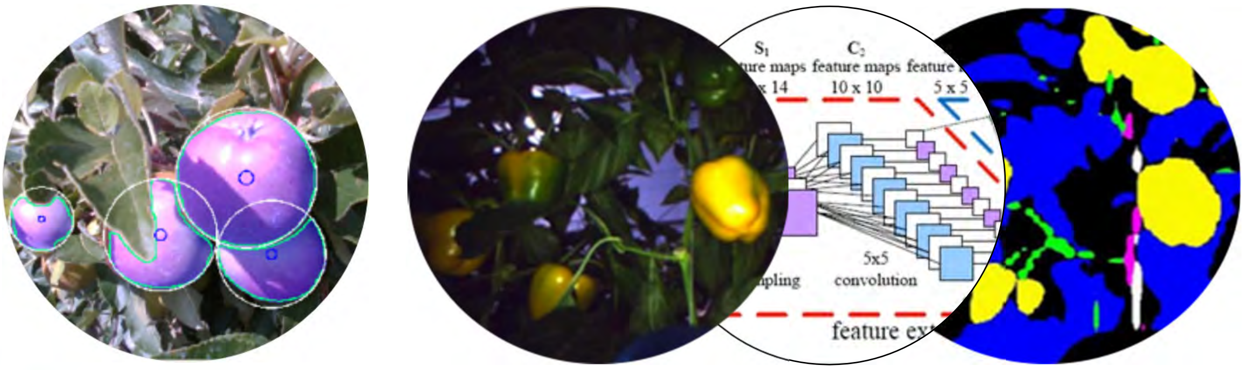
Deep learning and its application in horticulture and agriculture

Jochen Hemming
Wageningen University & Research, The Netherlands



WAGENINGEN UNIVERSITY & RESEARCH

RECPAD2022, October 28, 2022

1

Presentation outline

- Introduction Wageningen University & Research (WUR)
- Application of deep learning in horticulture and agriculture
  - Fruit, branch and obstacle detection for robotic harvesting
  - Robotic harvesting of gerbera flowers
  - Potato yield measurement
  - Selective broccoli harvesting (amodal masks)
  - Active learning
  - Insect and disease detection, determination and counting
  - Fish sorting
- Final Conclusions

WAGENINGEN UNIVERSITY & RESEARCH

2

Wageningen University and Research (WUR)

- A university and research centre dedicated to life science, based in the Netherlands.
  - Wageningen University
  - Wageningen Research
- WUR:
  - 6400 employees
  - 13000 BSc and MSc students
  - 2300 PhD students
  - From >100 countries

WAGENINGEN UNIVERSITY & RESEARCH

3

3

WUR Agro Food Robotics

- Joint program by research groups of Wageningen University & Research
- > 40 people working in the field of Computer Vision, AI and robotics for agri&food within the different science groups of WUR.
- Engineers and researchers work together with industrial partners on new robotic systems for agri and food.
- Concept development and functional design, prototype development, testing and validation, support for new product implementation.
- We specialise in artificial intelligence and sensing, especially spectral, machine learning and vision.

Why collaborate? These are your benefits:
- Access to innovative technologies
- Optimal R&D brainpower
- No need to start from scratch
- Be a frontrunner
- Networking / consortium Opportunities

www.wur.eu/agrofoodrobotics

4

General intro

- Agri-food domain: a need to produce more with less.
- Automation can help – machine vision is part of the automation.
- Natural environment is characterized by the high variability of the objects and environment and large changes over time.
- Deep learning methods have the capability to deal with this variablilty better than classical image analysis methods.
- Our application domain profits from the big players like Google and Facebook that are pushing the development on relevant topics (autonomous navigation, big-data and AI).

WAGENINGEN UNIVERSITY & RESEARCH

5

5

Sweeper **Swee**t-**pep**per **r**obot (2015-2018)

www.sweeper-robot.eu

- Horizon 2020 ICT use case project
- 6 partners from 4 countries (The Netherlands, Belgium, Sweden and Israel).

WAGENINGEN UNIVERSITY & RESEARCH

אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev

UMEÅ UNIVERSITY

PROEFSTATION

bogaerts greenhouse logistics

de tuindershoek

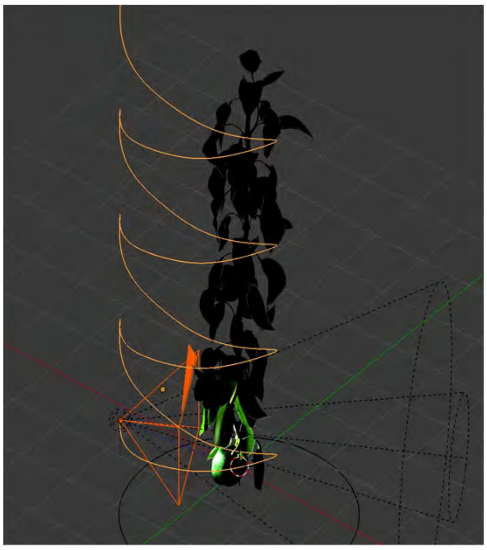
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No **644313**.

6

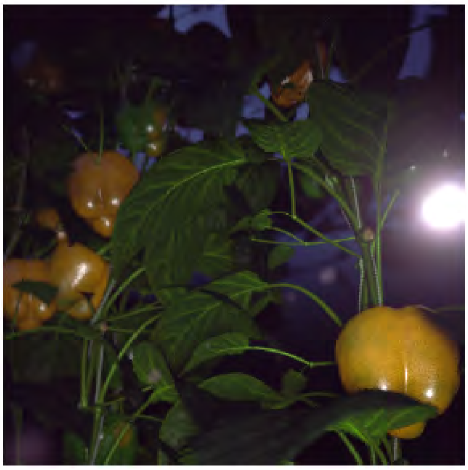
7

Synthetic bootstrapping of CNNs for semantic plant part segmentation

Current bottleneck is the requirement of large annotated datasets to train deep-learning networks.



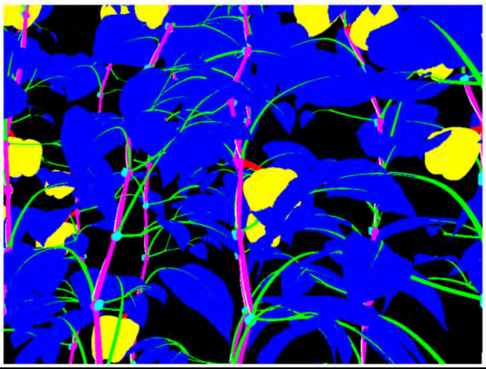
Synthetic



8

Deep-Learning for plant part localization in images

- Semantic segmentation with DeepLab V2 (per-pixel, no instance detection).
- Synthetic dataset is used to bootstrap the model.
- Trained network deployed for real-time obstacle detection and to determine best end-effector alignment.



