NPPN

The challenge of robust trait estimates in plant phenotyping with Machine Learning

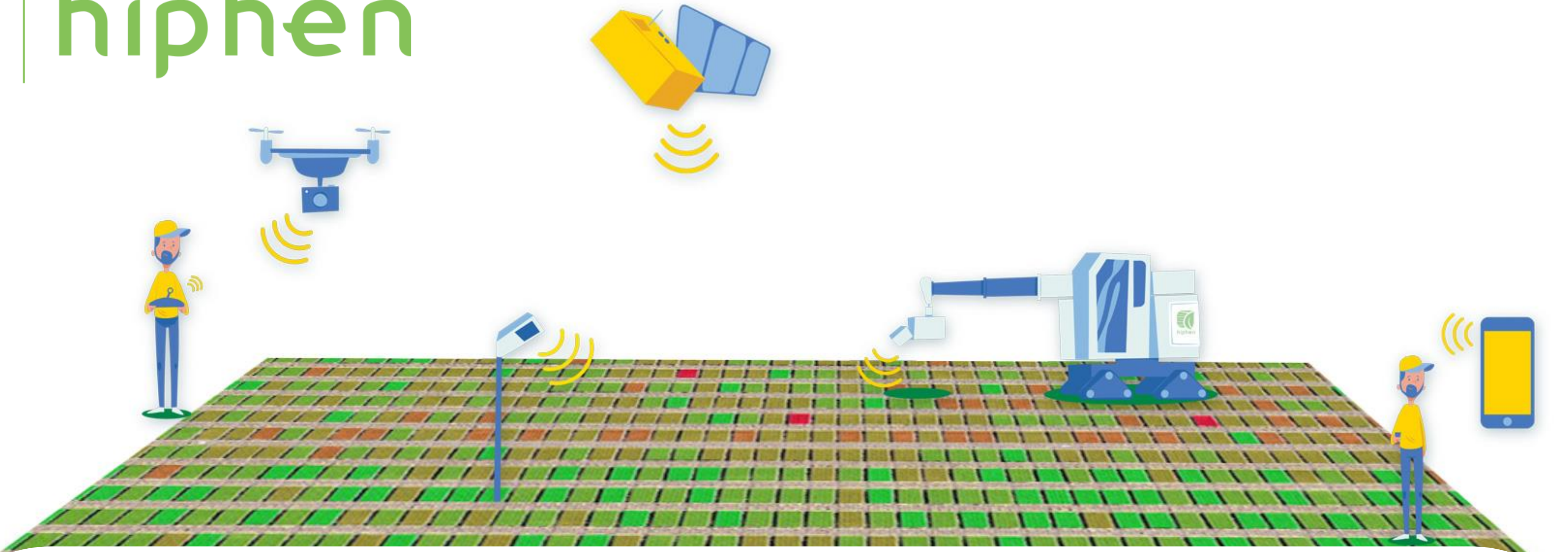
DATE

February 2022

AUTHOR(S)

Etienne David

We are expert in remote sensing for agriculture



Mechanistic crop knowledge

Automated data processing

Cloud-based data engine built for scale

Artificial intelligence

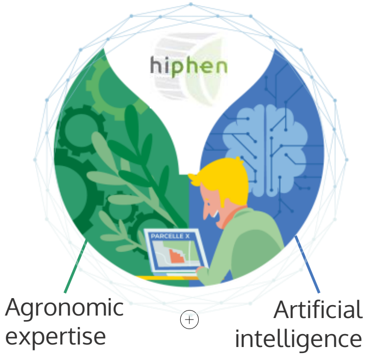
Decision making

Our 20+ agronomists and data scientists are expert at turning crop images into valuable agronomic traits

Our team is laser focused on **delivering excellence**

Discover our team in video

Our remote sensing expertise arises from agronomic knowledge combined with artificial intelligence



CAPTE

INRAE

hiphen

ARVALIS Institut du végétal

(...) And more...

Who am I ?

Master thesis – Computer vision for bioinformatics

Phd « Robust trait estimates with Deep Learning on high resolution RGB imagery »

