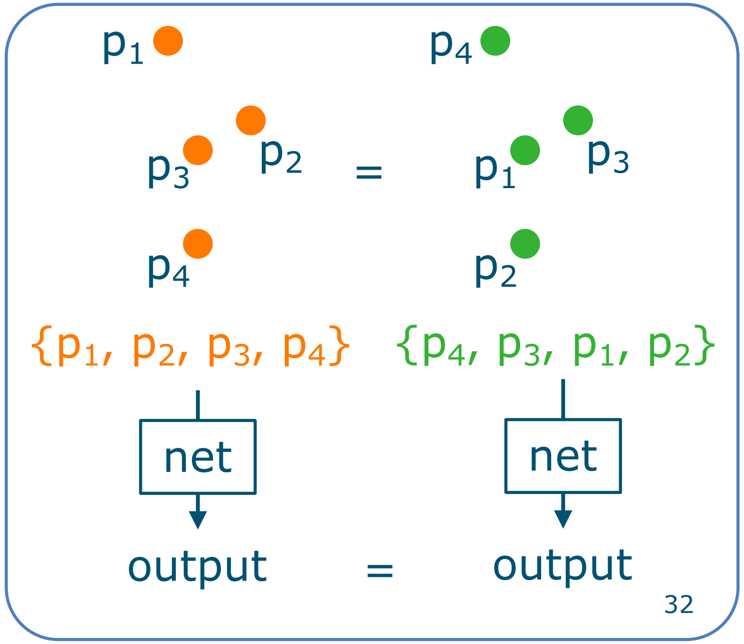
3D data processing

- Deep learning for 3D point clouds



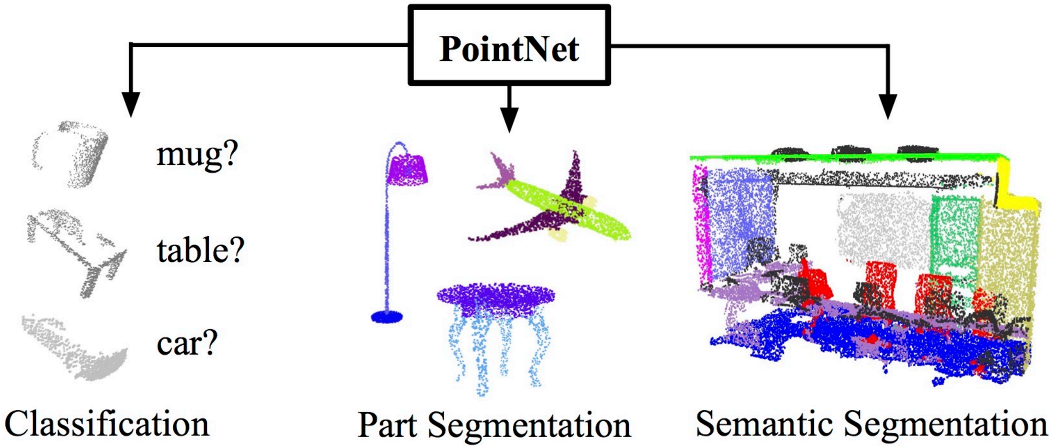
Deep learning for 3D point clouds

- 2D images contain structure: spatial ordering
  - 2D neural networks exploit this structure
- 3D point clouds do not contain this spatial structure
  - Point cloud is a set of points
  - A different permutation of the points is still the same point cloud
- 3D point nets need to be permutation invariant



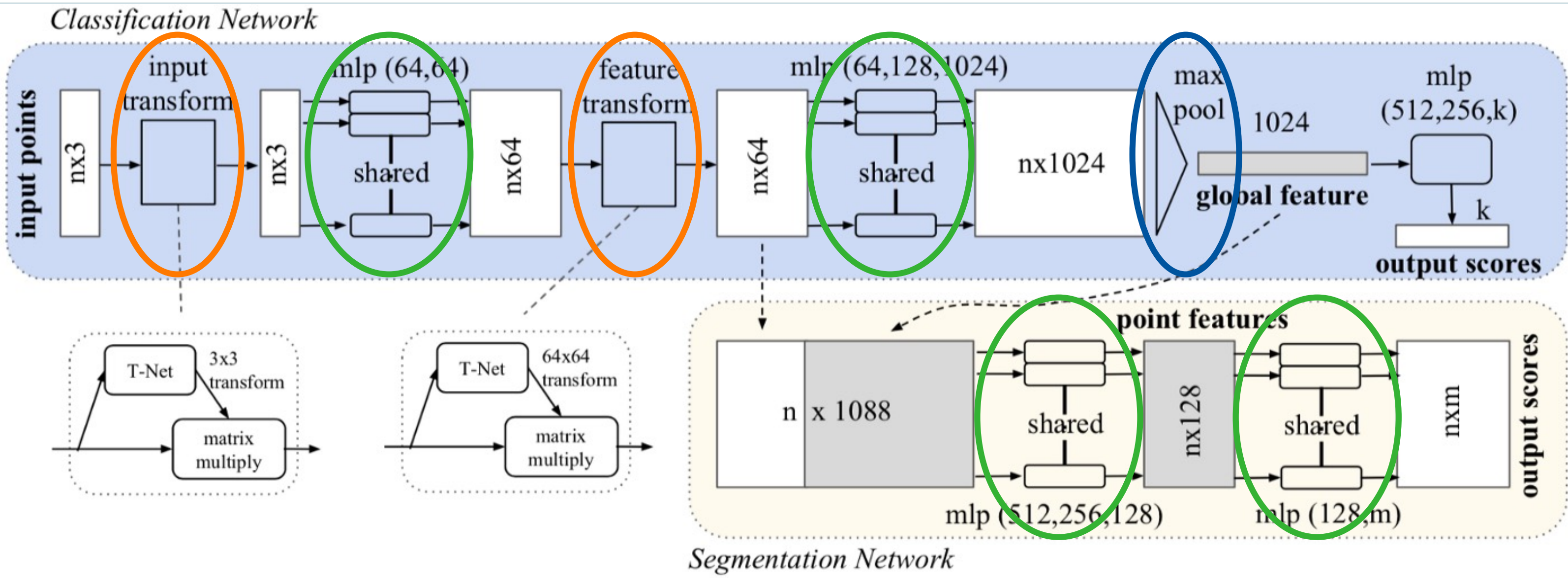
PointNet

- One of the first deep neural networks for 3D point cloud classification and segmentation



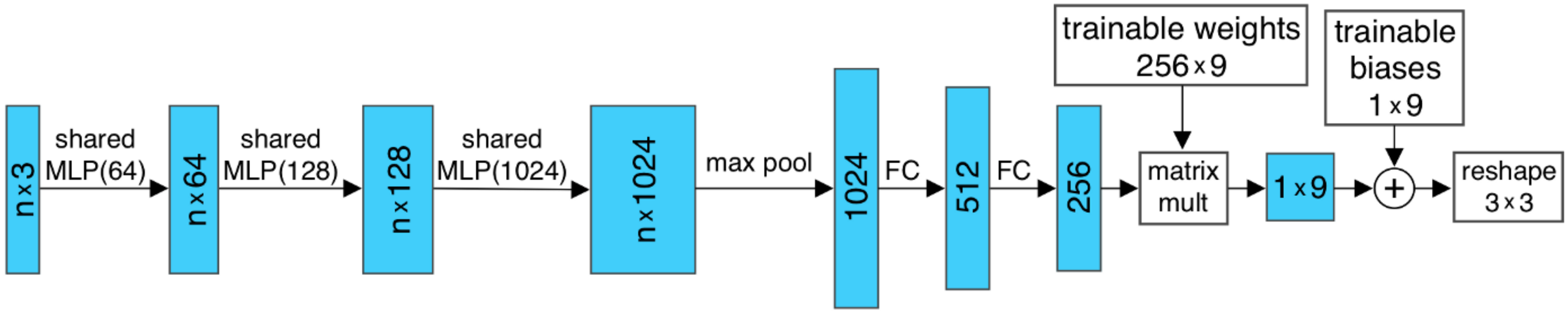
Qi, C.R. et al (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. CVPR 2017

PointNet: Permutation invariance

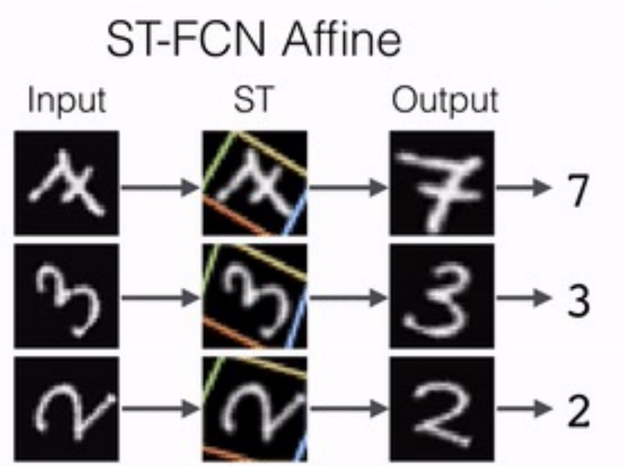


- Multiplication with affine transformation
- Same neural-network operations for all points
- Max pooling