WAGENINGEN UNIVERSITY & RESEARCH

9

Optimising Realism of Synthetic Images using Cycle-GANs (Generative Adversarial Networks)

- Dissimilarity gap remains caused by sub-optimal manual modelling
- Optimising the realism of synthetic images by unpaired image-to-image translation from the synthetic to empirical domain

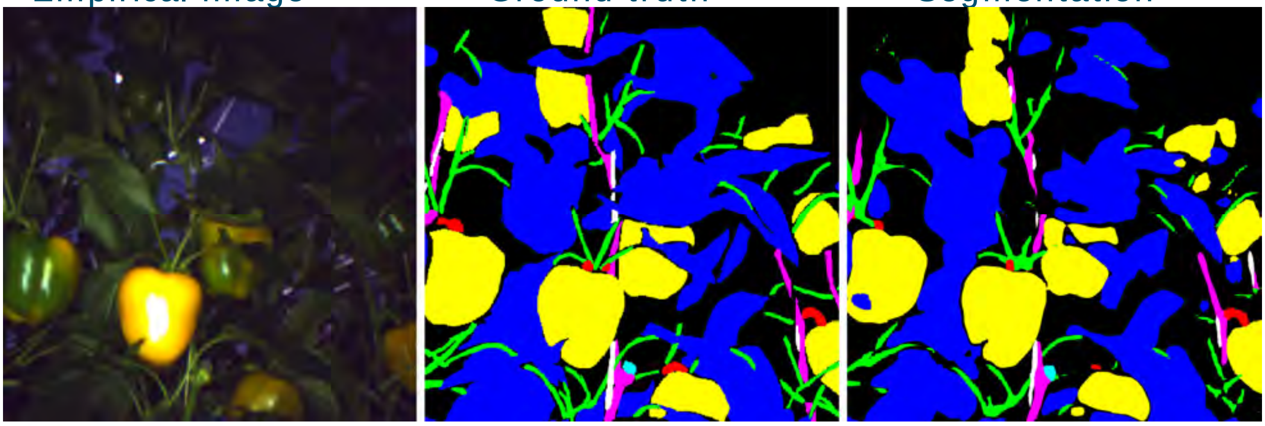
Barth, R. ; IJsselmuiden, J.M.M. ; Hemming, J. ; Henten, E.J. van (2017): Optimising Realism of Synthetic Agricultural Images using Cycle Generative Adversarial Networks. In: Proceedings of the IEEE IROS workshop on Agricultural Robotics 2017.

10

Results on empirical images

- Objective: plant main stem detection to calculate obstacle free approach direction for the robot

Empirical image | Ground truth | Segmentation



Barth, R., Hemming, J., Henten, E.J. van (2019). Angle estimation between plant parts for grasp optimisation in harvest robots. Biosystems Eng., 183, 26-46.

11

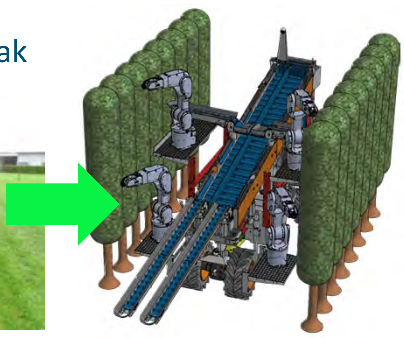
Apple harvesting robot

- Exchange human workers for robotic units
- Increase robotic picking rate step by step
- Human workers and a robots working together

MUNCKHOF Fruit Tech Innovators

RIWO

Munckhof Pluk-O-Trak



Orchard experiments 2020

12

Apple detection and localization using DL

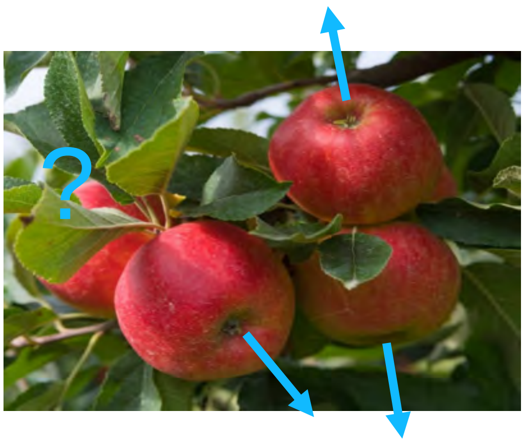
- RGB-Depth camera (Intel RealSense D435) mounted on platform.
- Real-time instance segmentation on RGB images with YOLACT++ [later switched to YolactEdge (ResNet-101 backbone)].
- Individual fruits are detected, even when fruit are touching and overlapping each other.
- Post-processing using depth channel, including point cloud sphere fitting.

WAGENINGEN UNIVERSITY & RESEARCH

13

Full pose estimation

- Low harvest succes rate due to apples hanging in all kinds of directions.
- Calyx detection model (YOLOv3 => v5)
    - Trained on 938 apple images
    - Input layer dimensions 416x416 pixel
    - Precision 0.84/recall 0.96
- Trained model converted to ONNX (Open Neural Network Exchange https://onnx.ai/) format for easier integration in robot.



15

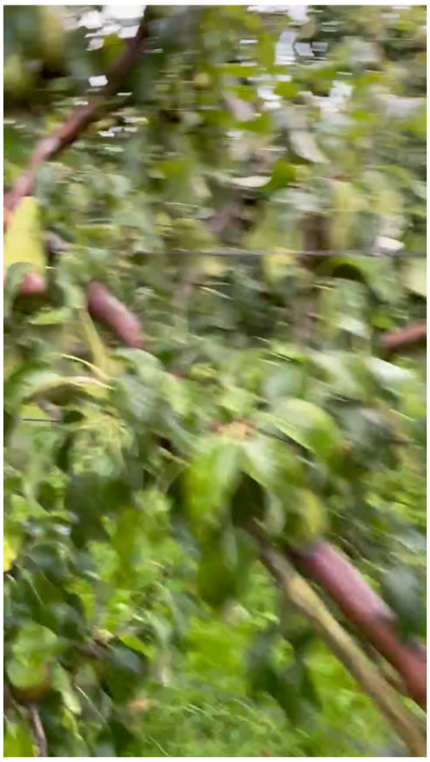
15

Orchard experiments September 2021
(source: https://www.npostart.nl/VPWON_1324351)