2017

2022

2015

2021

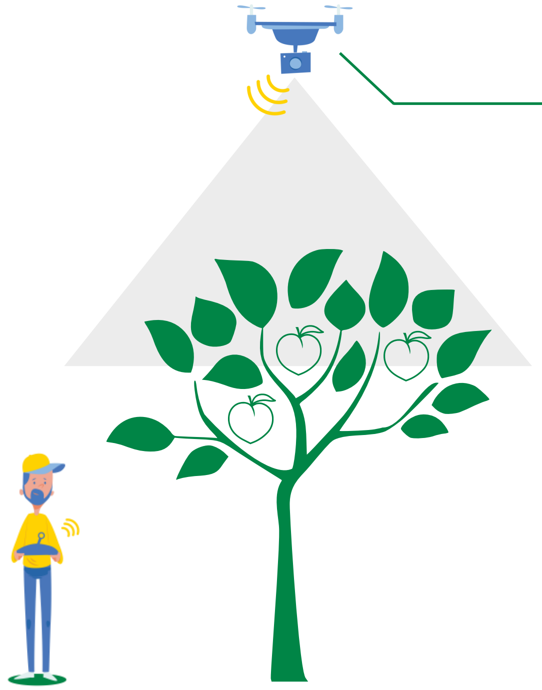
Data scientist

Science Manager

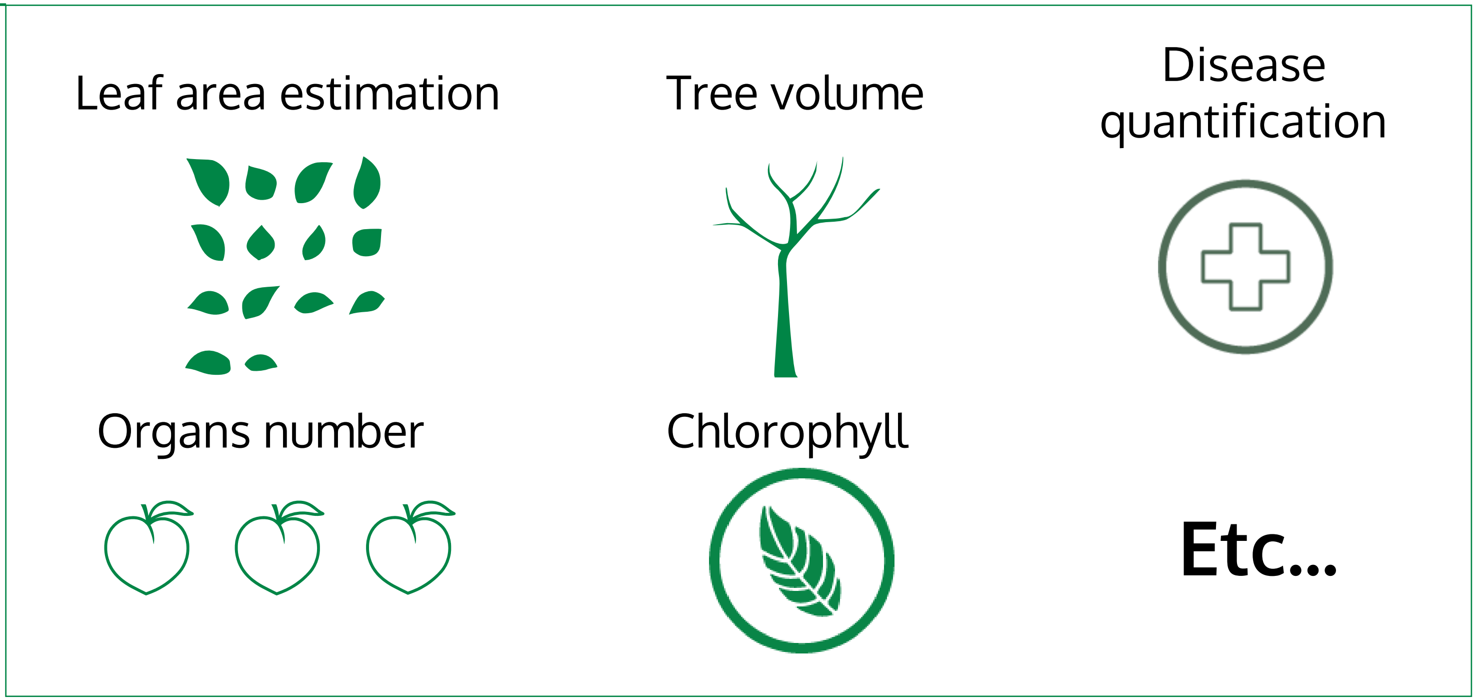
edavid@hiphen-plant.com

State trait estimates is the basis phenotyping

Drone

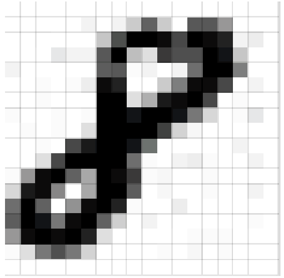


High-througput phenotyping provide numerous state traits (biophysical, biochemical, sanitary etc…)



Two main algorithm’s families can be used

Sensor’s signal



Measurement of a physical flux

Biophysical model

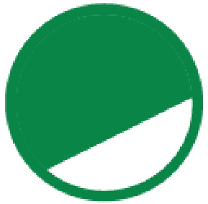
- Remote sensing
- « physical » calibration

Empirical model

- Computer science
- semantic calibration

Measured trait

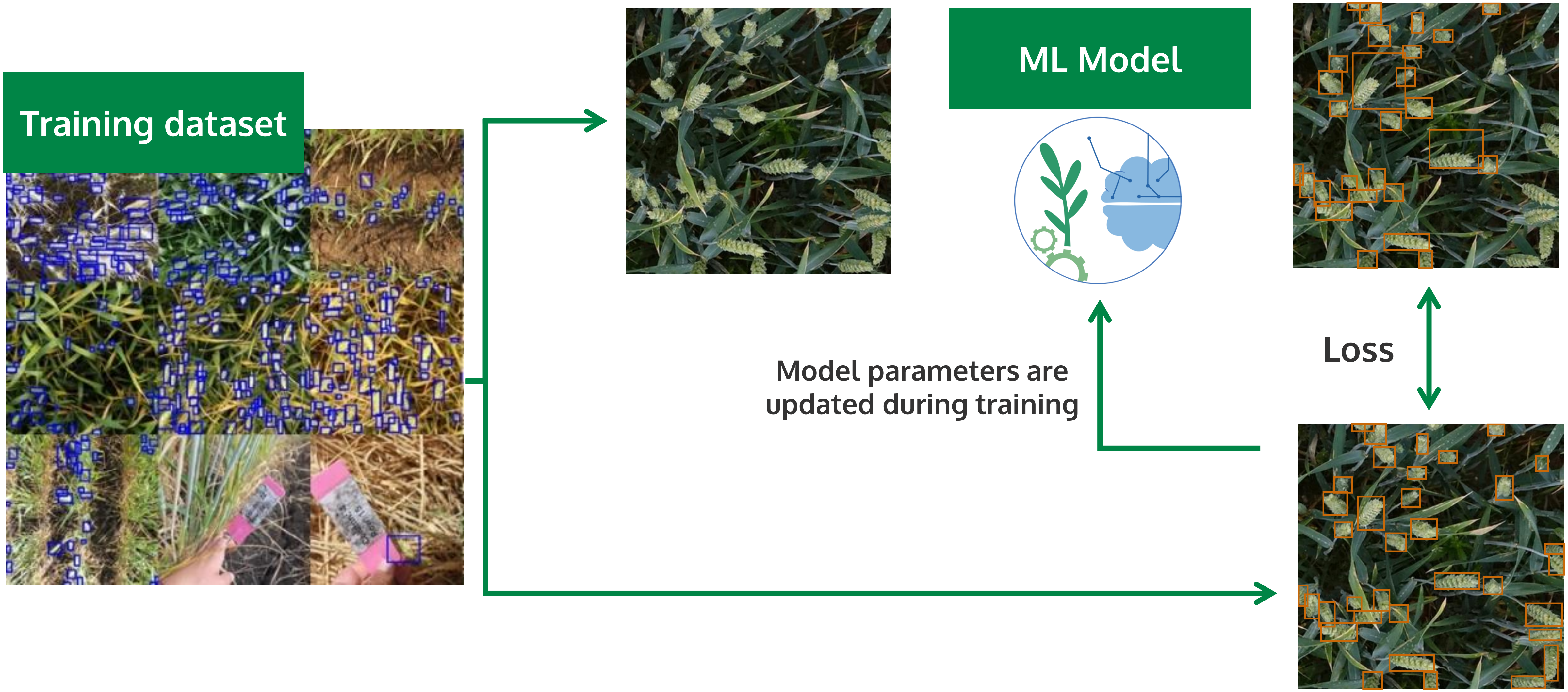
Green Cover Fraction



Calculated trait

Leaf Area Index

ML is a powerful family of empirical model



Deep Learning models learn representation

	Raw Data		Change of representation		Decision on better representation		
Traditional Machine Learning		→	Choice of handcraft features - Vegetation index - Color space	→	SVM (Learned)	→	
Deep Learning		→	Learned features	+	Learned classifier	→	