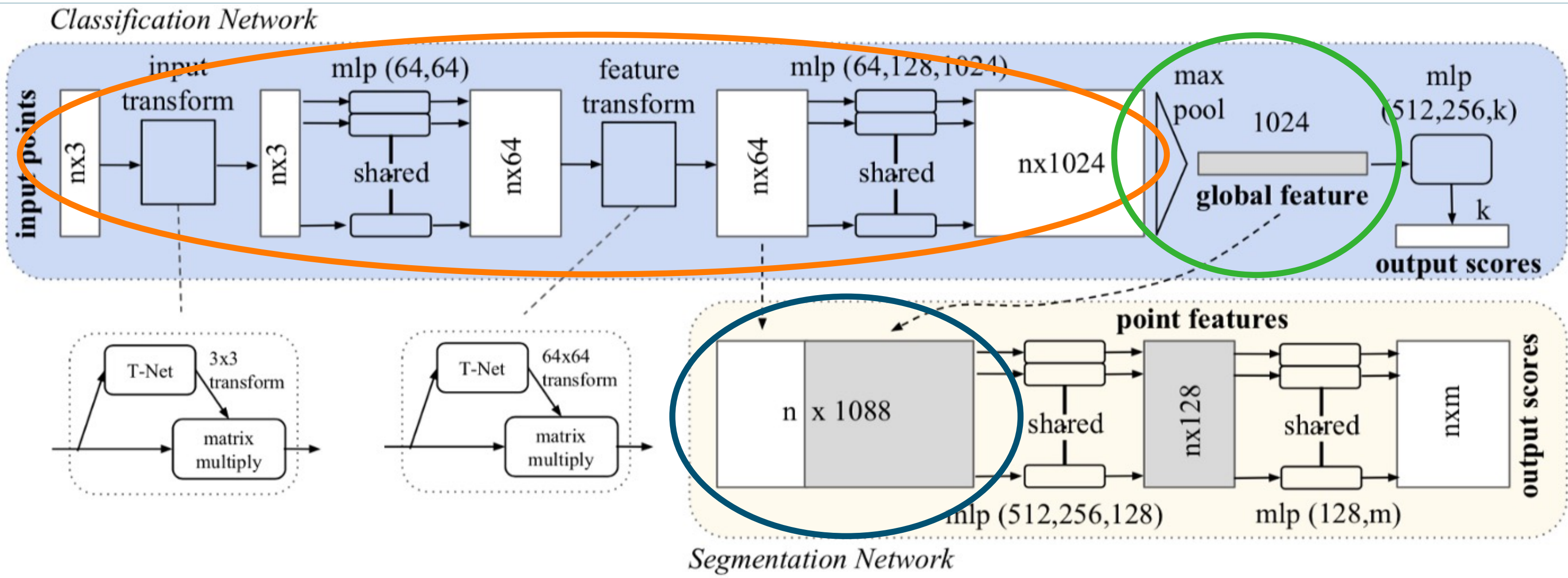
Transformation network



- Learning transformation invariance

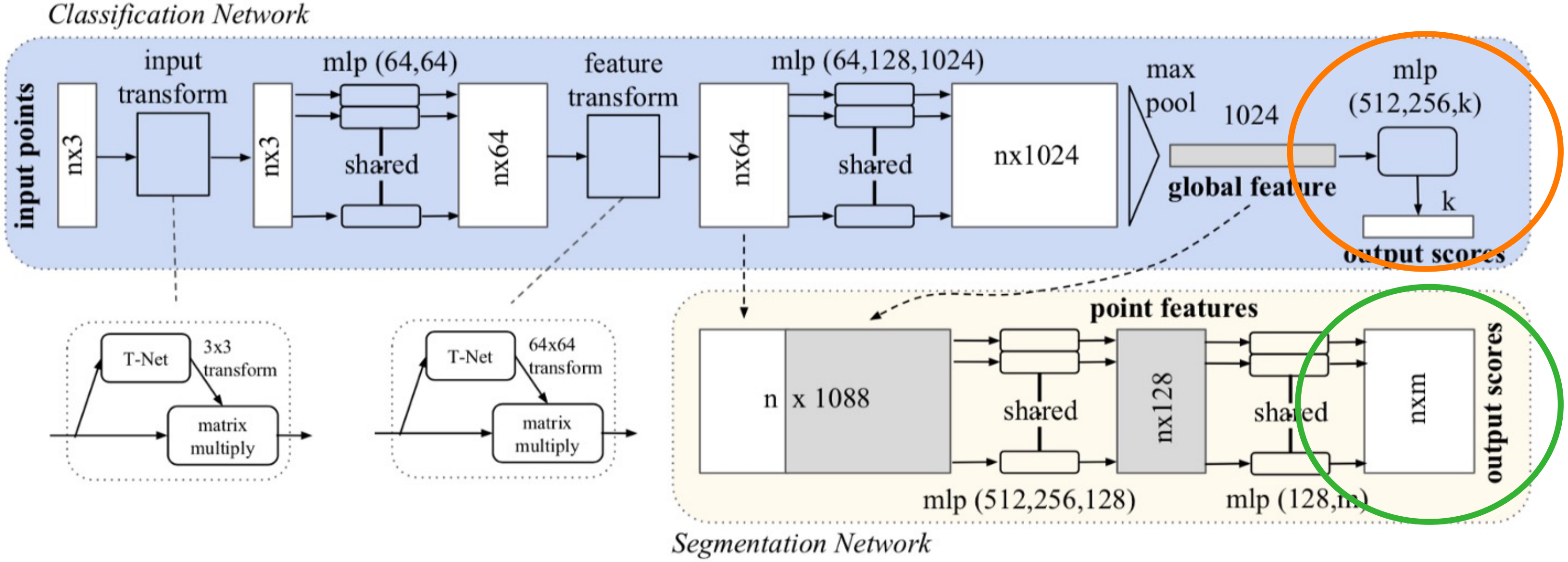


PointNet



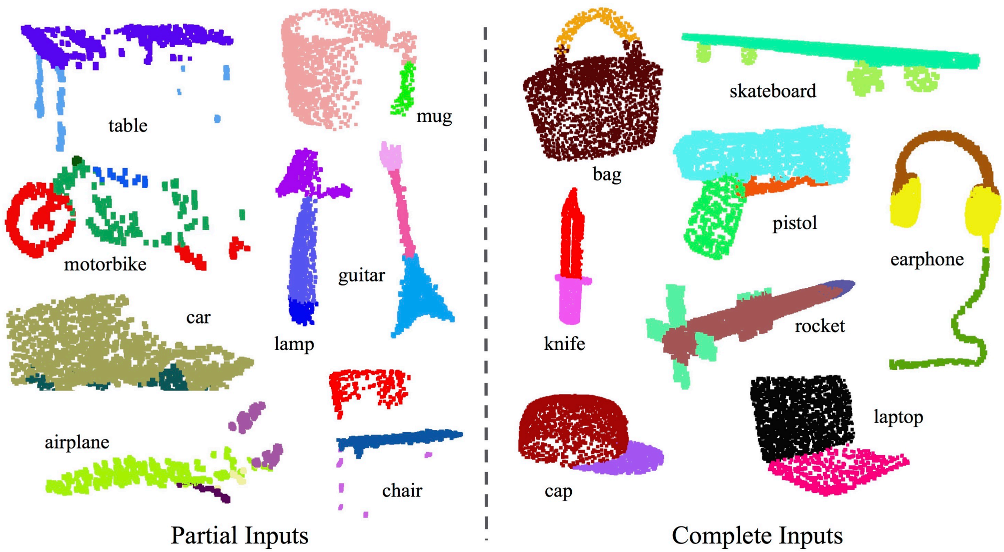
- Point features
- Global feature
- Combined features

PointNet



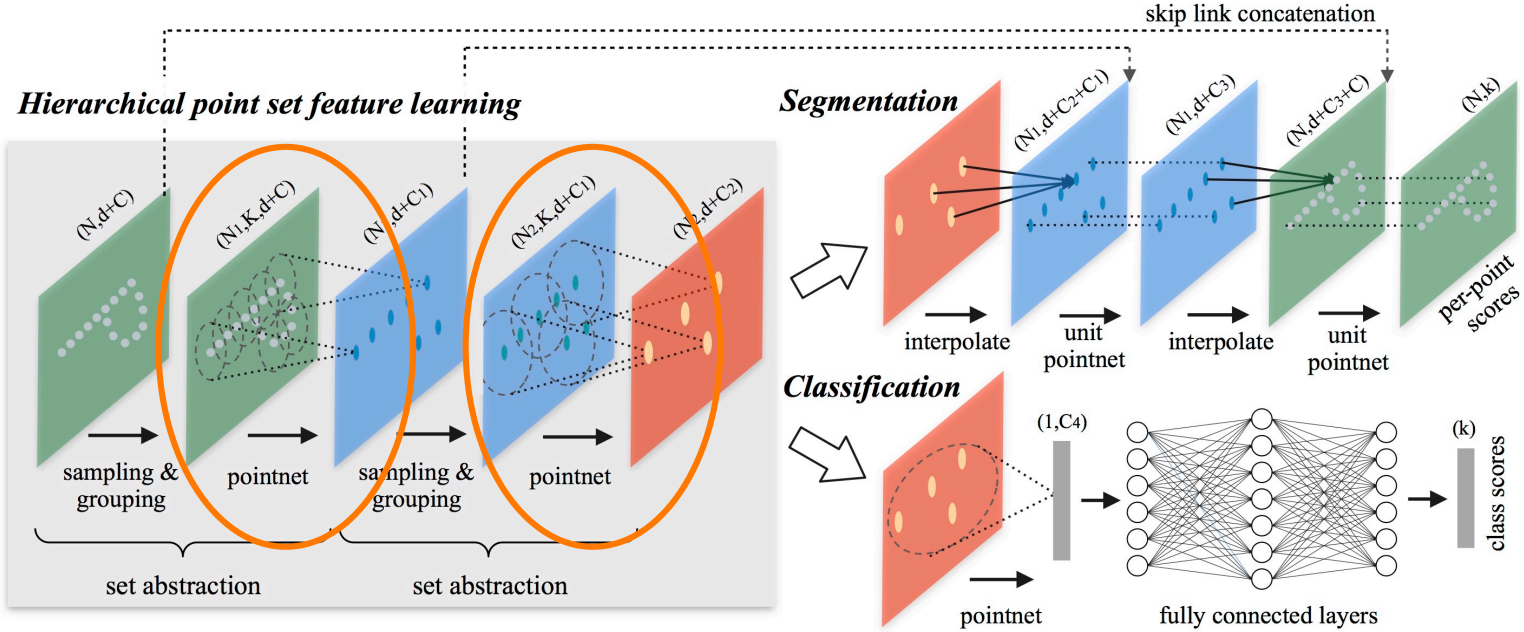
- Classification — What is this point cloud?
- Semantic segmentation — What is every point in the cloud

Some results



PointNet++

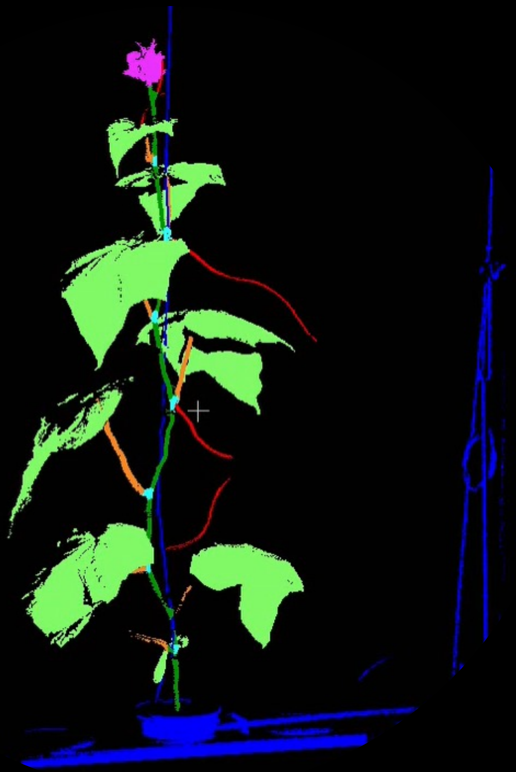
- PointNet is very flat: point level and global level
- PointNet++ creates more levels in a hierarchical structure



Qi, C.R. et al (2017) PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. NIPS 2017

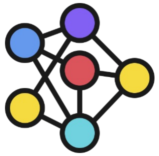
3D data processing

- 3D deep learning: spatial vs spectral data



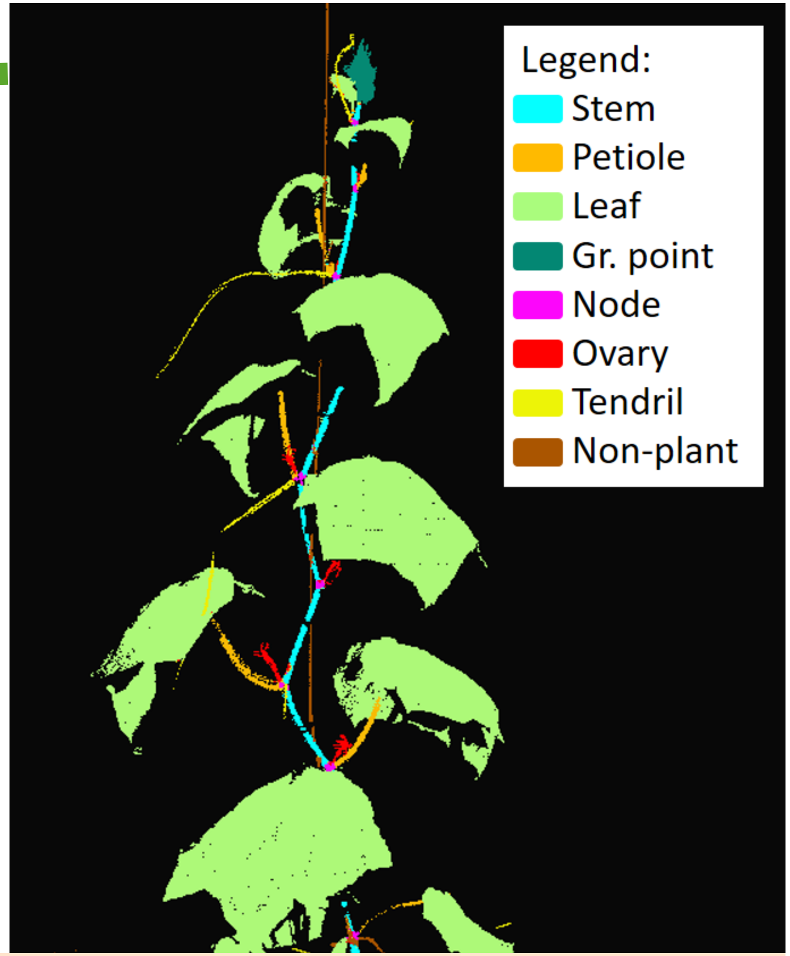
3D segmentation: PointNet++

3D + RGBI scan



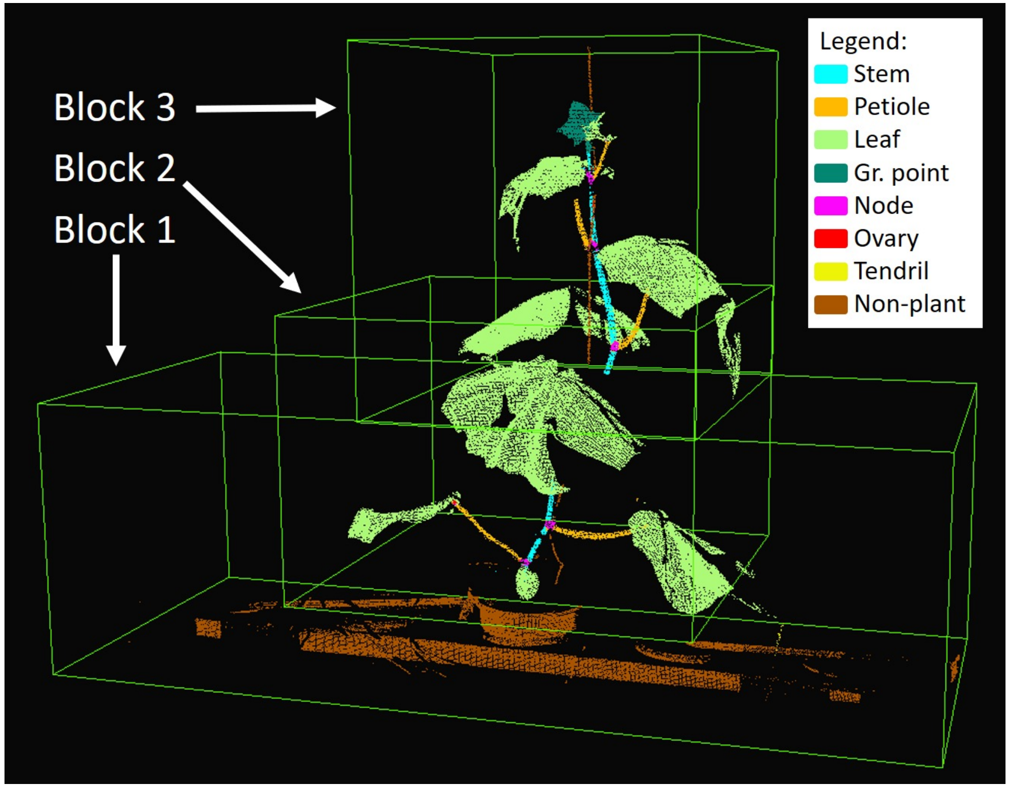
PointNet++

Segment into



Boogaard, F., van Henten, E.J. & Kootstra, G. (2021) Boosting plant-part segmentation from incomplete point clouds of cucumber plants by enriching the point clouds with spectral data. Biosystems Engineering, 211: 167-182. https://doi.org/10.1016/j.biosystemseng.2021.09.004