16

Pear harvesting robot

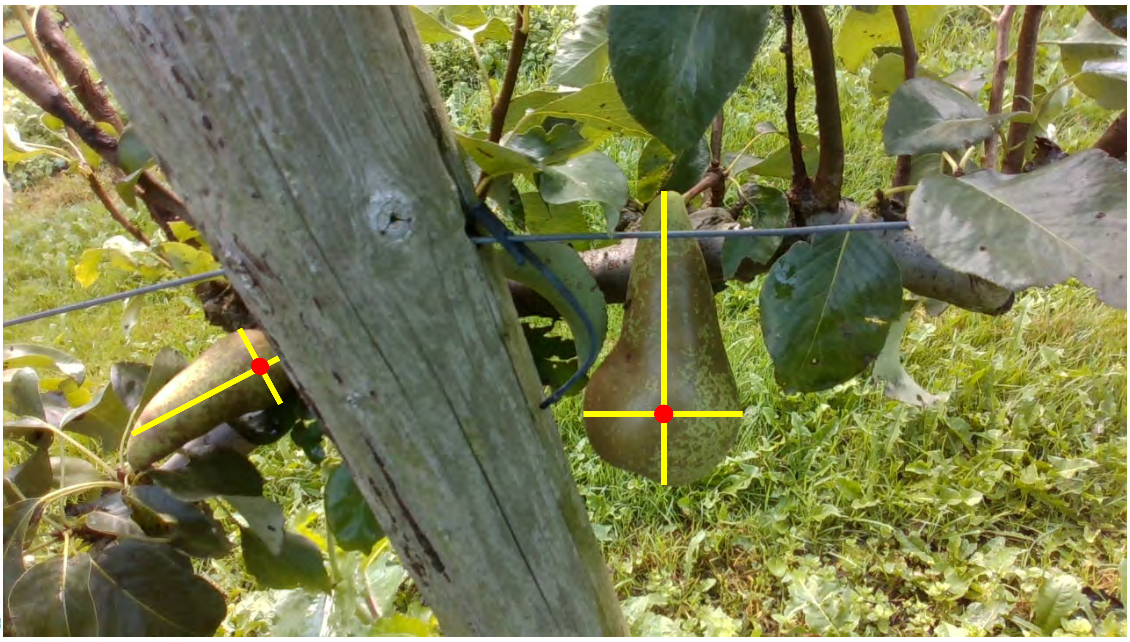
- Pears need to be harvested with their stem attached
- Detect individual pears and their stems
- Detect the grasping point



17

17

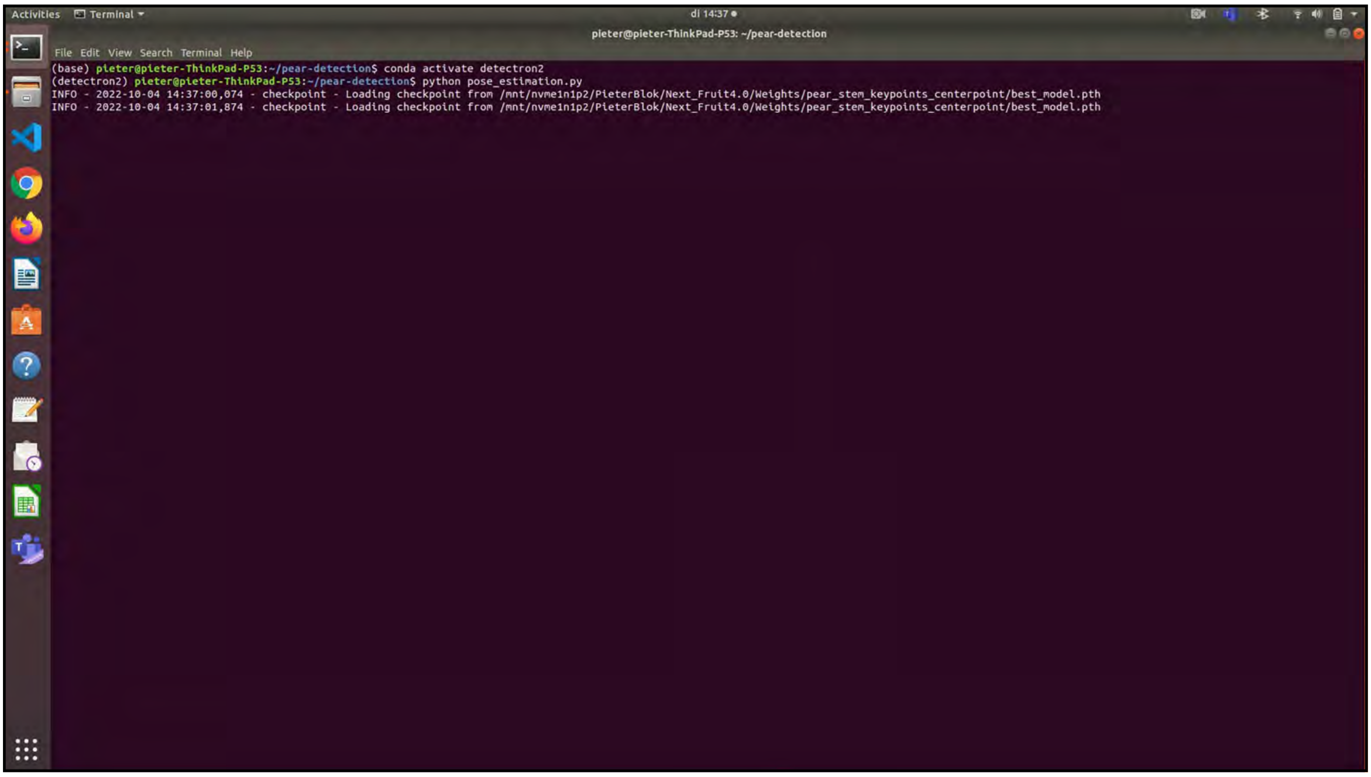
Keypoint R-CNN



Training and testing

- Test Keypoint R-CNN in its ability to:
  - Detect and segment each pear (instance segmentation)
  - Detect the grasping point (keypoint detection)
- 2778 images (RGB-Depth) were acquired:
  - 94 images were selected
  - 64 images for training
  - 15 images for validation (during training)
  - 15 images for testing (after training)