hiphen

Deep Learning allows to access new traits

Trait	Biophysical	Empirical	
		Machine Learning	Deep Learning
Green Cover Fraction	√	√	√
Leaf Area Index	√	√	√
Height	√		√
Lodging Score	√		√
Leaf Chlorophyl Content	√	√	√
Canopy Chlorophyll Content	√	√	√
Plants Density		√	√
Crop Cover Fraction			√
Senescent Fraction			√
Head Density			√
Disease Fraction			√

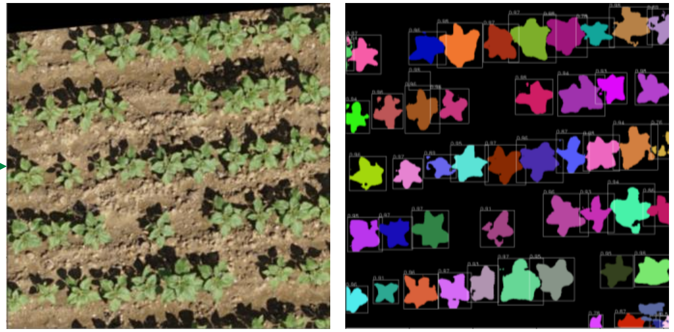
10

The domain shift problem

A **domain shift** is a change in the **data distribution** between an algorithm's training dataset and the encountered images when deployed.

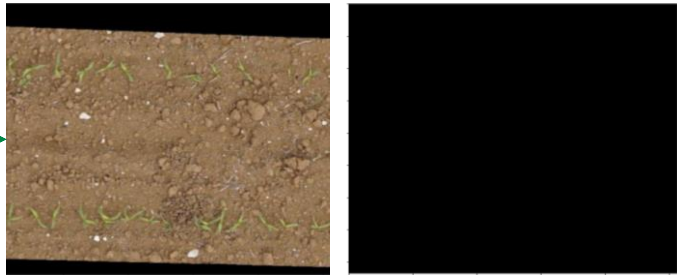
Sunflower detection

Sunflower



Same than source domain

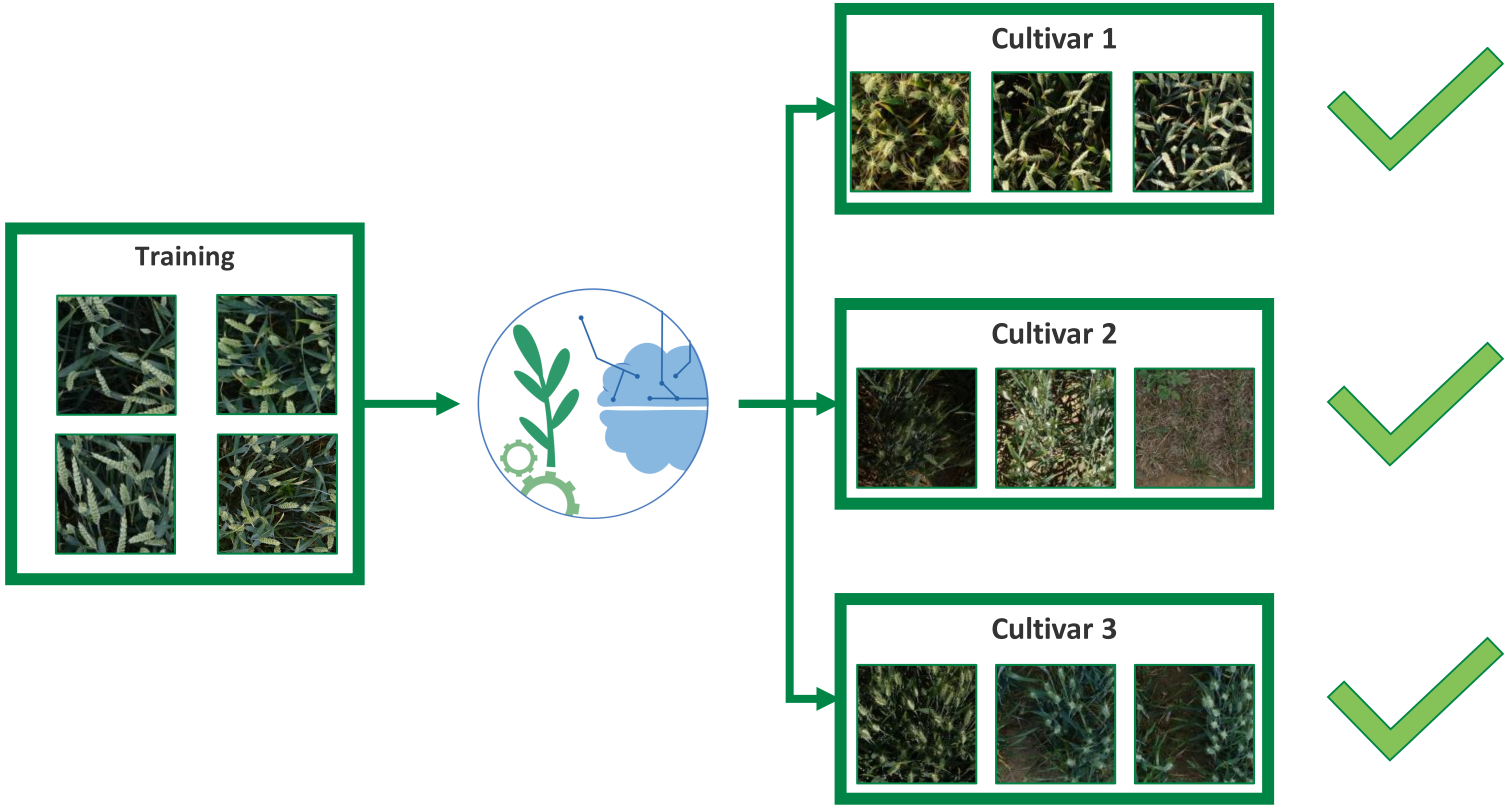
Maize



Different than source domain

Does it cause a robustness problem ?