3D segmentation: data



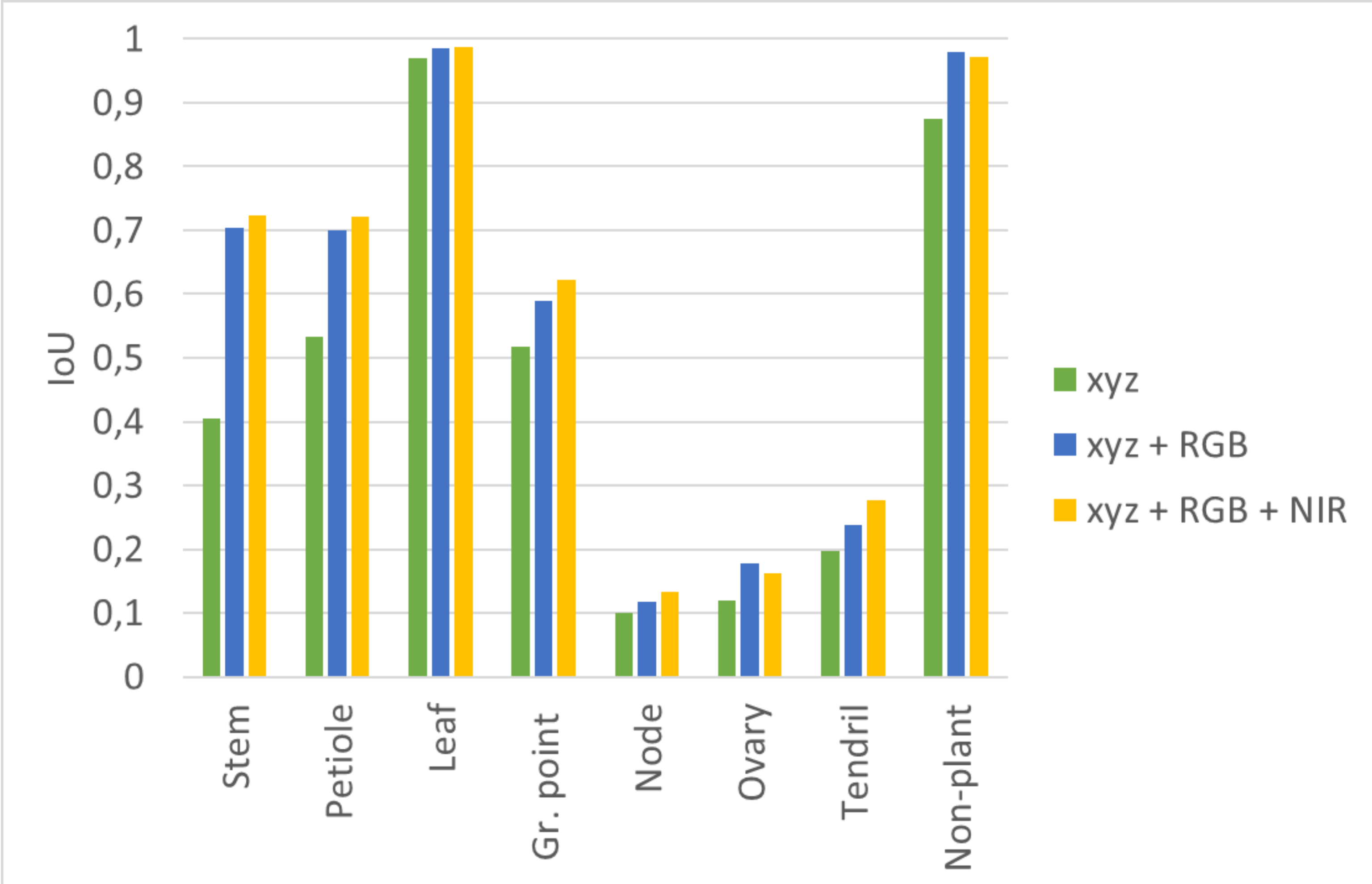
- 12 plants, 11 days, 2 sides
- 264 point clouds
- 200,000-700,000 points
- Voxel filter (2x2x2 mm^3)
- Split in blocks of 40,000
- Annotated manually twice

WAGENINGEN UNIVERSITY & RESEARCH

Boogaard, F., van Henten, E.J. & Kootstra, G. (2021) Boosting plant-part segmentation from incomplete point clouds of cucumber plants by enriching the point clouds with spectral data. *Biosystems Engineering*, 211: 167-182

3D segmentation: using spectral data

Mean IoU between manually segmented data and predictions



WAGENINGEN
UNIVERSITY & RESEARCH

Intra-observer variability

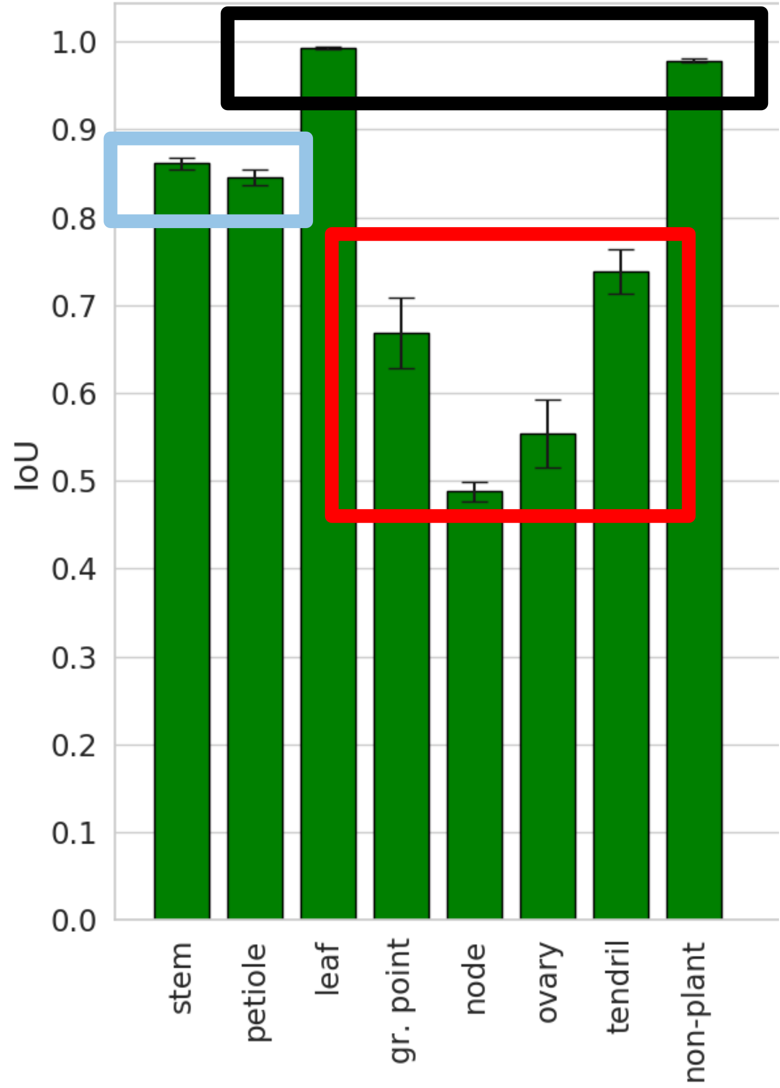
Intersection-over-Union (IoU) between two manual annotations

Presence in dataset:

- Leaf 83.7%
- Non-plant 9.5%

- Stem 2.8%
- Petiole 1.9 %

- Growing point 0.5%
- Node 0.5%
- Ovary 0.5%
- Tendril 0.6%

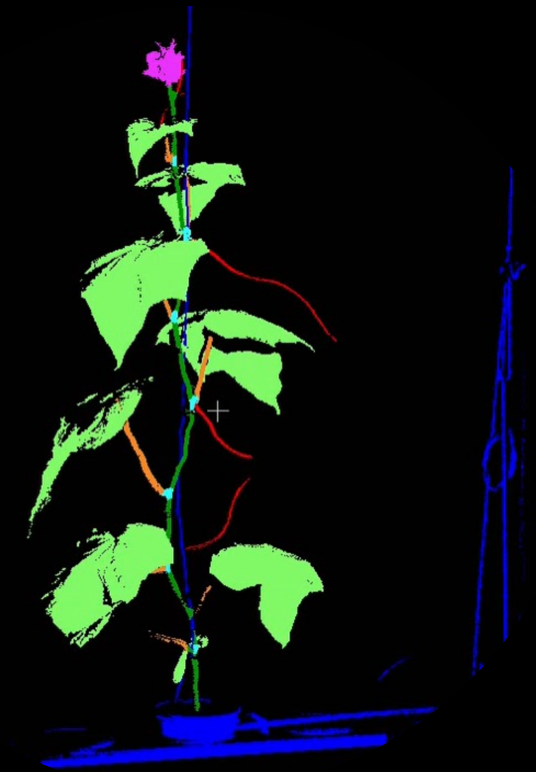


3D segmentation: conclusions

- Spectral data boosts segmentation performance
- Stem, petiole, leaf, growing point and non-plant material can be segmented well
- Node, ovary and tendril are difficult to segment accurately
- This corresponds with the intra-observer variability

Challenges in 3D plant phenotyping

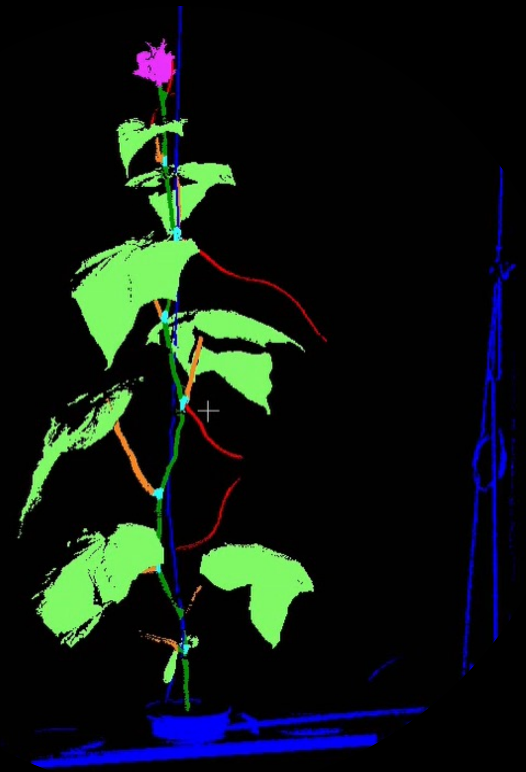
- Class imbalance
- Limited training data
- Variation
- Occlusion