WAGENINGEN UNIVERSITY & RESEARCH



20

Results

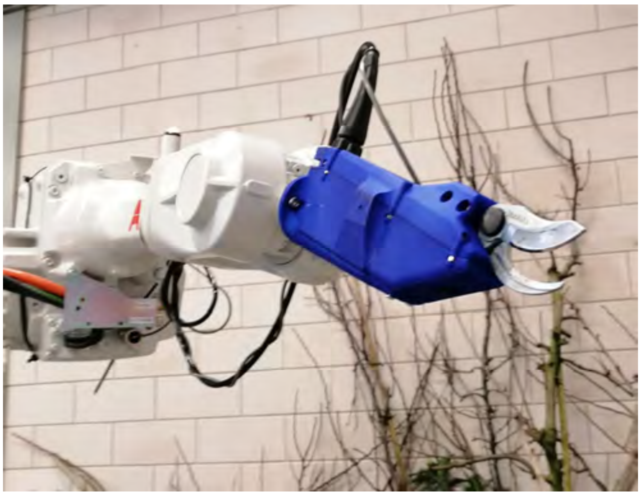
- The images in the test set contained 45 (annotated) pears
- Evaluation metric: 3D Euclidean distance between the annotated grasping point and the predicted grasping point

Metric	Value (mm)
Average Euclidean distance (3D)	**4.7**
Minimum Euclidean distance (3D)	1.2
Maximum Euclidean distance (3D)	21.1
Median Euclidean distance (3D)	3.4

WAGENINGEN UNIVERSITY & RESEARCH 21

21

Pruning robot for soft fruit bushes



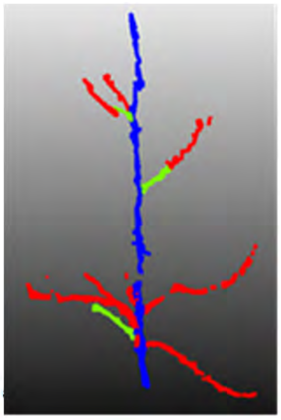
22

Sensor system for 3D recordings

23

Coloured point cloud (RGB-D) analysis

- Manual annotation of 160 point clouds.
- Model trained with PointNet deep learning network

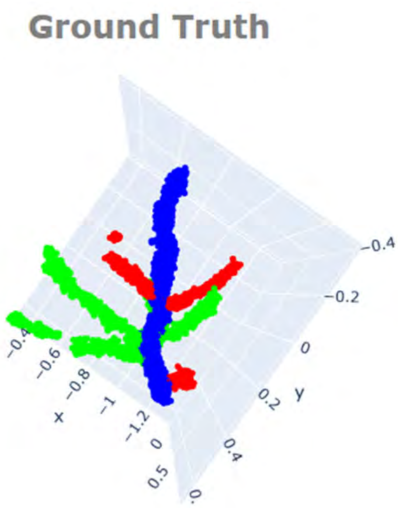


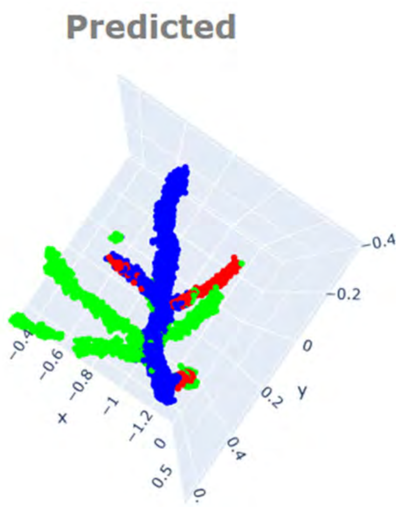
3 classes
- Trunk (blue)
- 1-y branch (red)
- 2-y branch (green)

24

Preliminary results (work in progress)

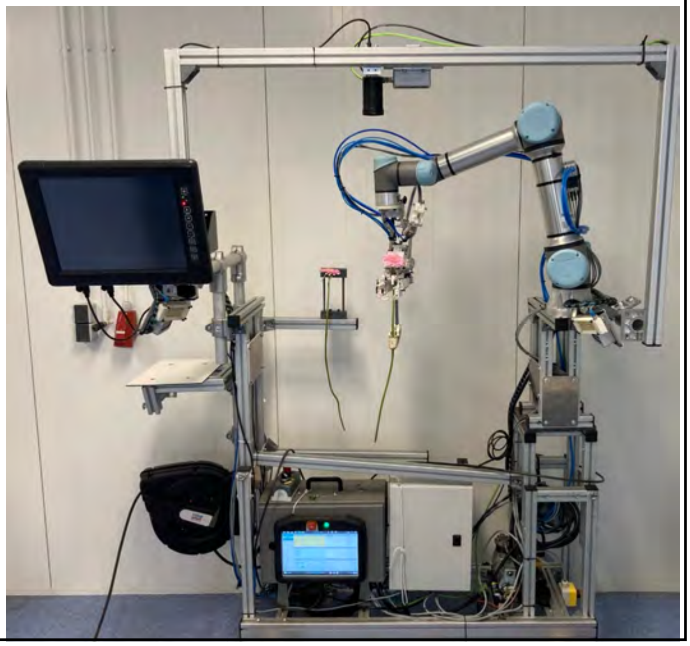
- Noisy data, many gaps
- 75% correct classification