Robustness in plant phenotyping is to the capacity of an algorithm to produce an **unbiased** trait estimate for **all images** acquired with the **same** protocol of acquisition.



How much does domain shift affects ML algorithms performance in Plant Phenotyping ?

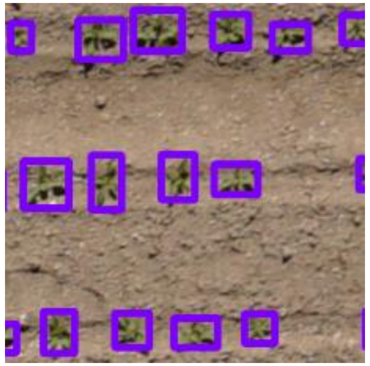
Session of acquisition as a proxy for diversity

Session of acquisition

Vector X Time unit X Experimental unit
(ex: 1 Flight over one trial)

Fixed unit of diversity

UAV plant counting



3-4 leaves

Loam soil

Row spacing of 45 cm

Sony alpha 6000

Cloudy condition

GSD of 4.5 mm

Wheat head localization

Post-Flowering

Row spacing of 17,5 cm

Sony alpha 6000

Cloudy condition

GSD of 4.5 mm

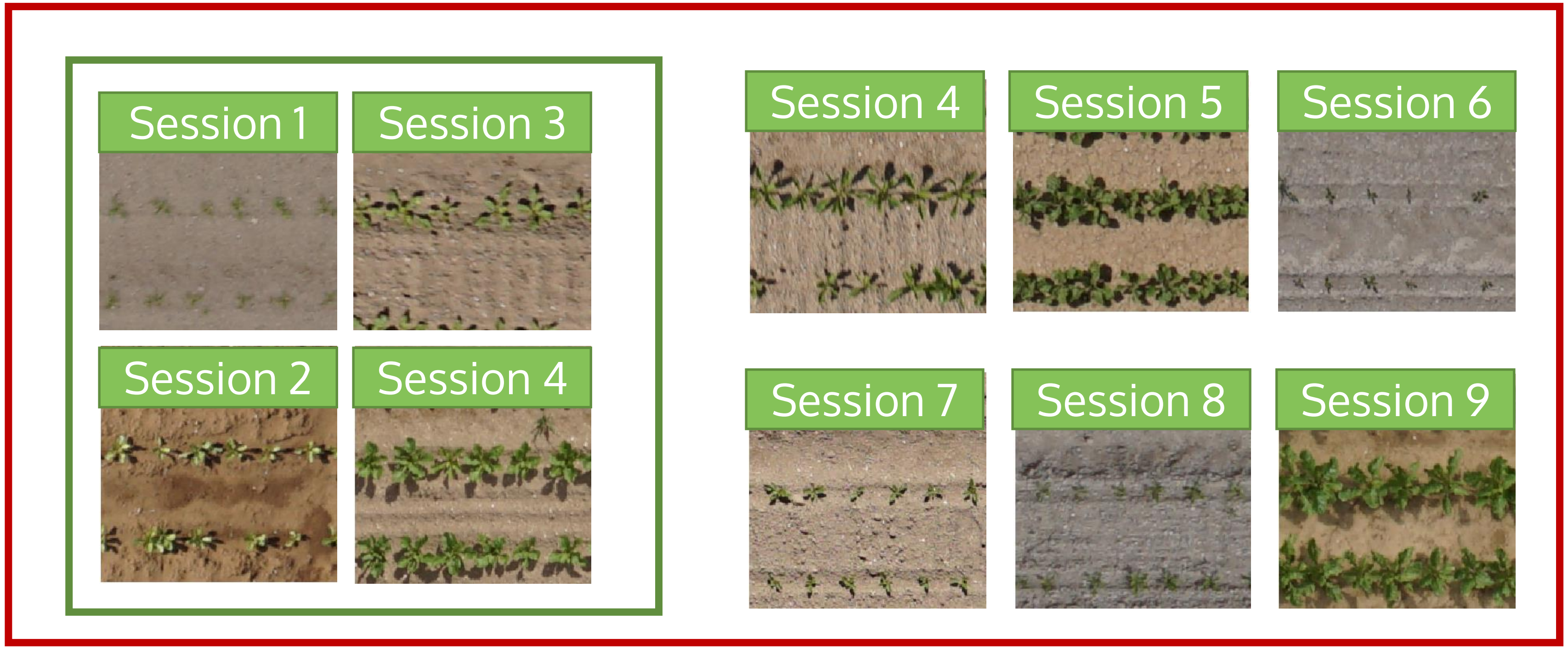
Studies are conducted on few sessions

Example on plant counting

Study	UAV	Crop	Sessions	Localization	Test independency
(Quan et al., 2019)		Maize	10	✓	
(Ribera et al., 2017)	✓	Sorghum	2		
(Valente et al., 2020)	✓	Spinach	1		
(Liu et al., 2020)	✓	Maize	2	✓	
(Lin and Guo, 2020)	✓	Sorghum	2	✓	
(Madec et al., 2019)		Wheat	2	✓	✓
(Xiong et al., 2019)	✓	Wheat	10+		✓

The ideal dataset covers all posibilities

Training diversity = Application diversity



The ideal dataset should include as many sessions of acquisition as possible to cover **expected diversity** !