WAGENINGEN
UNIVERSITY & RESEARCH

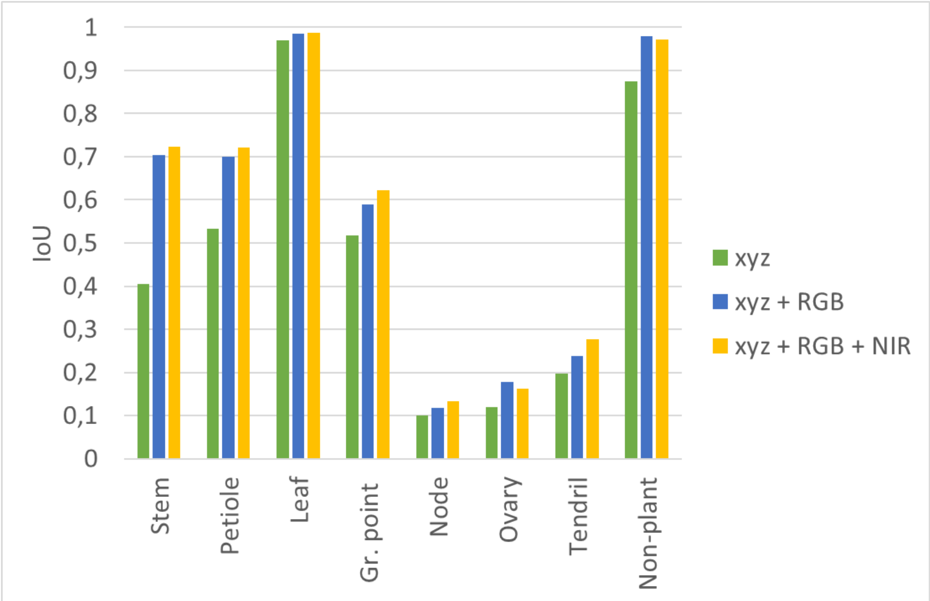
Challenges in 3D plant phenotyping

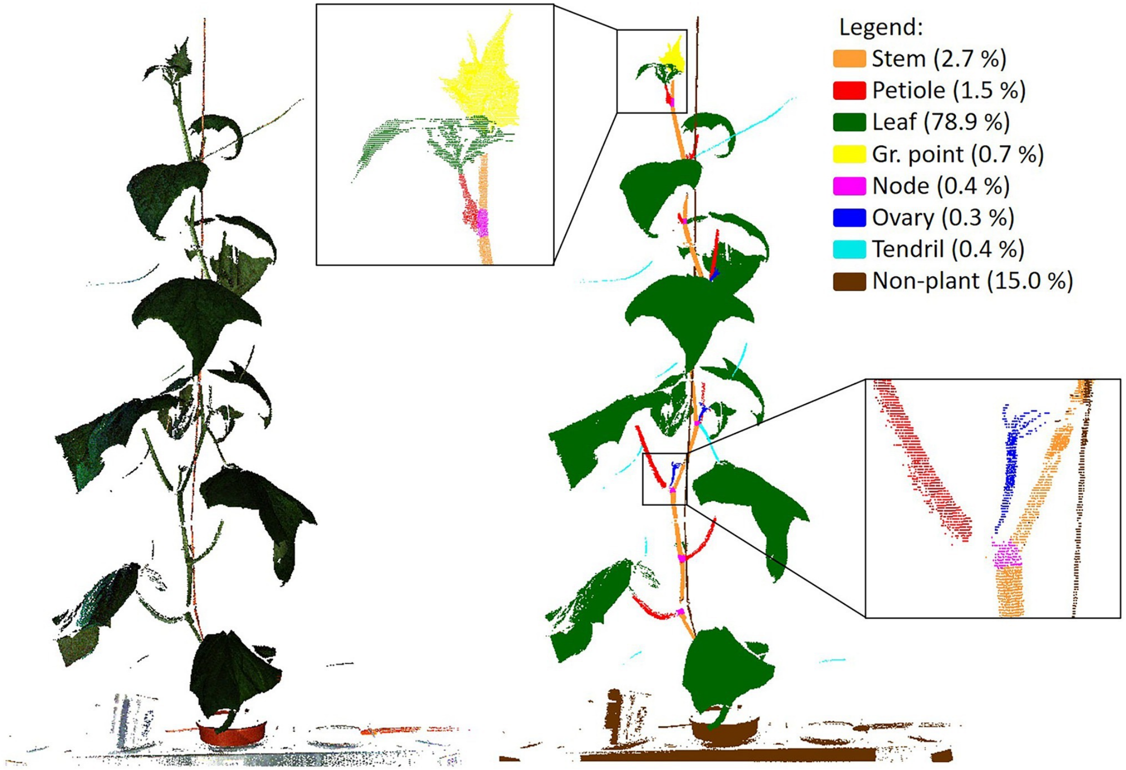
- Class imbalance



WAGENINGEN
UNIVERSITY & RESEARCH

Class imbalance has influence on performance

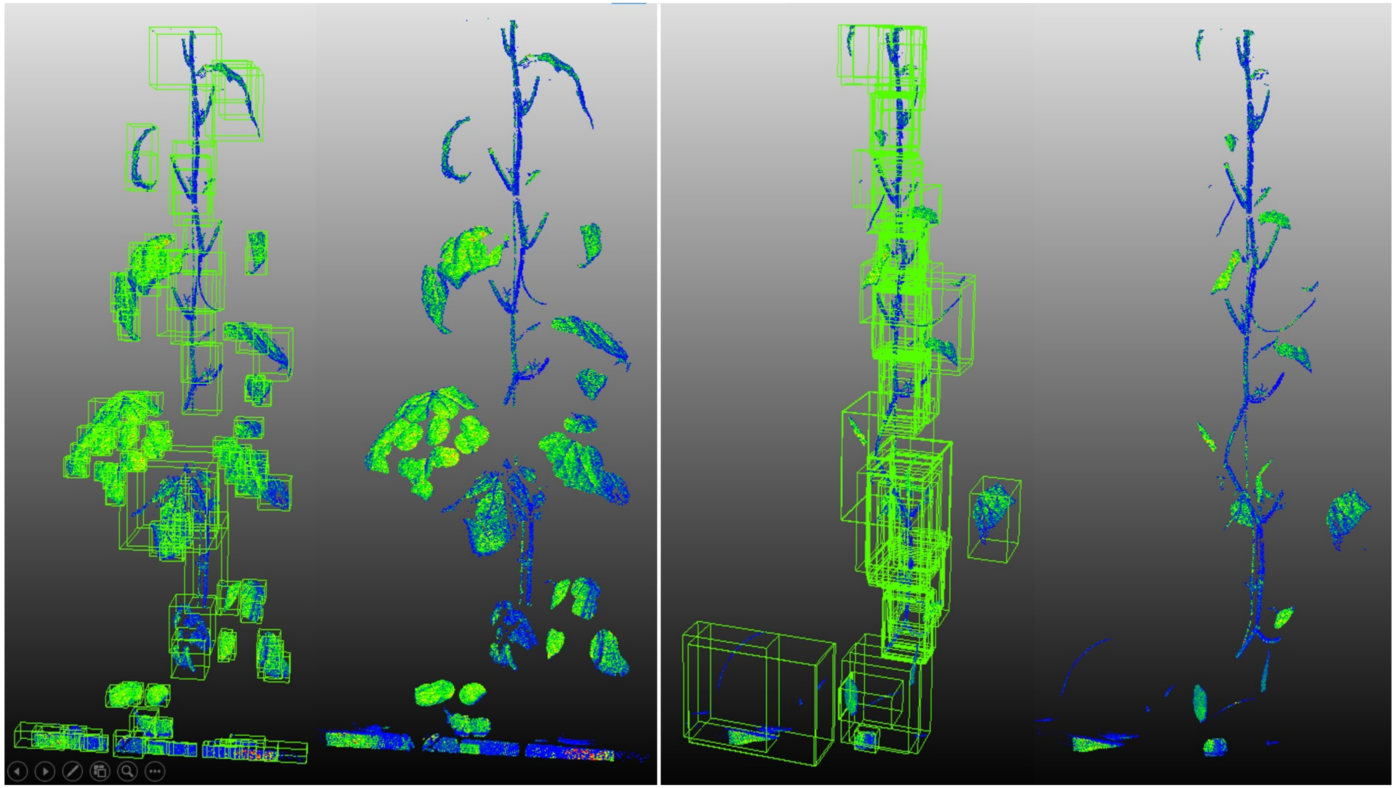




Boogaard, F., van Henten, E.J. & Kootstra, G. (2021) Improved Point-Cloud Segmentation for Plant Phenotyping Through Class-Dependent Sampling of Training Data to Battle Class Imbalance. Frontiers in Plant Science, Sec. Technical Advances in Plant Science.

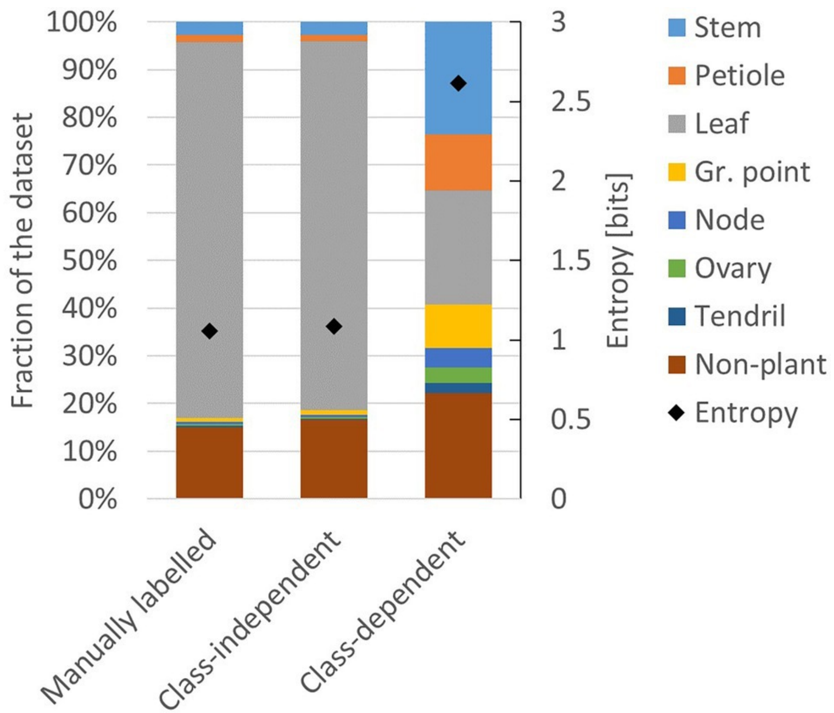
Class-dependent sampling

- Sample points selected with inverse relation to size of the class
- Select the N closest points to the sample point as training sample

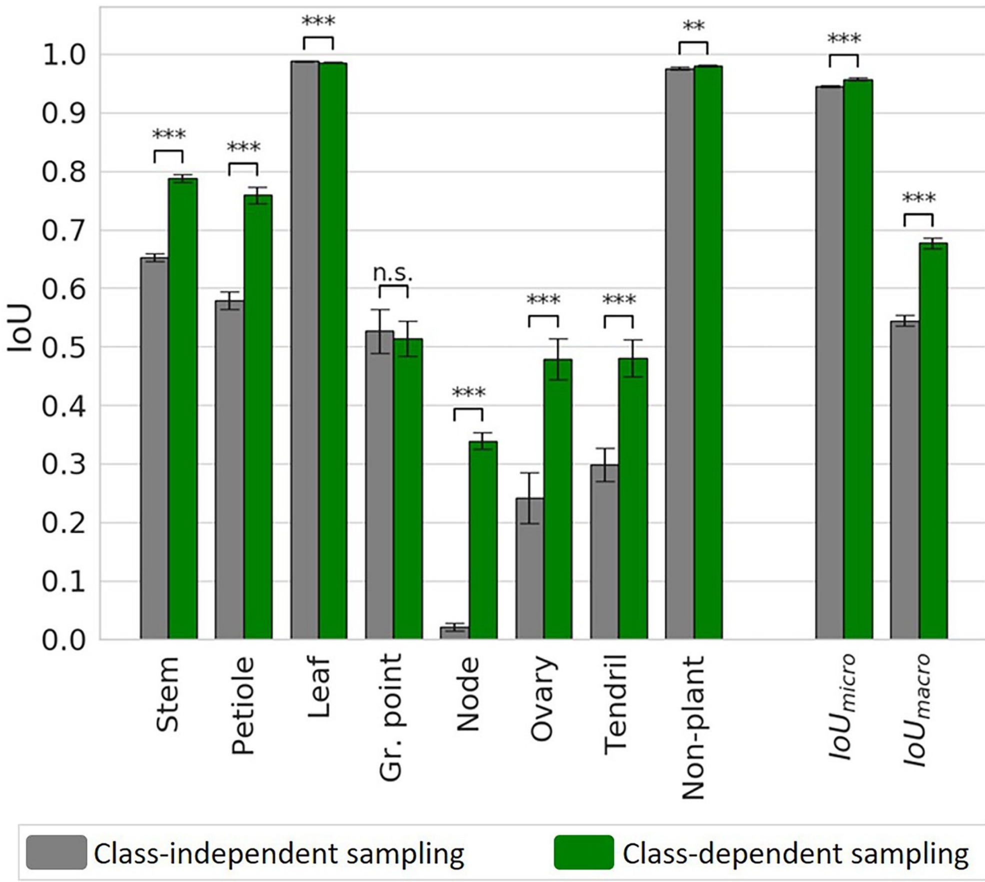


Class-dependent sampling

- More balanced training set



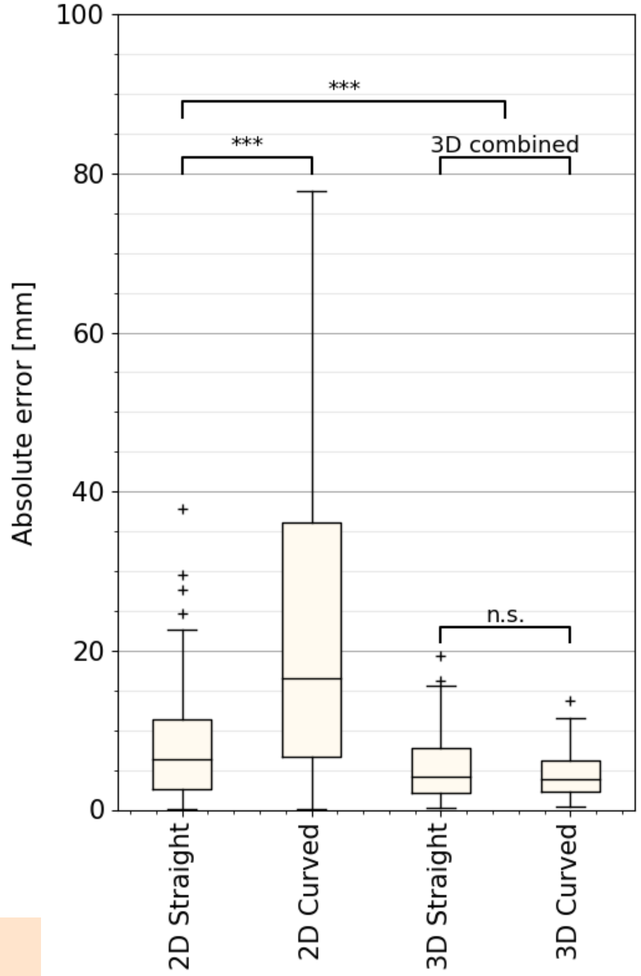
- Improved performance



WAGENINGEN
UNIVERSITY & RESEARCH

Results: Internode-length estimation

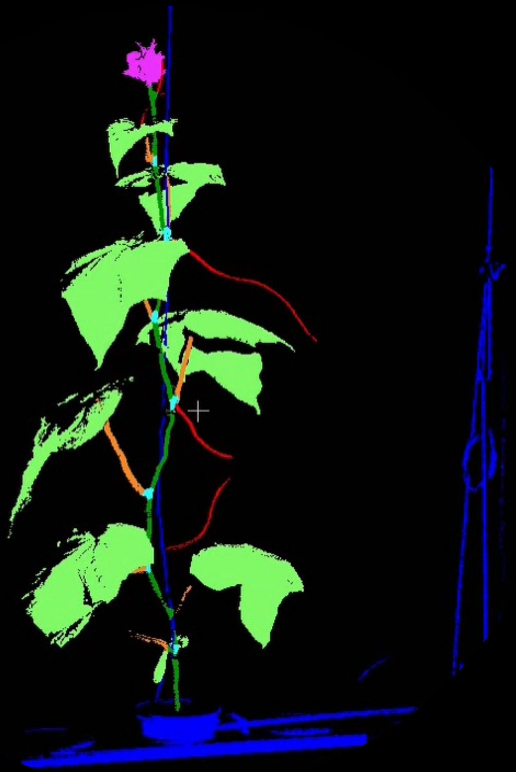
- With 3D points clouds, using spectral data and dealing with class imbalance, we can estimate the internode length automatically…
- … also of the curved plants
- Outperforming 2D plant phenotyping



Boogaard, F., van Henten, E.J. & Kootstra, G. (in prep) The added value of 3D point clouds for digital plant phenotyping – a case study on internode length measurements in cucumber.