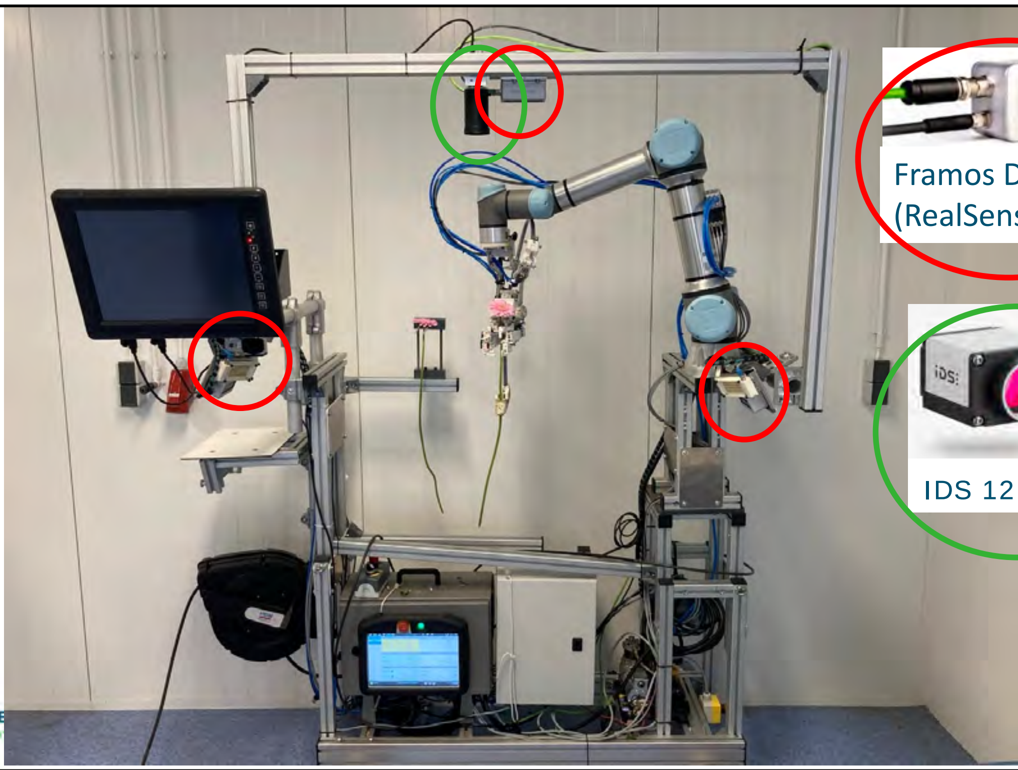
25

Robotic harvesting of Gerbera



WAGENINGEN
UNIVERSITY & RESEARCH

26



Framos D415e
(RealSense D415)

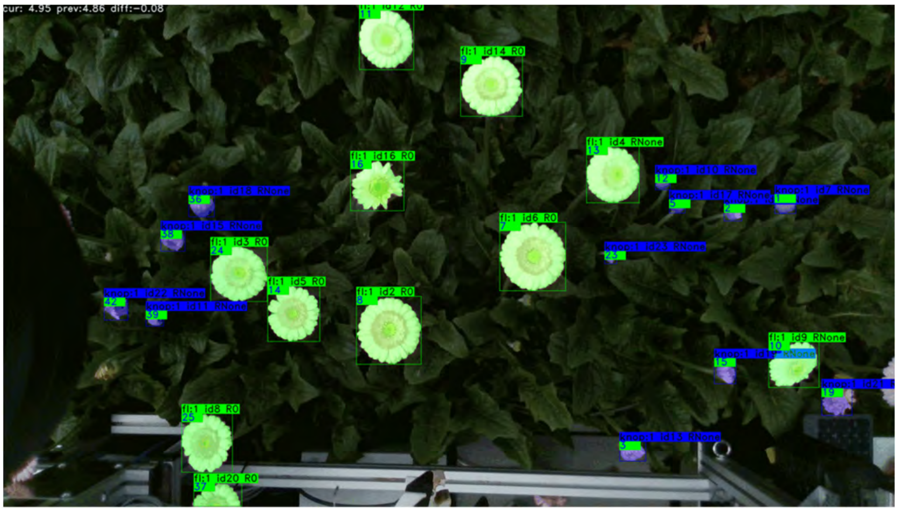
IDS 12 MP RGB

27

27

Ripe flower detection (top view camera)

- Object detection using Mask-RCNN (flowers and buds)
- Tracking over multiple images



WAGENINGEN UNIVERSITY & RESEARCH

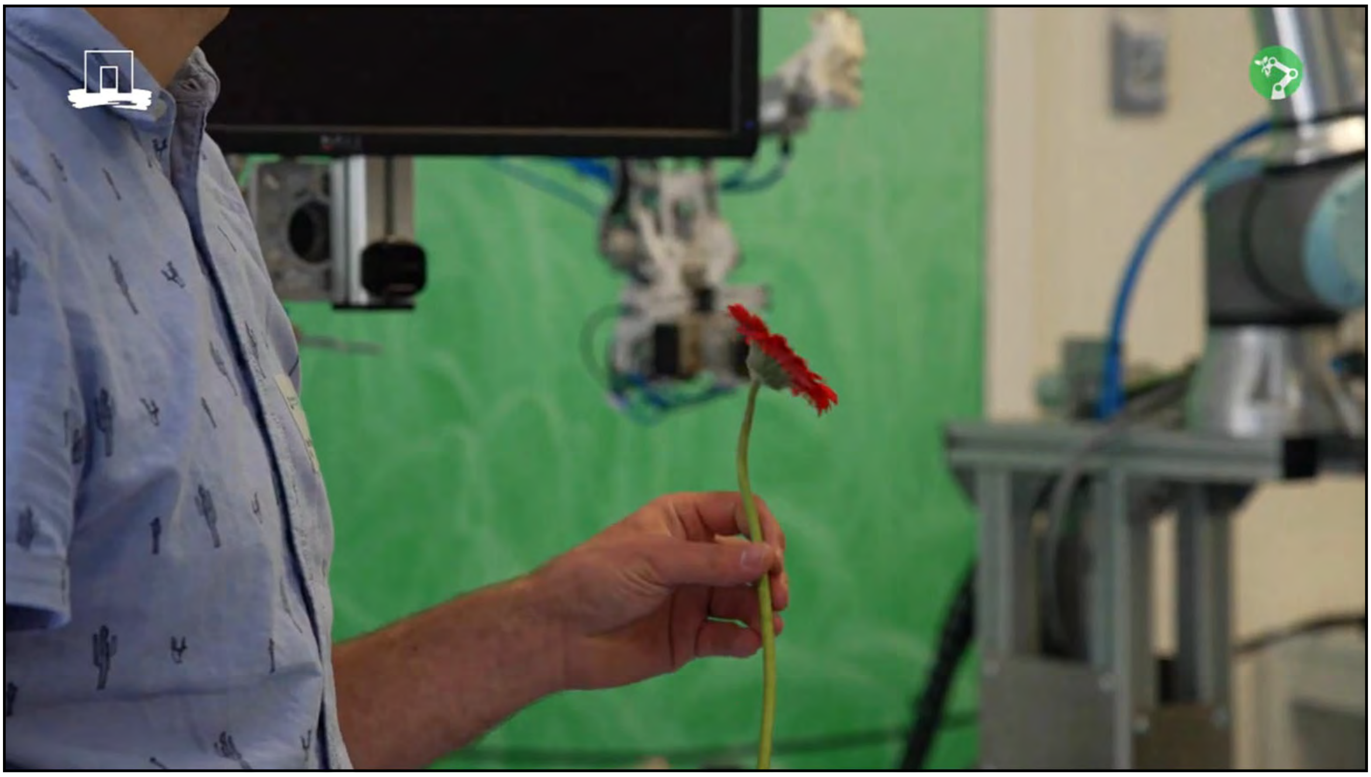
28

Stem recognition

- Two instance segmenation networks tested: YOLACT and Mask-RCNN
- Only Mask-RCNN was able to detect long vertical objects
- Match stems and flower
- Post-processing: convert masks into 3d cylinders
- Calculate gripping point
- Calculate best approach for arm