Sharing the labelling burden

Institution #1

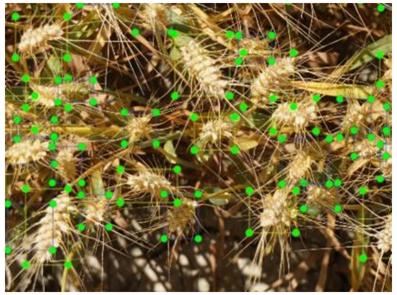
Dataset #1

Institution #2

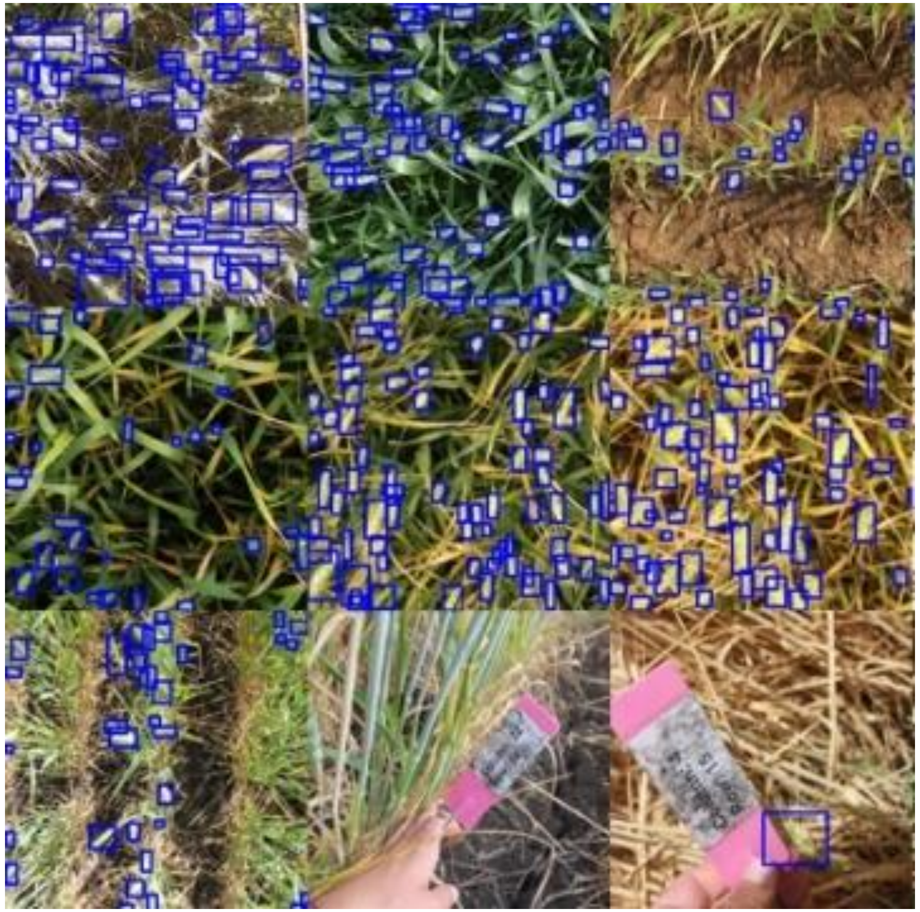
Dataset #2

Institution #3

Dataset #3



An example of large, collaborative dataset



Global WHEAT Dataset

Database/Software Article

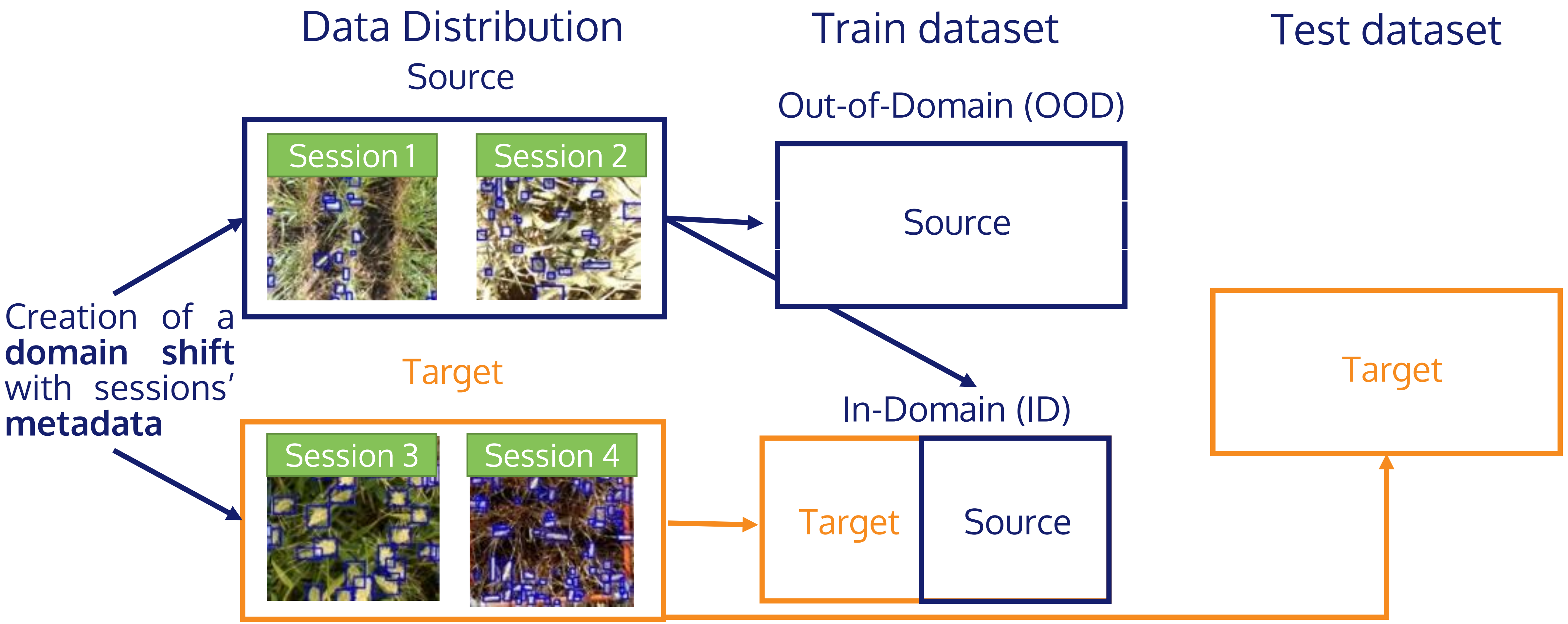
Global Wheat Head Detection 2021: An Improved Dataset for Benchmarking Wheat Head Detection Methods