Challenges in 3D plant phenotyping

- Limited training data

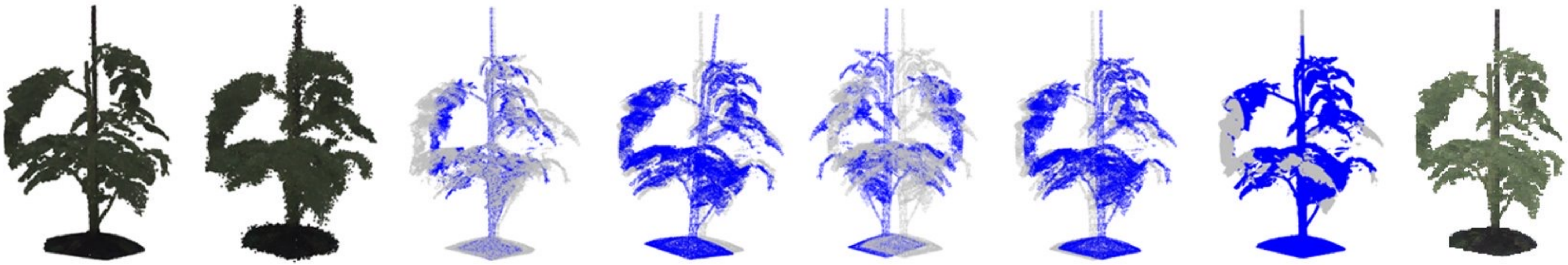


Small datasets

- Annotation of 3D point-cloud data is very time consuming
- This results in small dataset
- Running the risk of
    - Overfitting on the training set
    - Poor generalization to the test set

- Use of data augmentation
    - Artificially increasing the variation in the training set

3D data augmentation: global

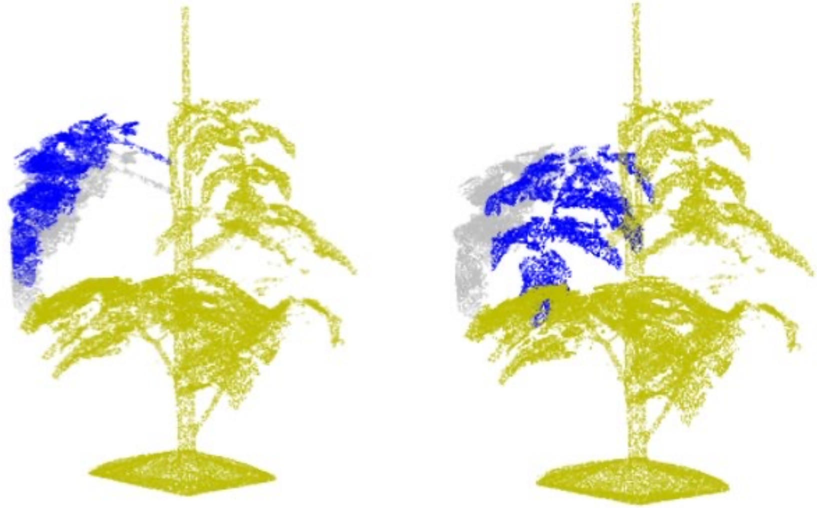


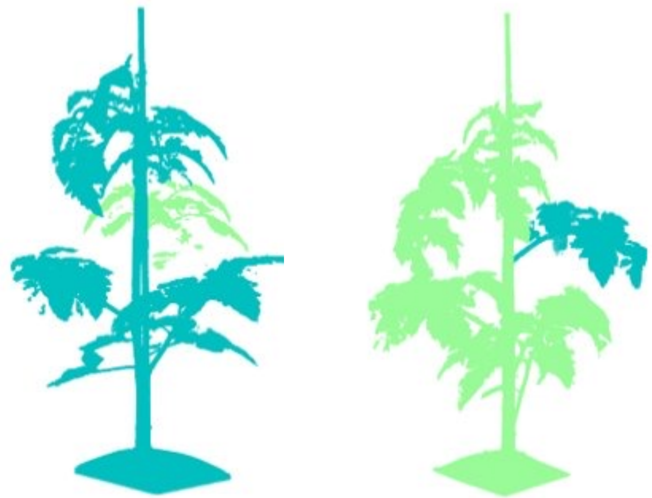
- Down-sample
- Jitter
- Scaling
- Rotation
- Translation
- Cropping
- Brightness

Xin, B, Bartholomeus, H. and Kootstra, G. (review) 3D Data-Augmentation Methods for Semantic Segmentation of Tomato Plant Parts. *Frontiers in Plant Science*.

3D data augmentation: local

- Leaf translation
- Leaf rotation
- Leaf cross-over



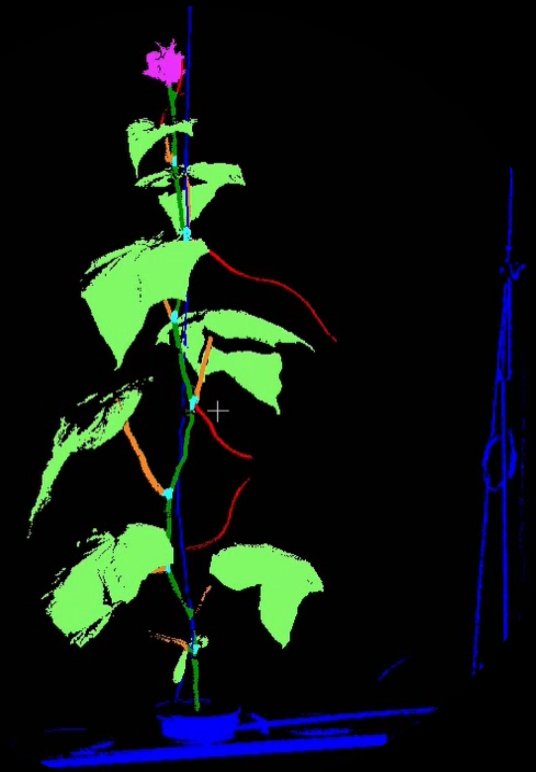


(j) (e)

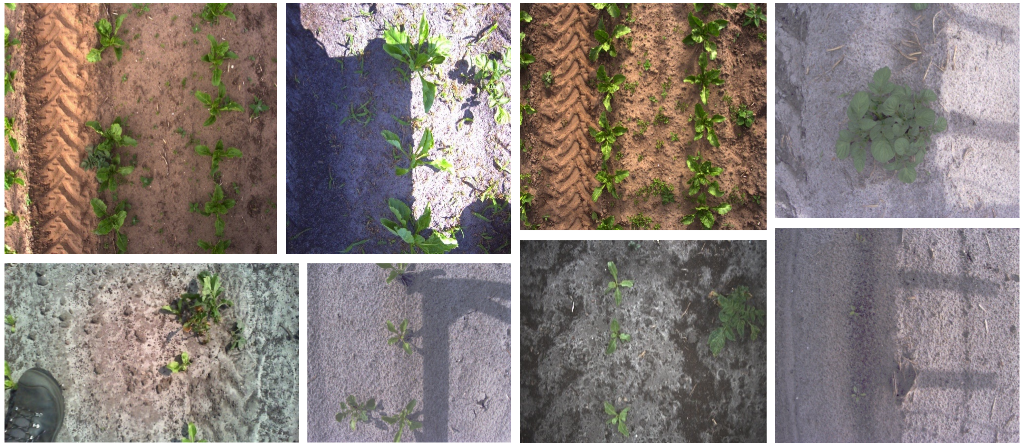
WAGENINGEN UNIVERSITY & RESEARCH

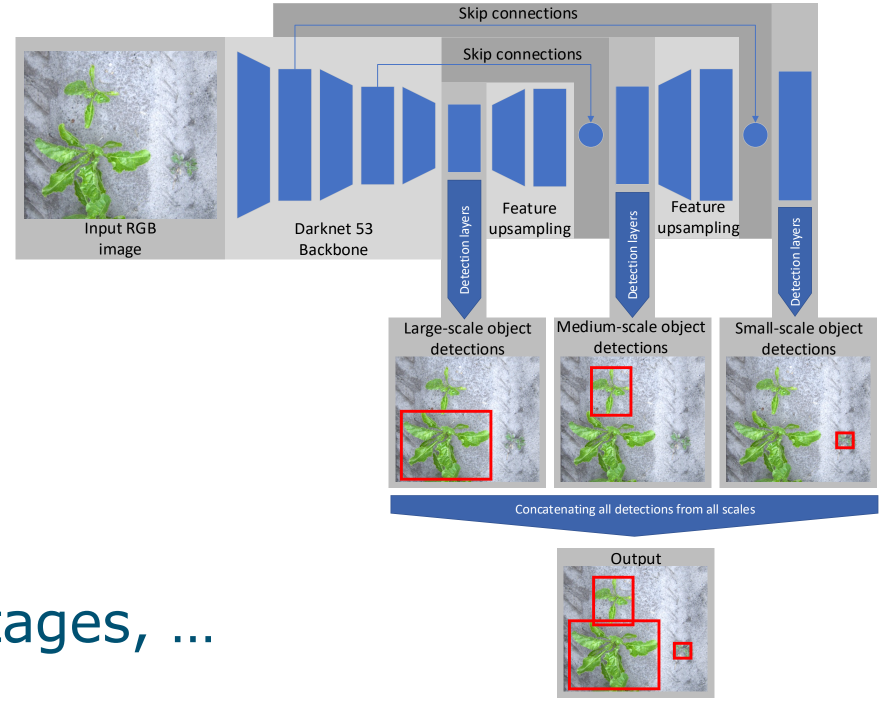
Challenges in 3D plant phenotyping

- Variation



Challenge of variation: Different sugarbeet fields



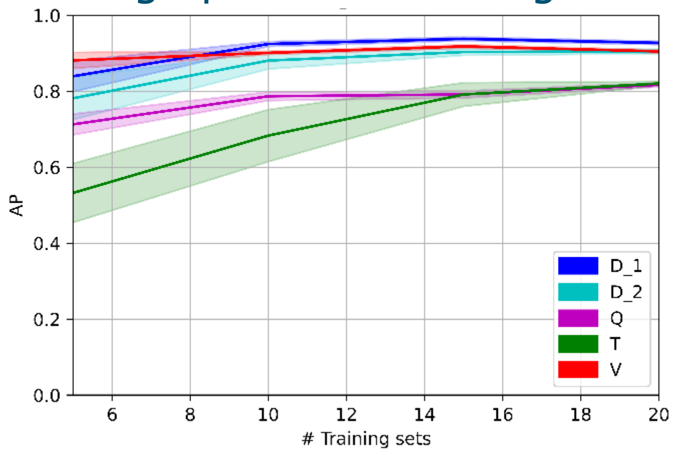


- Data from 22 different fields
- Different cultivars, soils type, growth stages, …

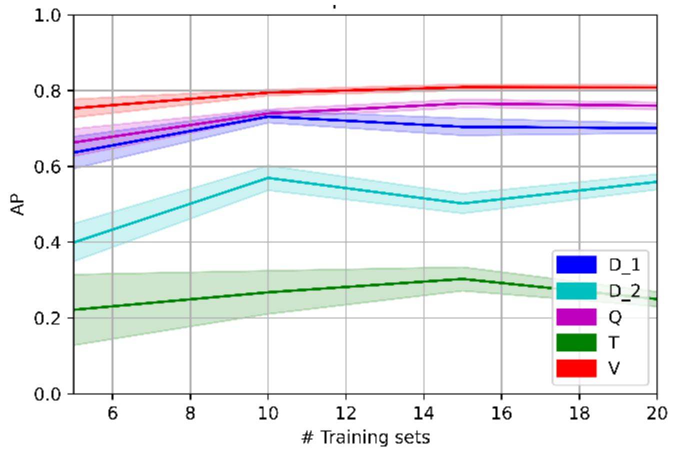
Challenge of variation: Different sugarbeet fields

- Generalization improves when the training sets contains more variation (more fields, not more data)

Average precision for sugarbeet



Average precision for potato



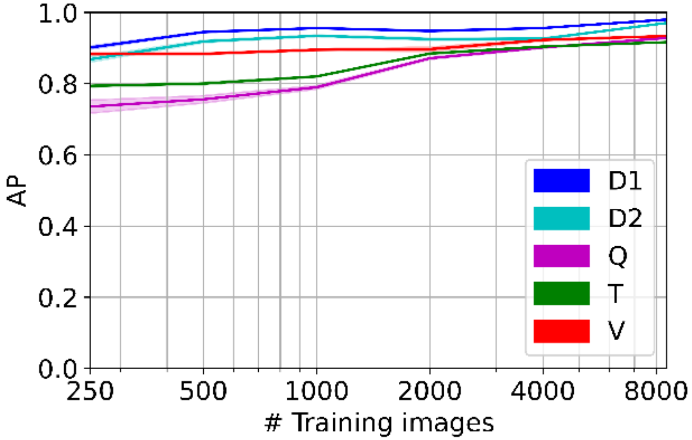
Training set consists of 500 images in all cases