WAGENINGEN UNIVERSITY & RESEARCH

29



30

Yield measurement during the harvest of seed potatoes

- Final goal: yield potato per m2 soil.
  - Gross yield (weight)
  - Net yield (weight) → tare estimation
  - Size distribution potatoes

WAGENINGEN UNIVERSITY & RESEARCH

31

31

Image acquisition module

WAGENINGEN
UNIVERSITY & RESEARCH

32

32

33

Image analysis

- Mask-RCNN for instance segmentation
- Training dataset
  - 250 labelled images
  - 8000 potatoes
  - 1070 clods
  - 4 potato varieties

WAGENINGEN UNIVERSITY & RESEARCH

34

34

Example of instance segmentation result

35

35

Better size estimation of occluded crops

- Goal: better estimate the size of crops that are occluded (partly visible)
- Humans can rely on the *amodal perception*, which is the ability to reason about the invisible and occluded parts of objects.

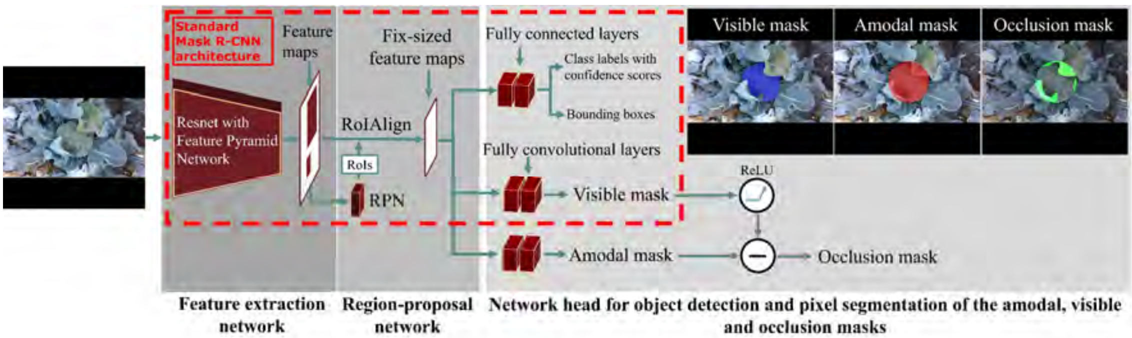
WAGENINGEN UNIVERSITY & RESEARCH

36

36

Occlussion RCNN + deep-learning regression

- We developed an algorithm, called Occlusion R-CNN (ORCNN) to better estimate the size of crops that are occluded (partly visible)
- ORCNN is an extension of Mask R-CNN, that can estimate three masks instead of one!

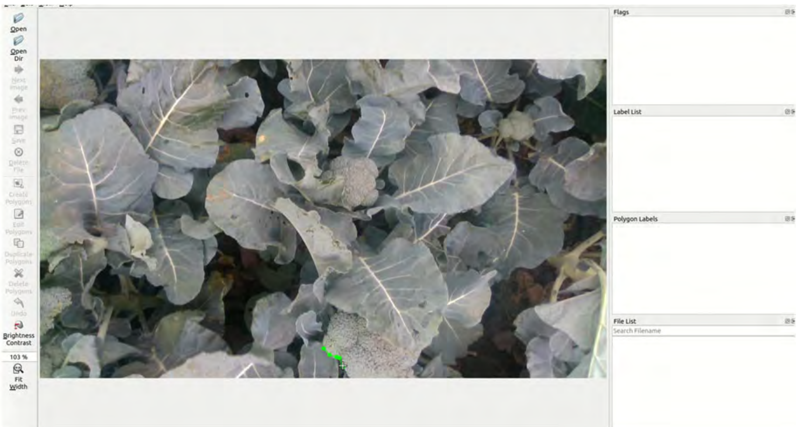


Pieter M. Blok, Eldert J. van Henten, Frits K. van Evert, Gert Kootstra (2021): Image-based size estimation of broccoli heads under varying degrees of occlusion, Biosystems Engineering, 208, 213-233.

37

Active learning

- To properly optimise a DL network there are usually a lot of labelled images needed (usually hundreds or thousands of images)
- The image labelling process is time-consuming, especially when it can only be done by experts with specific domain-knowledge.



WAGENINGEN UNIVERSITY & RESEARCH

38

38

maskAL: active learning for Mask R-CNN

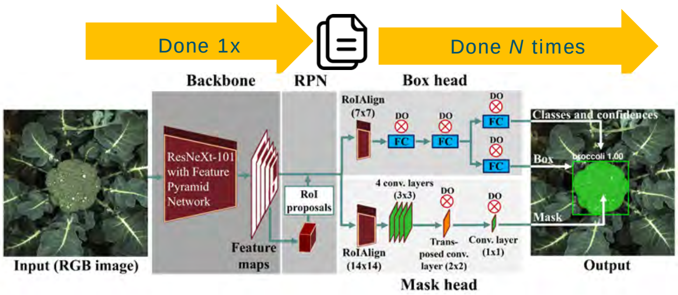
- In active learning, the most informative images are automatically gathered from a large pool of unlabelled images
- The hypothesis is that the generalisation performance of the CNN significantly improves when the training is done on the most informative images
- Because only the most informative images need to be annotated with active learning, the annotation effort will be reduced whilst preserving or improving the performance of Mask R-CNN.

39

39

Monte-Carlo dropout

- Basis for the uncertainty sampling is Monte-Carlo dropout (MCD)
- With MCD the image is repeatedly analysed with dropout. This dropout causes the random disconnection of some of the CNN's neurons, and this can lead to different model outputs.
- If the model outputs deviate, there seems to be uncertainty about the image, indicating that it can be a candidate for selection and annotation.