Etienne David,[1,2] Mario Serouart,[1,2] Daniel Smith,[3] Simon Madec,[1,3] Kaaviya Velumani,[2,4] Shouyang Liu,[5] Xu Wang,[6] Francisco Pinto,[7] Shahameh Shafiee,[8] Izzat S. A. Tahir,[9] Hisashi Tsujimoto,[10] Shuhei Nasuda,[11] Bangyou Zheng,[12] Norbert Kirchgessner,[13] Helge Aasen,[13] Andreas Hund,[13] Pouria Sadhegi-Tehran,[14] Koichi Nagasawa,[15] Goro Ishikawa,[16] Sébastien Dandrifosse,[17] Alexis Carlier,[17] Benjamin Dumont,[18] Benoit Mercatoris,[17] Byron Evers,[6] Ken Kuroki,[19] Haozhou Wang,[19] Masanori Ishii,[19] Minhajul A. Badhon,[20] Curtis Pozniak,[21] David Shaner LeBauer,[22] Morten Lillemo,[8] Jesse Poland,[6] Scott Chapman,[3,12] Benoit de Solan,[1] Frédéric Baret,[2] Ian Stavness,[20] and Wei Guo[19]

ARVALIS Institut du végétal
INRAE
The University of Queensland Australia
University of Saskatchewan
東京大学 The University of Tokyo
Rothamsted Research
ETH zürich
CIMMYT International Maize and Wheat Improvement Center
Kansas State University
京都大学 Kyoto University
Norwegian University of Life Sciences
Liège université Gembloux Agro-Bio Tech
ARC
Terraphenotyping Reference Platform
NARO National Agriculture and Food Research Organization

hiphen

How to measure the domain shift ?



Faster-RCNN to localize and count

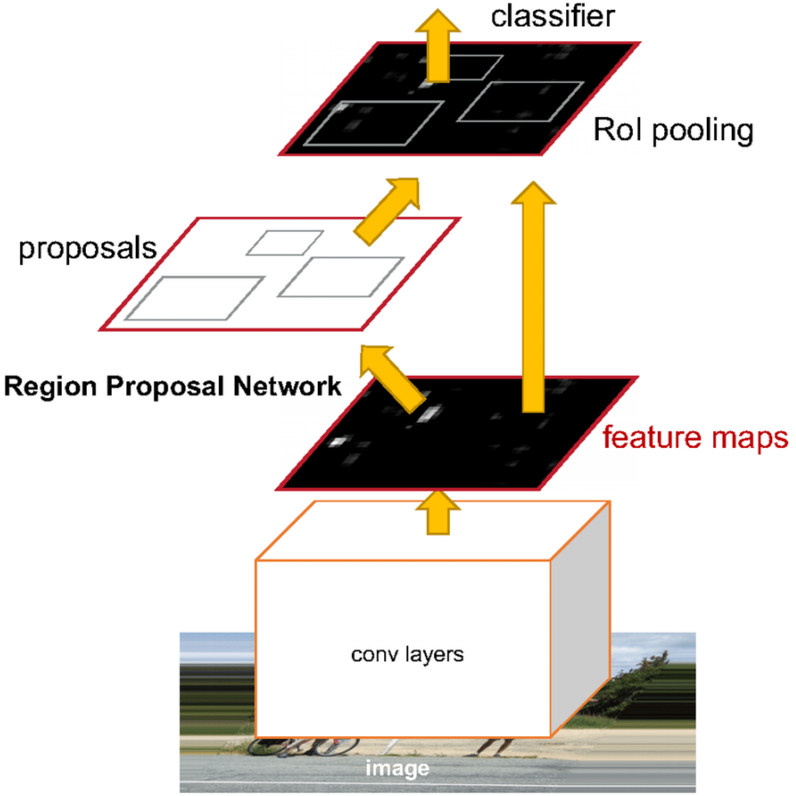
Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun

$$\mathcal{L}(p,u,t^u,v) = \mathcal{L}_{\text{cls}}(p,u) + 1[u \geq 1]\mathcal{L}_{\text{box}}(t^u,v)$$

$$\mathcal{L}_{\text{cls}}(p,u) = -\log p_u$$

$$\mathcal{L}_{\text{box}}(t^u,v) = \sum_{i\in\{x,y,w,h\}} L_1^{\text{smooth}}(t_i^u - v_i)$$

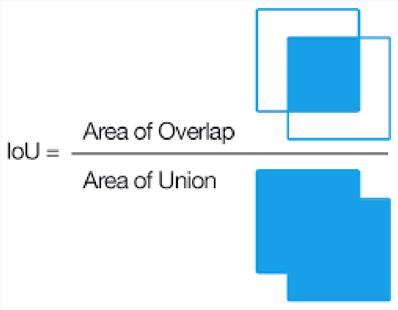


Design metrics to evaluate robustness

Localization

TP: True positive, FP: False positive , FN: False negative

Intersection over Union (IoU)



Accuracy (Acc)

- $\sum \frac{TP(IoU)}{TP(IoU)+FP(IoU)+FN(IoU)}$

Average Domain Accuracy (ADA)