WAGENINGEN UNIVERSITY & RESEARCH

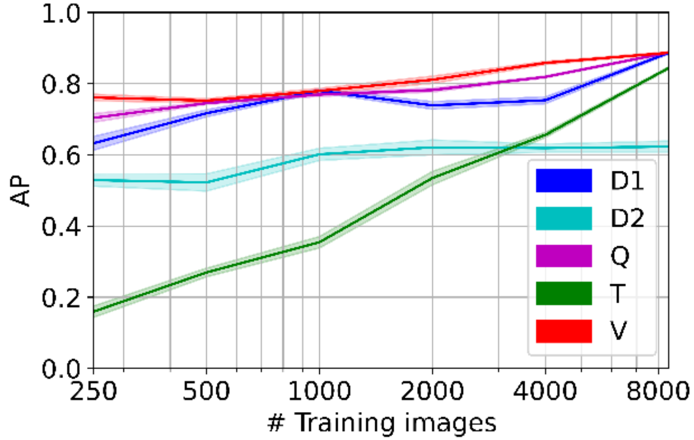
Effect of the size of the training set

- The performance further improves if more training samples are used

Average precision for sugarbeet



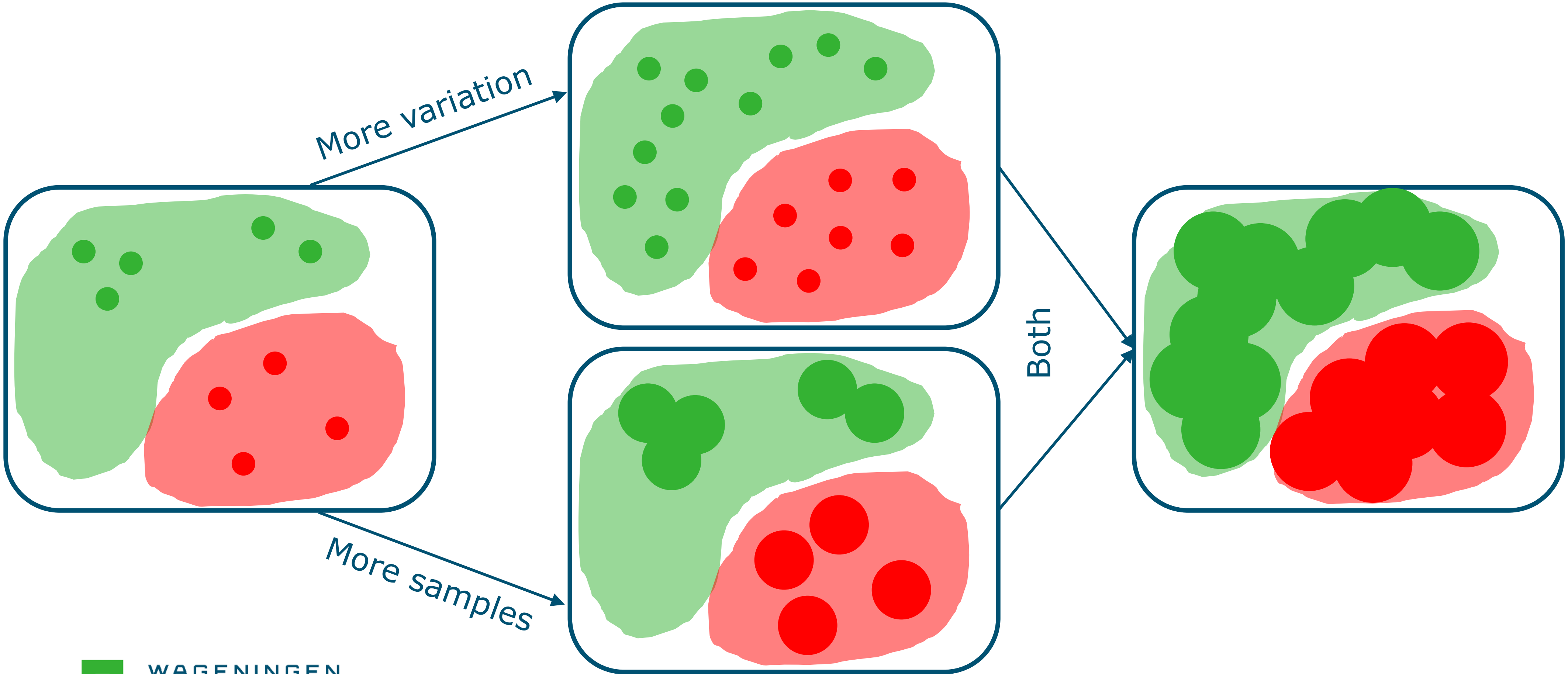
Average precision for potato



Weed detection: dealing with variation

Ruigrok, *et al.* (2020). Application-specific evaluation of a weed-detection algorithm for plant-specific spraying. *Sensors*, 20

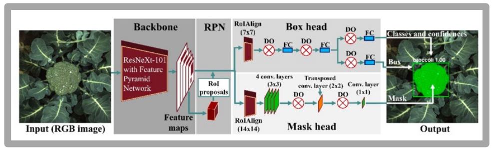
Increasing variation and size of the training set



How to select training data?
Active learning

Selecting new training data based on uncertainty

Pool of unlabelled images



Train model

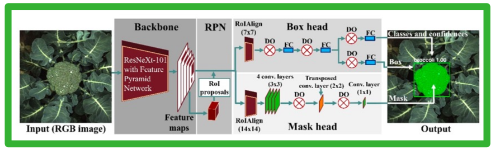
AL 1

Initial training data

Blok, P., Kootstra, G., ..., van Henten, E.J. (2022) Active learning with MaskAL reduces annotation effort for training Mask R-CNN. *Computers and Electronics in Agriculture*, 197: 106917

Selecting new training data based on uncertainty

Pool of unlabelled images



AL 2

Train model

Initial training data

AL 1

Blok, P., Kootstra, G., …, van Henten, E.J. (2022) Active learning with MaskAL reduces annotation effort for training Mask R-CNN. *Computers and Electronics in Agriculture*, 197: 106917

Selecting new training data based on uncertainty

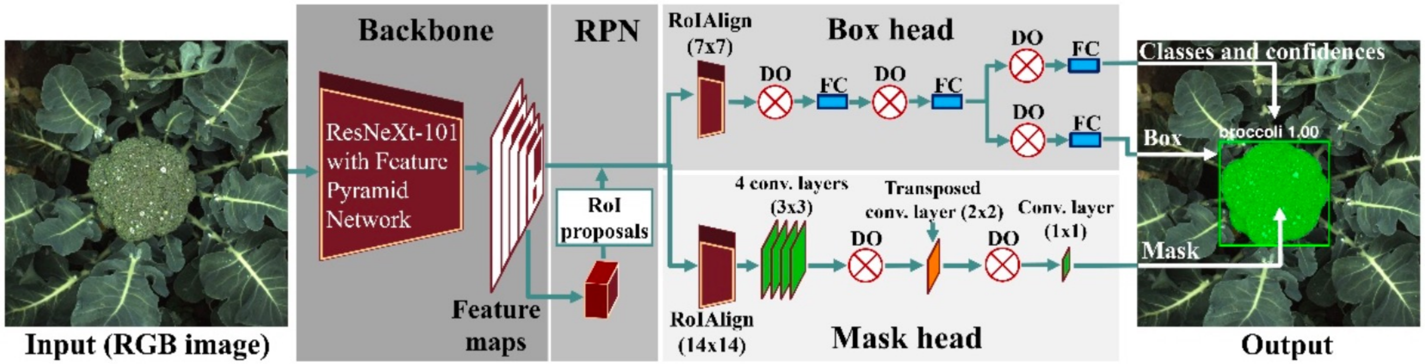
Pool of unlabelled images

Etc…

Initial training data | AL 1 | AL 2

Blok, P., Kootstra, G., …, van Henten, E.J. (2022) Active learning with MaskAL reduces annotation effort for training Mask R-CNN. *Computers and Electronics in Agriculture*, 197: 106917

Uncertainty estimation: Monte-Carlo dropout



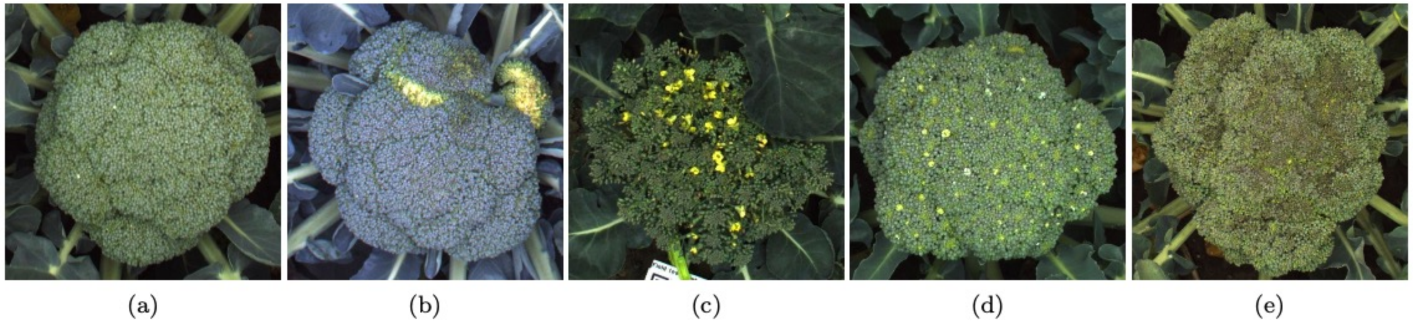
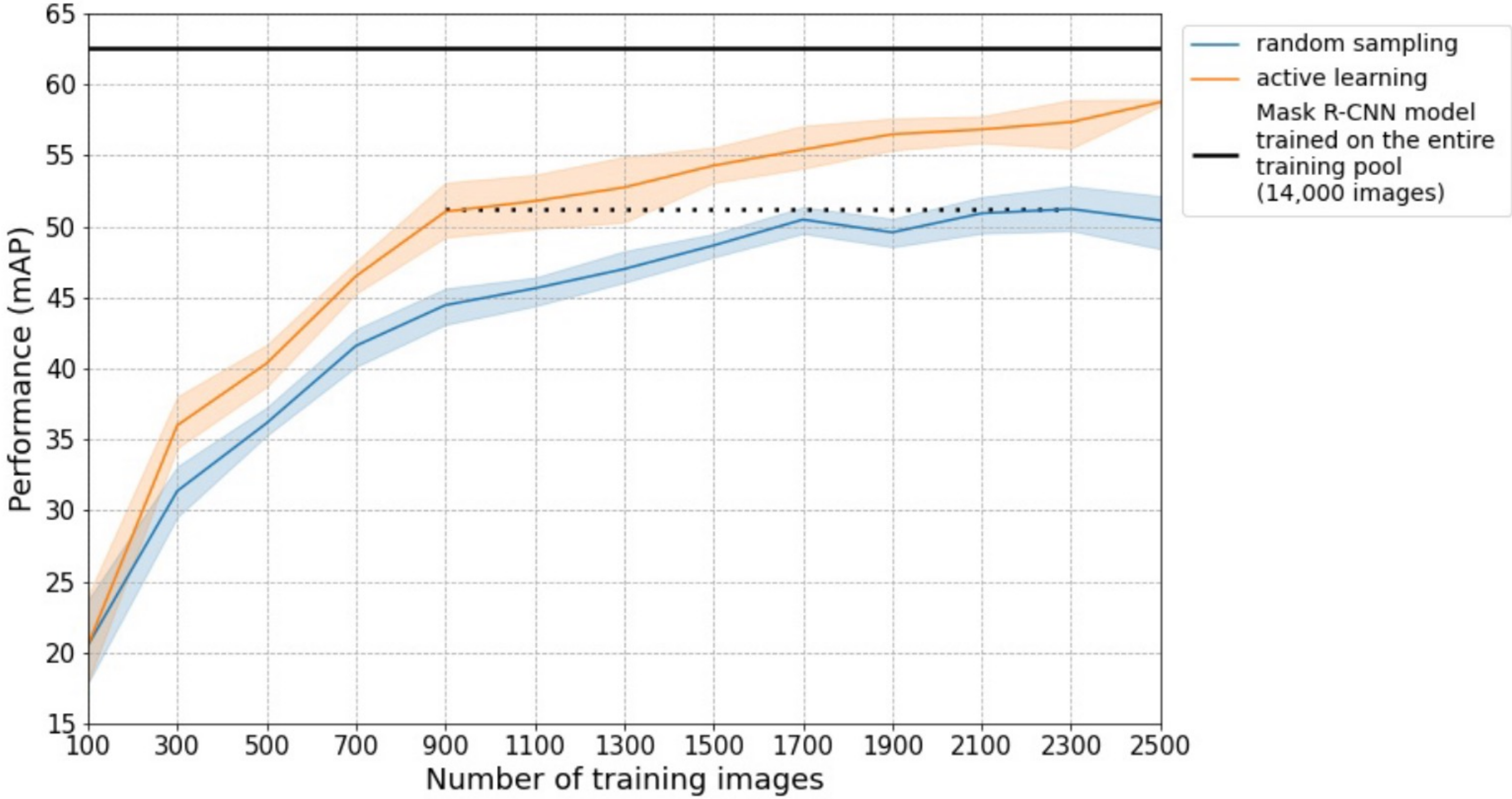


Figure 4: Examples of the five broccoli classes that were annotated in our data set: (a) healthy (b) damaged (c) matured (d) cateye (e) headrot. The displayed images were all cropped from a bigger field image.

Blok, P., Kootstra, G., …, van Henten, E.J. (2022) Active learning with MaskAL reduces annotation effort for training Mask R-CNN
Computers and Electronics in Agriculture, 197: 106917

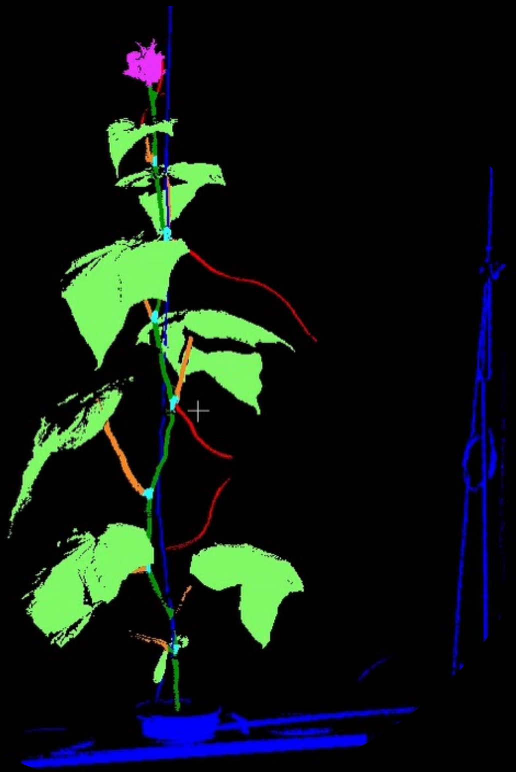
Results



Performance (mAP)
65
60
55
50
45
40
35
30
25
20
15
100
300
500
700
900
1100
1300
1500
1700
1900
2100
2300
2500
Number of training images
random sampling
active learning
Mask R-CNN model trained on the entire training pool (14,000 images)