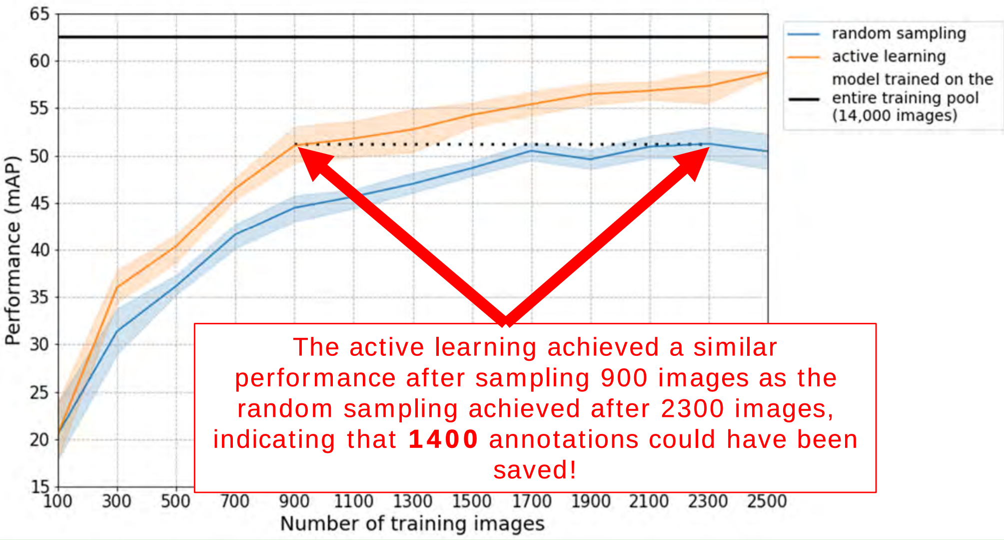


WAGENINGEN UNIVERSITY & RESEARCH

40

40

Experiment with MaskAL



Pieter M. Blok, Gert Kootstra, Hakim Elchaoui Elghor, Boubacar Diallo, Frits K. van Evert, Eldert J. van Henten (2022): Active learning with MaskAL reduces annotation effort for training Mask R-CNN on a broccoli dataset with visually similar classes, Comp. and Electronics in Agriculture., 197.

41

Pest detection and scouting

- Whiteflies, winged aphids and other (also beneficial) insects are attracted by the yellow color.
- They fly towards the trap and get stuck in the glue.
- Used to determine how many and which pests occur in the crop.
- Labor intensive to count.

WAGENINGEN UNIVERSITY & RESEARCH

42

Detection of insects with object detection

- F-RCNN Inception Resnet V2 deep-learning network for automatic counting of white fly (*Bemesia und Trialeurodes*) and beneficial insects (*Macrolophus* und *Nesidiocoris*) trapped on yellow sticky traps.
- Counting is used to forecast the population dynamics of pest and beneficial species.
- The final aim: decision support system

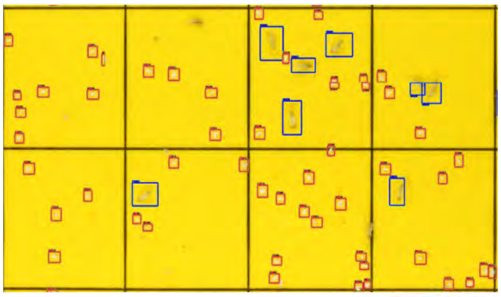


Photo reference: CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=39889

WAGENINGEN UNIVERSITY & RESEARCH
PROEFSTATION VOOR DE GROENTETEELT
Universiteit Antwerpen
JUNTA DE ANDALUCIA
NWO Netherlands Organisation for Scientific Research

43

Insect counting on sticky traps

- Trained with > 5000 manually annotated object instances
- Detection/classification results:

	Class	Predicted [%]				
		Wf	Mr	Nc	None	**Total**
Ground Truth	Wf	88.12	0.00	0.00	11.88	**100**
	Mr	0.00	91.80	1.95	6.25	**100**
	Nc	0.00	3.23	77.42	19.35	**100**
						85.78

Wf=White fly; Mr=Macrolophus; Nc=Nesidiocoris

Nieuwenhuizen, A.T. ; Hemming, J. ; Suh, H.K. (2018): Detection and classification of insects on stick-traps in a tomato crop using Faster R-CNN. Proceedings of the Netherlands Conference on Computer Vision NCCV 2018, September 1, 2018. Eindhoven, Netherlands.

44

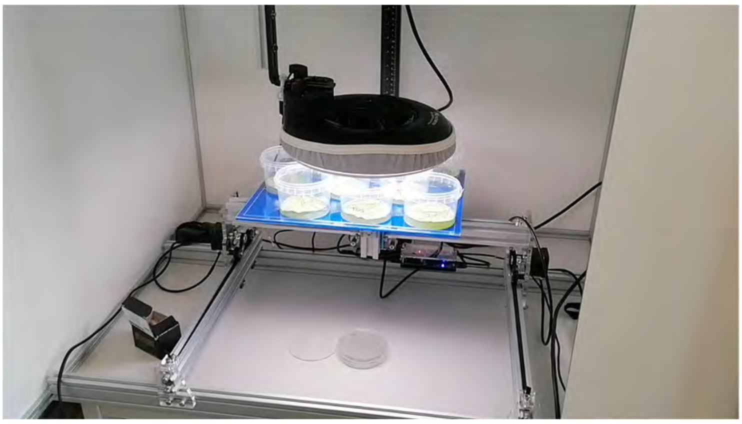
Automated evaluation of pest and disease bioassays

- Manual counting of objects (insect eggs) and estimating the diseased area on leaf samples is time consuming.
- Counting is partly subjective and person dependent.
- Can counting of objects and quantification of damaged leaf areas be done by deep learning?