$$\frac{1}{D}\sum_{d=1}^{D}\frac{1}{n_d} * \sum_{i=1}^{n_d} Acc_{di}$$

Counting

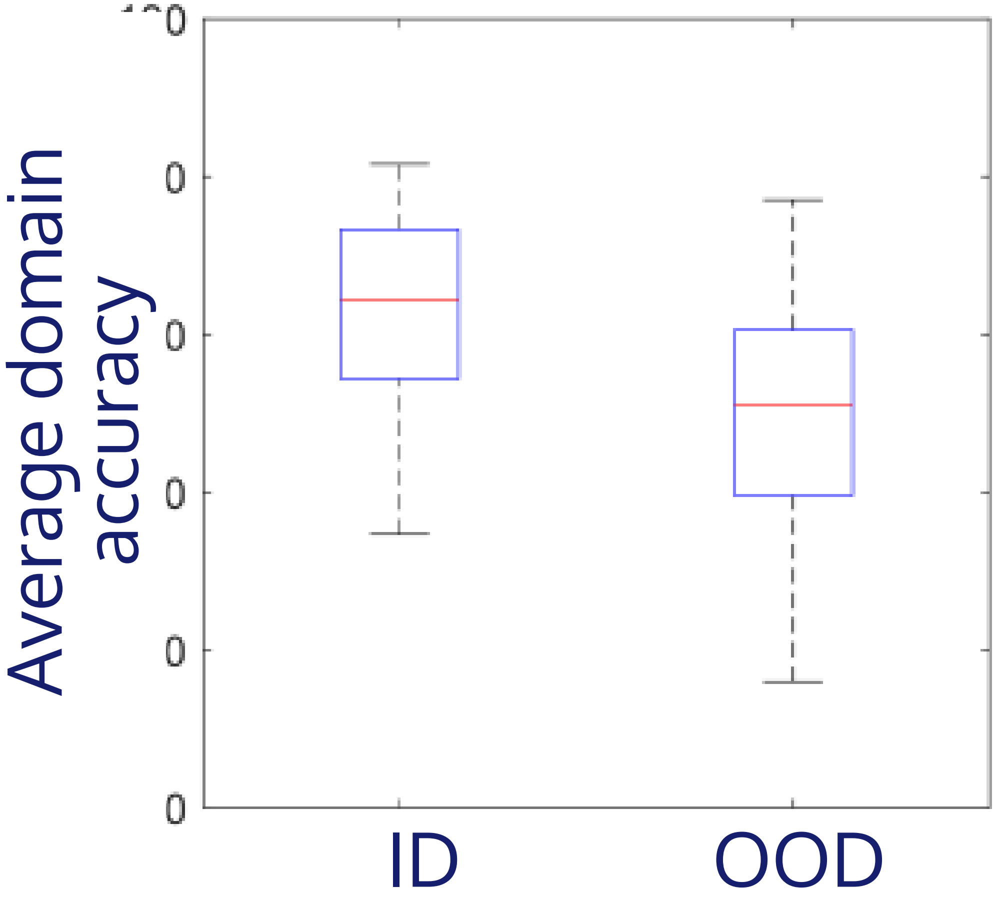
Root Mean Square Error (RMSE)

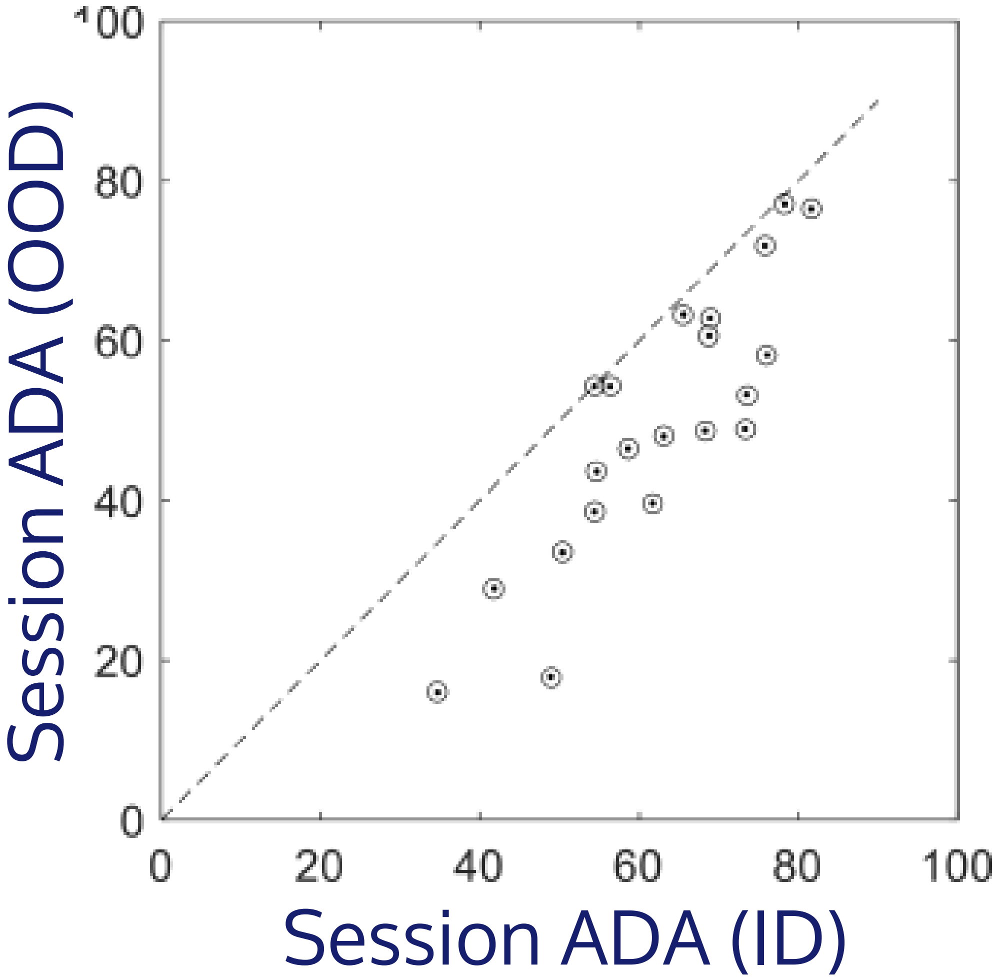
$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}$$

Relative RMSE (rRMSE)

$$rRMSE = \frac{RMSE}{\hat{x}}$$

OOD largely underperform against ID model





hiphen

How to solve the domain shift problem ?