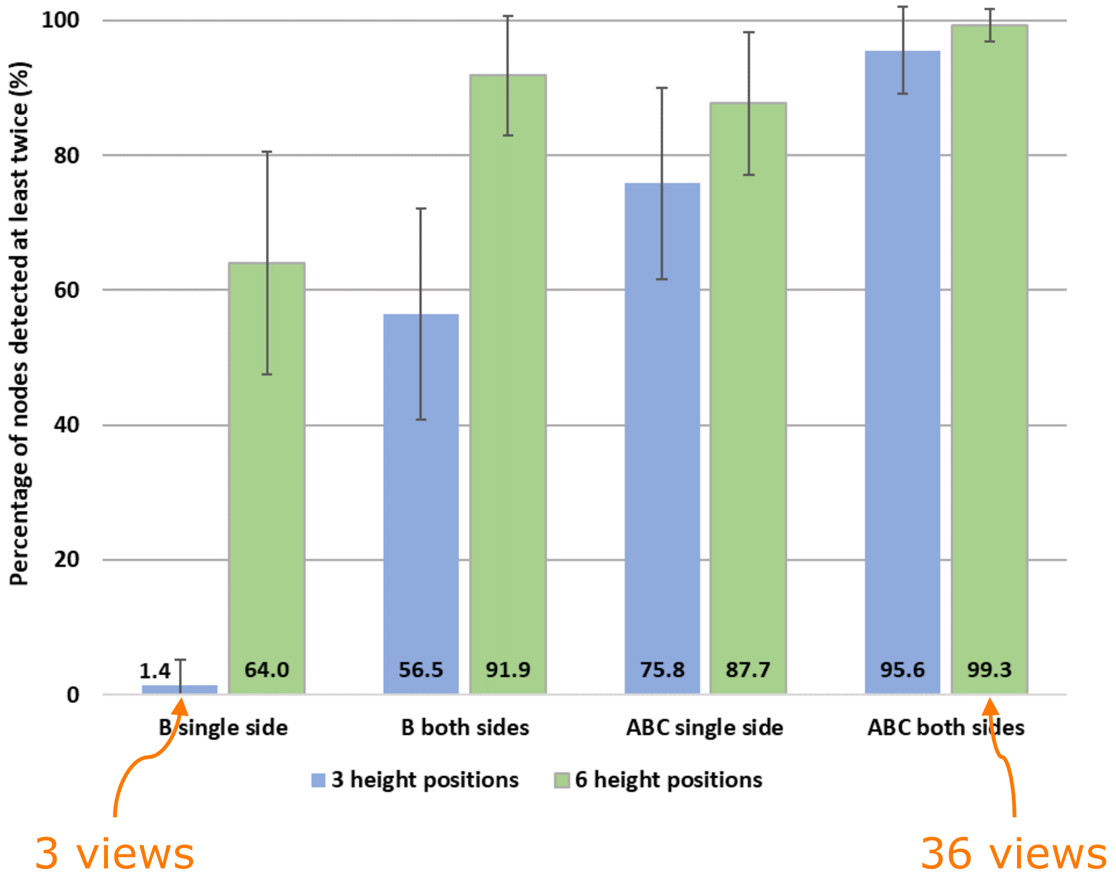
Challenges in 3D plant phenotyping

- Occlusion



Multi-view analysis

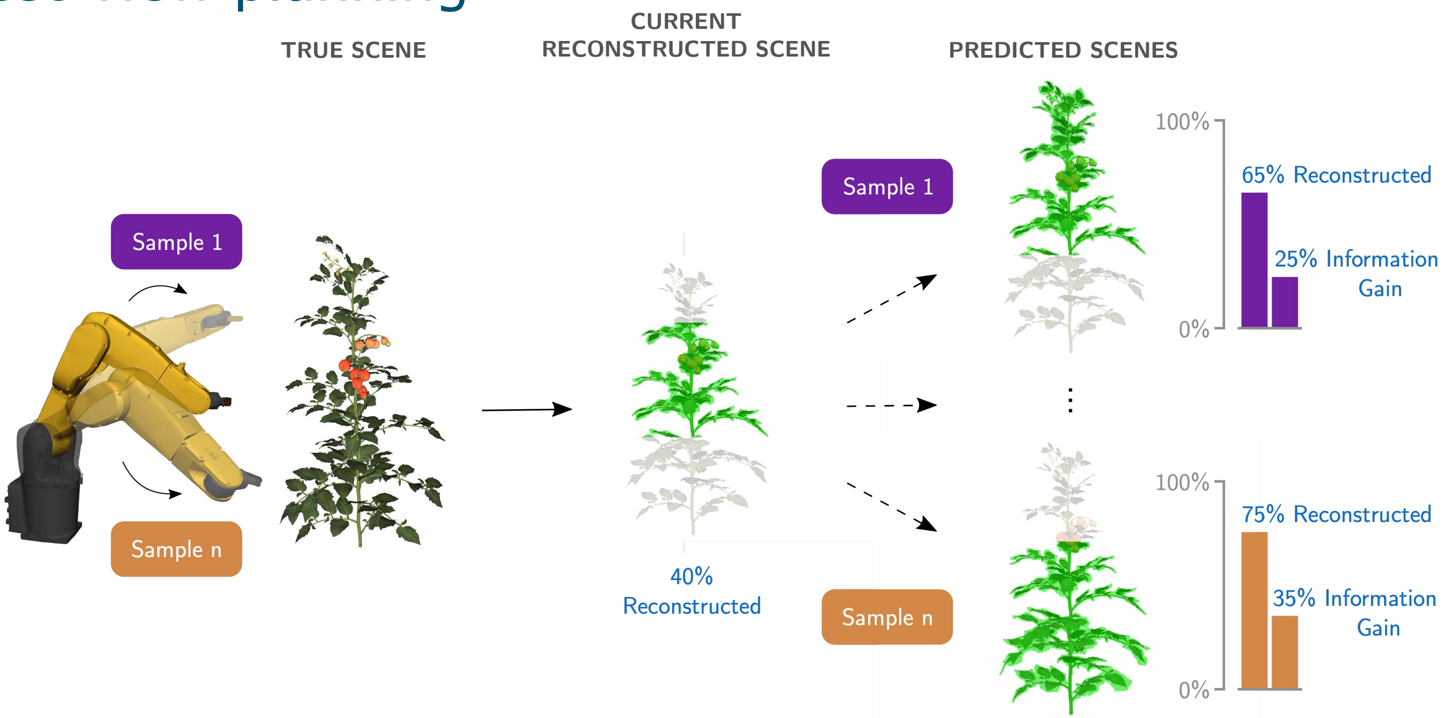
- We need many viewpoints to detect all nodes on a cucumber plant
- How to find the correct viewpoints?



WAGENINGEN UNIVERSITY & RESEARCH

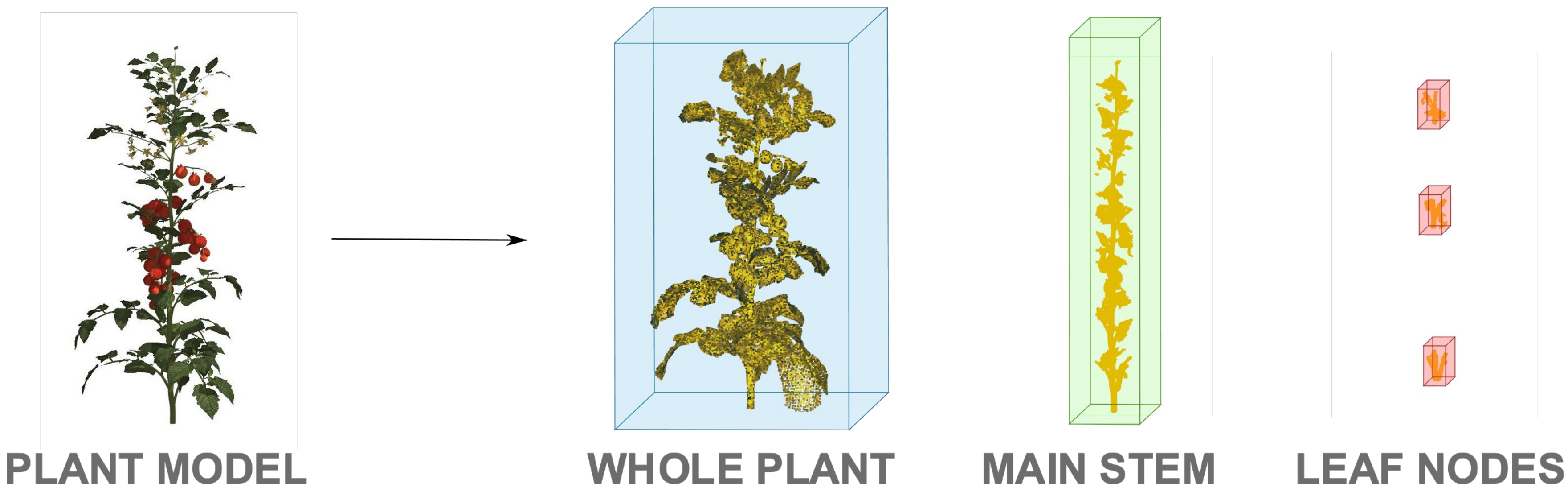
Robots need to reason about viewpoints

- Reasoning about seen and unseen space
- Next-best-view planning

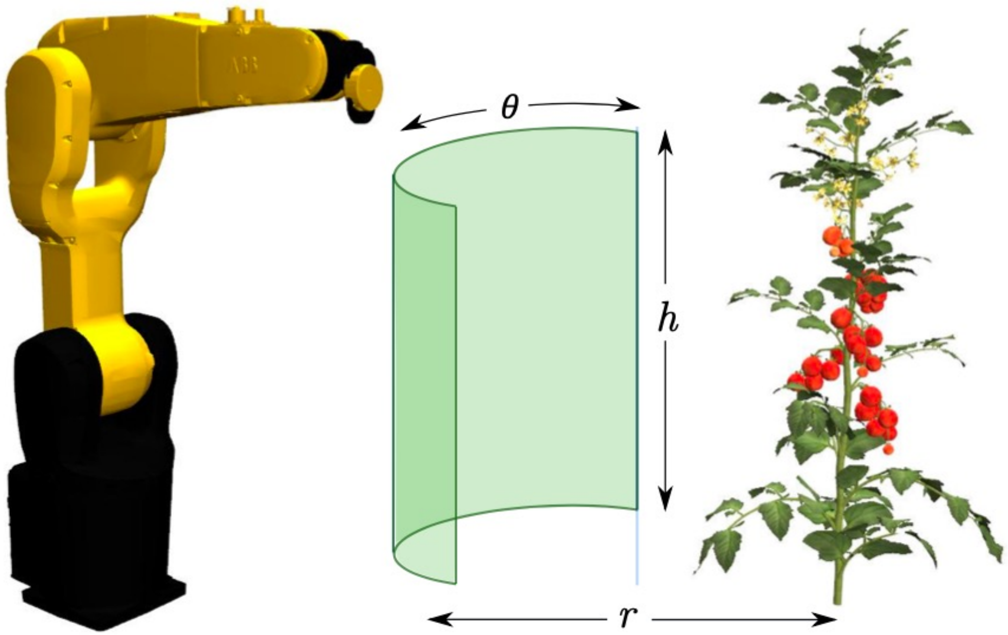


Burusa, A.K, van Henten, E.J. and Kootstra, G. (review) Attention-driven Active Perception for Efficient Reconstruction of Plants and Targeted Plant Parts. Computers and Electronics in Agriculture.