WAGENINGEN UNIVERSITY & RESEARCH

45

Developed recording and analysis device

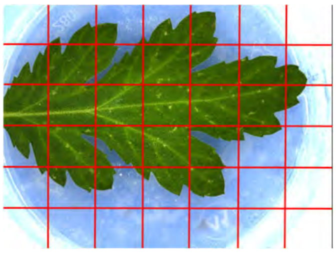


- Extreme high resolution camera/optics is needed
- Camera with autofocus

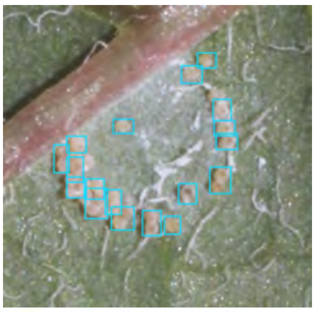
46

46

Whitefly egg counting



- Network: YoloV5, object detection
- Diameter whitefly egg is about 0.15 mm
  - 350 pixel/millimetre



WAGENINGEN UNIVERSITY & RESEARCH

47

Validation experiment (whitefly eggs)

	MAPE (human)	MAPE Deep learning
Tomato	15%	4%
Cucumber	17%	3%
Poinsettia	13%	5%
Sweet pepper	23%	27%

MAPE = mean absolute percentage error

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

- Results of deep learning outperforms human.
- Issues image quality:
  - Condense on petri-dish
  - Unsharp images because of bended leaves

WAGENINGEN UNIVERSITY & RESEARCH

48

OPTimised Integrated Pest MAnagement

OPTIMA

http://www.optima-h2020.eu/

Gerrit Polder, Pieter M. Blok, Tim van Daalen, Joseph Peller, Nikos Mylonas (2021). Early Disease Detection in Apple and Grape Using Deep Learning on a Smart-Camera. Proceedings of the European Conference on Agricultural Engineering AgEng2021. Universidade de Évora, pp. 51-56.

49

Fully Documented Fisheries

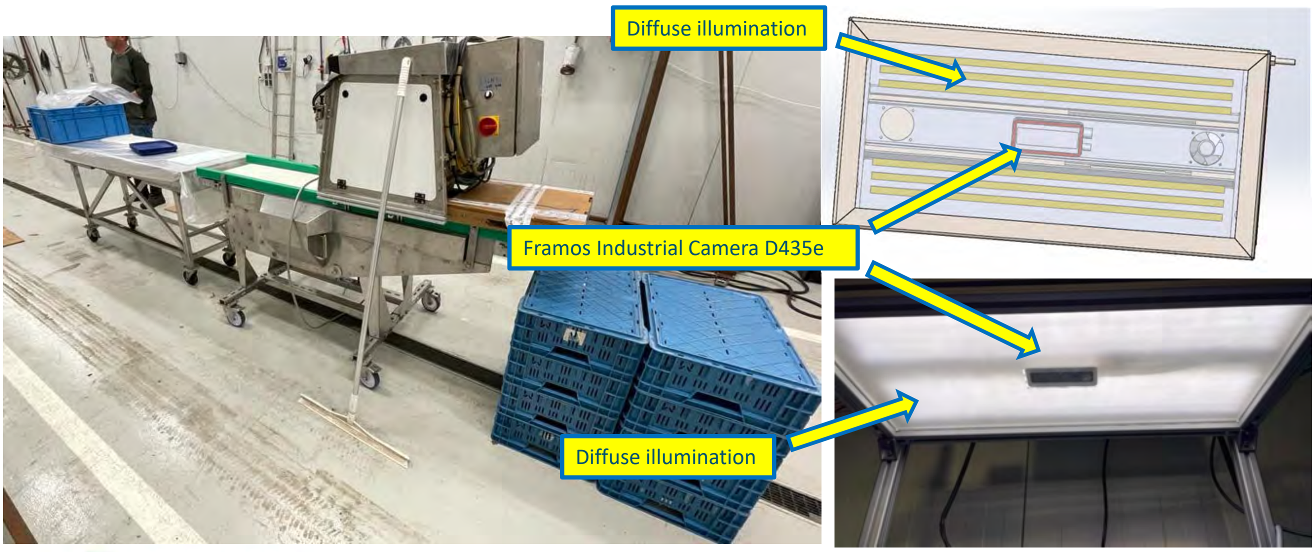
- Landing obligation in the EU obliges fishing vessels to, sort, store, and land all undersized fish.
- Development of an automatic detection system that is able:
  - to acquire images of the discard;
  - to classify and count the different species separating fish from the debris;
  - to estimate the weight and length of the discard by species;

WAGENINGEN UNIVERSITY & RESEARCH

Founded by the European Maritime and Fisheries Fund (EMFF)

50

Fully Documented Fisheries – prototype



WAGENINGEN UNIVERSITY & RESEARCH

51

51

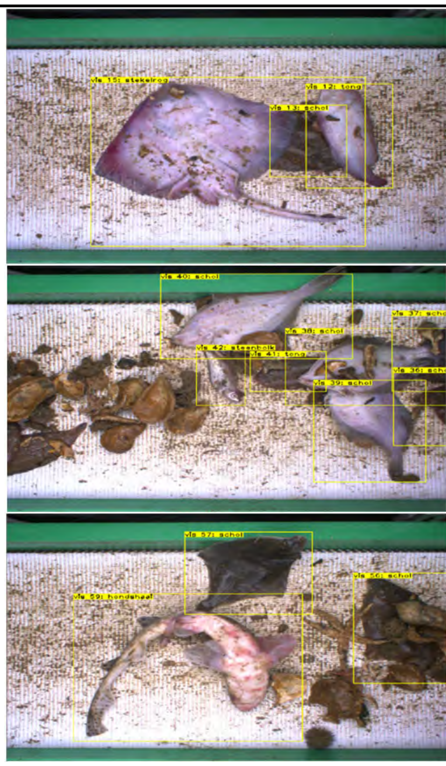
YOLOv3

Van Essen et al. 2021 "Automatic discard registration in cluttered environments using deep learning and object tracking: class imbalance, occlusion, and a comparison to human review". ICES Journal of Marine Science (Under review)

52

Identification results

- Recognizes 80% of fish with a 'count' error of 20%
- Issues: Occlusion, debris and location of the fish



WAGENINGEN UNIVERSITY & RESEARCH

53

53

Final conclusions

- In agri food almost every new machine vision project is using DL
- DL is just one piece of the puzzle in an application.
- Training process is based on large amounts of data, labour intensive labelling is needed.
- Expert knowledge is needed to select architecture, and to bootstrap and fine tune the network.
- Black box: deep-learning is incapable of providing arguments why it has reached a certain conclusion.
- Resource-Demanding Technology. High computational costs.
- DL methods clearly show superior classification results.

WAGENINGEN UNIVERSITY & RESEARCH 54

54

Thank you for your attention

Acknowledgements to my WUR colleagues:

Pieter Blok, Bart van Marrewijk, Joseph Peller, Gerrit Polder, Angelo Mencarelli, Jos Ruizendaal, Menno Sytsma, Toon Tielen, Ard Nieuwenhuizen and many more.

Contact:

jochen.hemming@wur.nl

www.wur.eu/agrofoodrobotics

WAGENINGEN UNIVERSITY & RESEARCH 55

55