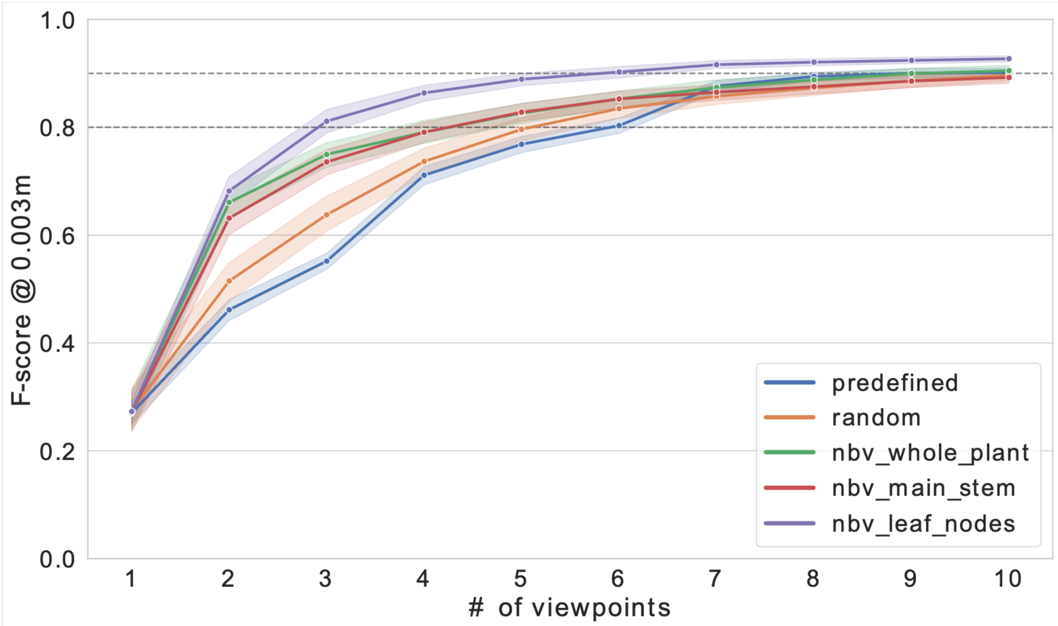
Attention mechanism



Methods

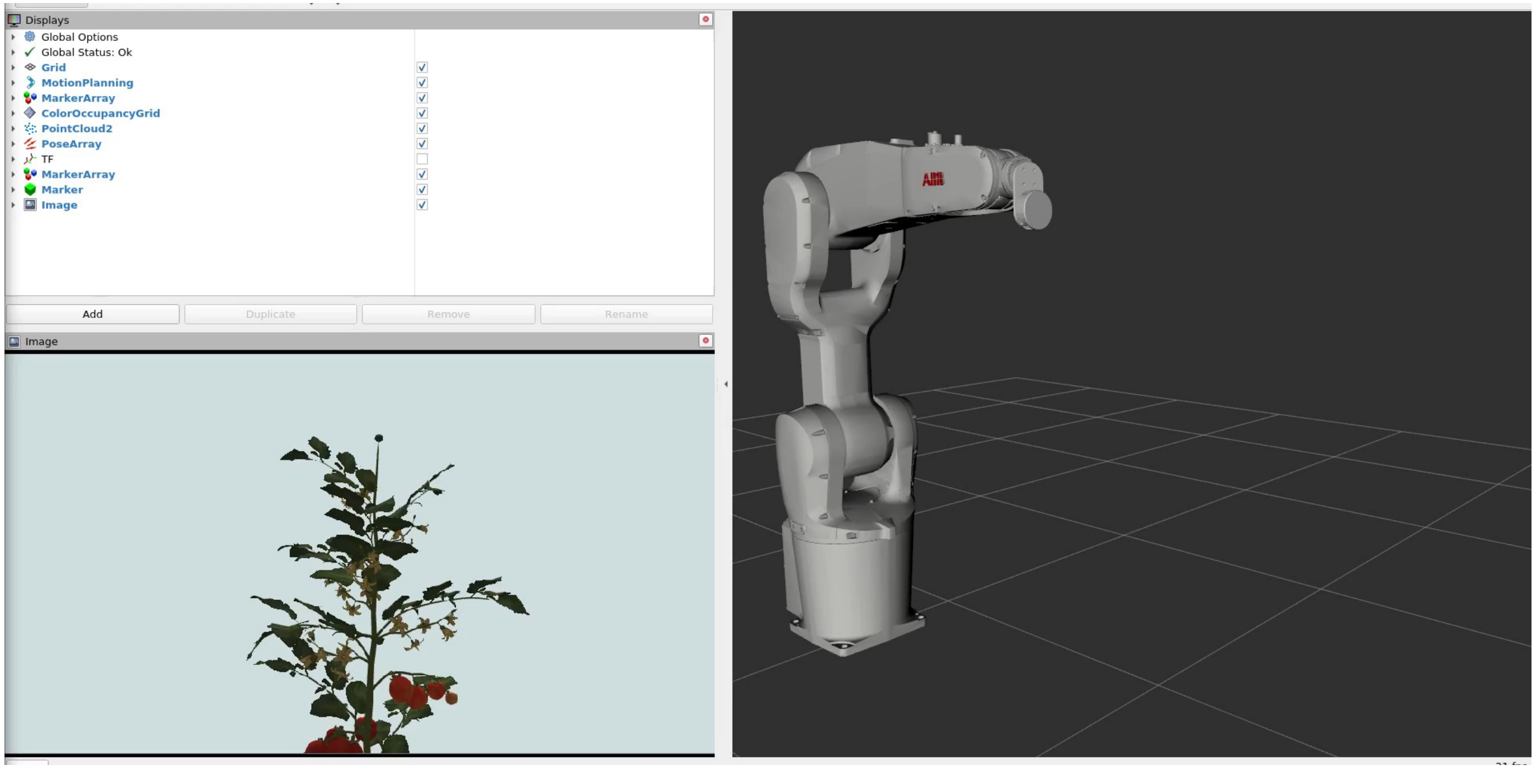


Results



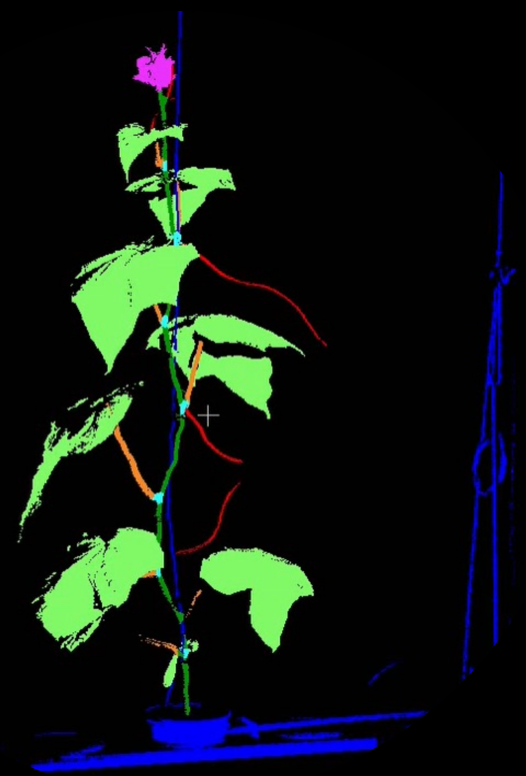
Task: leaf node reconstruction

Example



Burusa, A.K, van Henten, E.J. and Kootstra, G. (review) Attention-driven Active Perception for Efficient Reconstruction of Plants and Targeted Plant Parts. Computers and Electronics in Agriculture.

Conclusion

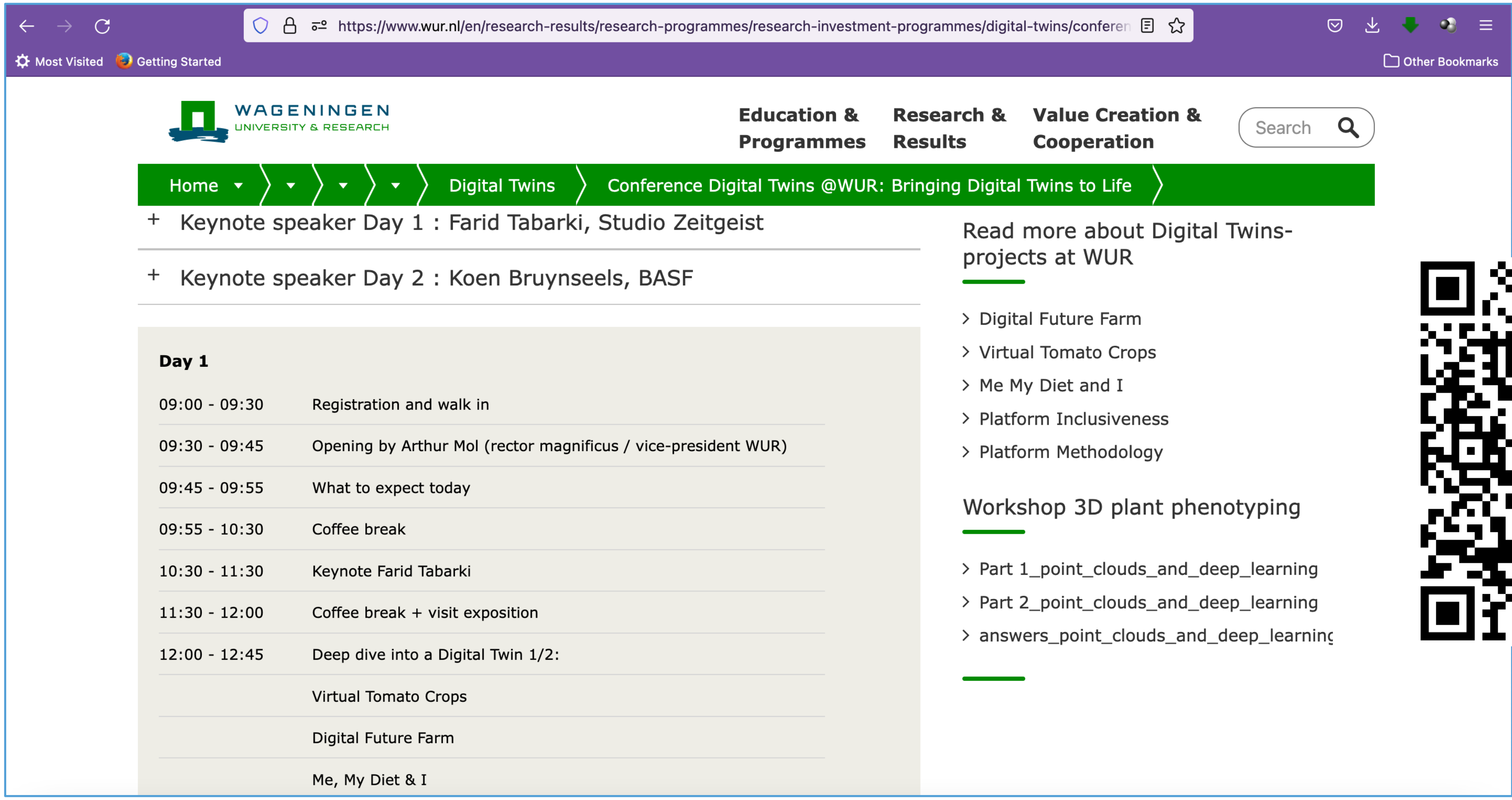


Conclusion

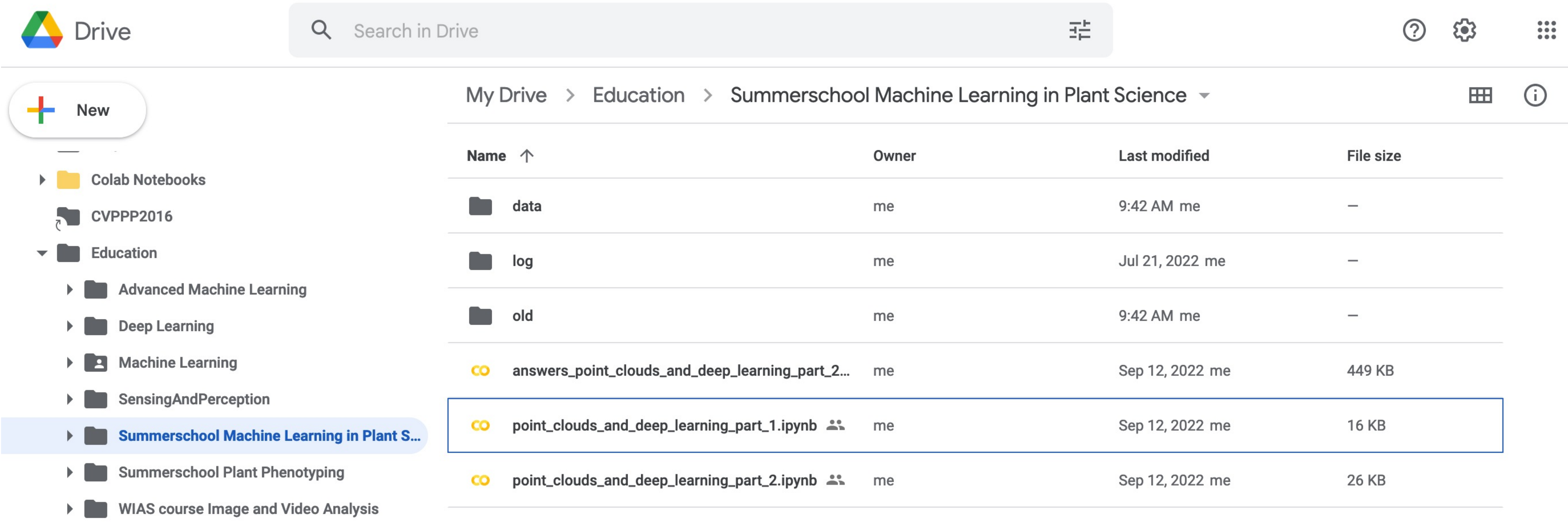
- Many different 3D acquisition systems exist
- Segmentation and detection of plant parts is better in 3D
- PointNet(++) can learn to segment 3D point clouds
- Adding spectral information to the spatial information improves segmentation
- The training set needs to of good quality (variation, class imbalance, data augmentation)
- We need active scanning to deal with occlusions


Hands-on tutorial

Download notebooks



Google Drive – Google Colaboratory



Two notebooks

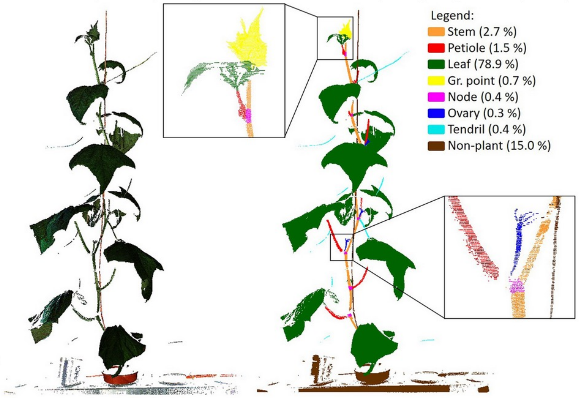
point_clouds_and_deep_learning_part_1.ipynb
File Edit View Insert Runtime Tools Help Last edited on September 12

+ Code + Text

3D plant phynotyping

In previous tutorials, you have seen examples of computer-vision methods to process 2D colored images. In the past years, many powerful methods became available to get relevant information out of the images. In the field of plant science, these have been applied, for instance, to detect plants in a field (Ruigrok *et al*, 2020), to segment individual leaves (Shi *et al*, 2019), and to detect nodes on a plant (Boogaard *et al*, 2020).

However, as pointed out in (Boogaard *et al*, 2020), 2D data is very limiting if one needs to estimate geometrical properties, such as lengths and surface areas, as they need to be estimated from the 2D projection of the 3D world onto the image. Instead, if 3D point clouds are used, such measurements can be done much more accurately.



The field of 3D point-cloud processing, especially using deep neural networks, is not yet as mature as that for 2D image processing. A lot of progress is expected in the coming years and already now, there are some good examples of using 3D point clouds for plant phenotyping, e.g., (Shi *et al*, 2019; Golbach *et al*, 2016; Boogaard *et al*, 2022).

In this tutorial, you will make some first steps exploring the use of 3D point-cloud data.

Learning goals
After successful completetion of this tutorial, you will be able to:

- mention the advantages of 3D imaging data over 2D imaging data,
- describe the concept of several sensor systems to acquire 3D data,
- apply some basic 3D point-cloud processing methods,
- explain the functioning of a 3D neural network conceptually,
- apply PointNet, a 3D neural network, to segment a point cloud of tomato ,seedlings in leaf and stem, and

point_clouds_and_deep_learning_part_2.ipynb
File Edit View Insert Runtime Tools Help Last edited on September 12

+ Code + Text

Point clouds and deep learning

In this part of the tutorial, you will learn how to train a deep neural network to learn to segment plants in different plant parts.

After finishing this part of the tutorial, you will be able to:

- explain how training data for 3D semantic segmentation looks like
- use the PointNet deep neural network
- compare the performance of a model trained for only a few epochs to one trained for many epochs

1: Getting started

Run the blocks below to load the required libraries and get the data

[ ]
```
import numpy as np
import matplotlib.pyplot as plt
import plotly.graph_objects as go
from plotly.subplots import make_subplots
import plotly.express as px
import pandas as pd
import os
import pickle
from pathlib import Path
import random
import numpy as np
import h5py

# TensorFlow is a library to implement and run neural networks
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers

```

WAGENINGEN
UNIVERSITY & RESEARCH

See in 3D!

gert.kootstra@wur.nl

bolai.xin@wur.nl

WAGENINGEN UNIVERSITY & RESEARCH