Improve size of training dataset → Use engineering best practices → Explicitly train for robustness

Global WHEAT Dataset

kaggle

AIcrowd

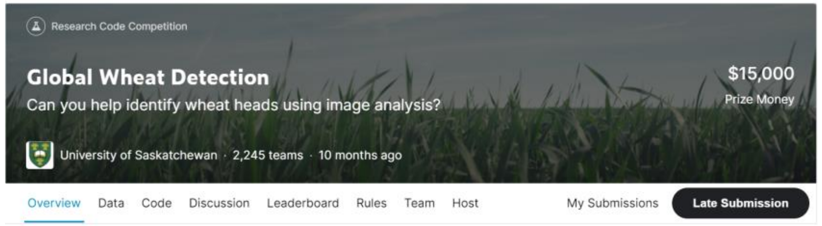
Stanford University

WILDS

Global Wheat Challenges

2020

kaggle



2200+

2021

AIcrowd

400+

GIFS | GLOBAL INSTITUTE FOR FOOD SECURITY

Growing science for life

Nutrien - a Founding Partner

Kubota

#DigitAg

hiphen

Plant Phenomics

A SCIENCE PARTNER JOURNAL

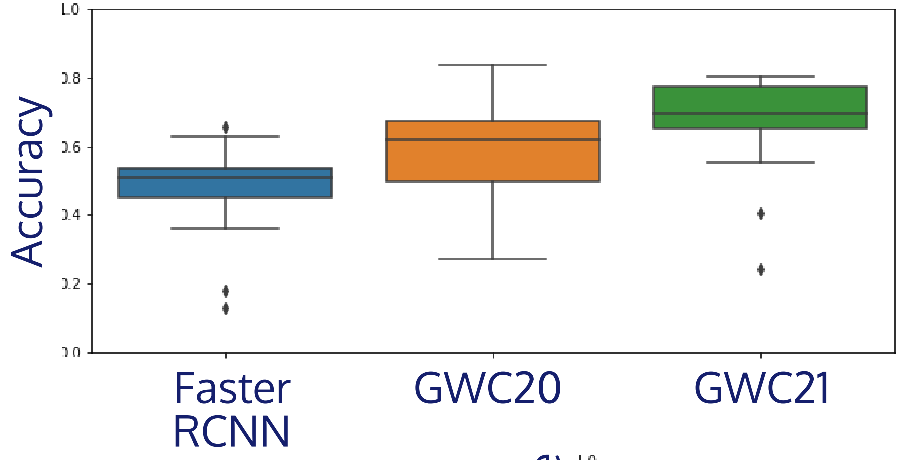
Quantomics

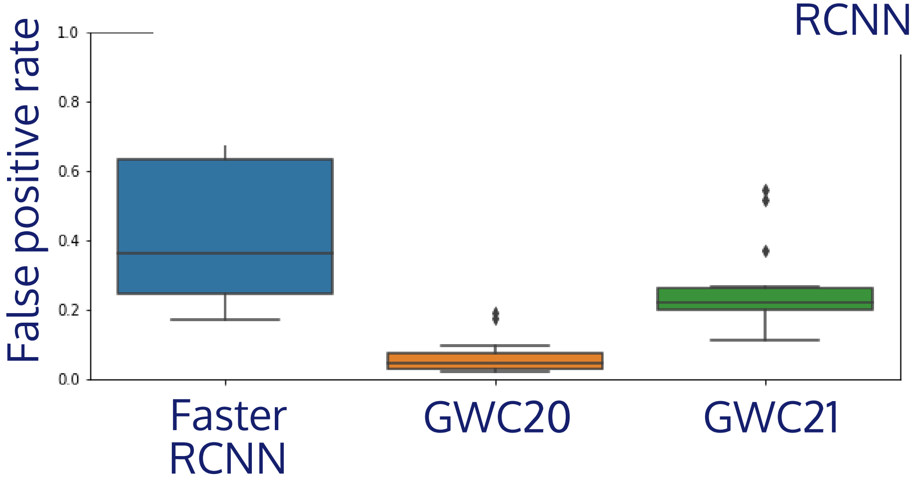
Navigator for Future Agriculture

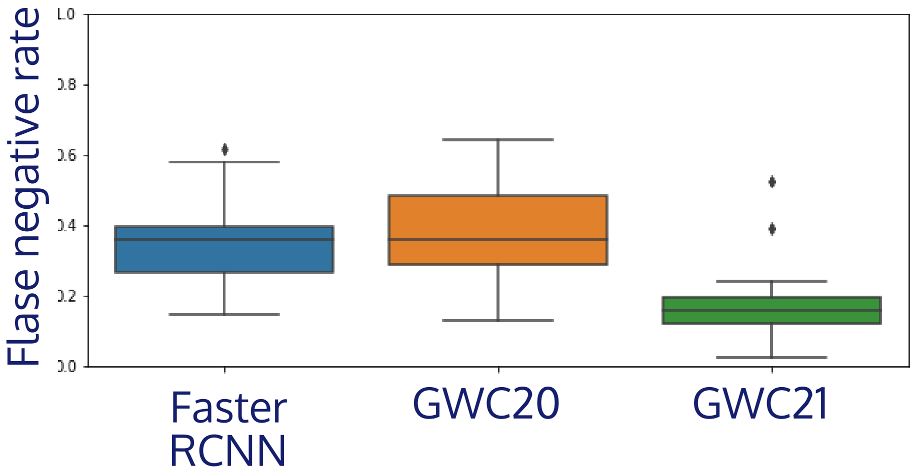
Common characteristics of winning solutions

	Rank	Solution name	Data preparation	Data Augmentation	Architecture	Ensemble training strategy
Baseline		Madec	No	No	Faster-RCNN	No
GWC_2020	1	DungNB	No	Mixup ; Custom mosaic	EfficientDet; FasterRCNN	Random subsampling
	2	OverFeat	Jigsaw	Mixup, Cutmix	Efficentdet	Random subsampling
	3	Javu	No	Mixup	YoloV3	No
GWC_2021	1	RandomTeamName	No	Mosaic	Yolov5	Domain subsampling
	2	David_jeon	Model is applied on 1600 px images	Mosaic; CutMix	Yolov5	No
	3	SMART	**Network to correct image**	CutMix	Yolov4	Yes

Winning solutions in 2020 and 2021 overperforms our baseline







How to solve the domain shift problem ?

Improve size of training dataset → Use engineering best practices → Explicitly train for robustness

Global WHEAT Dataset

kaggle

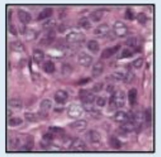
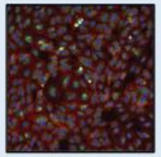
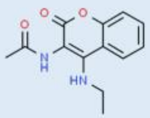
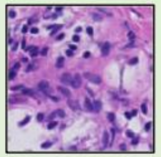
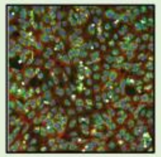
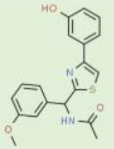
AIcrowd

Stanford University

WILDS

Explicitly train for robustness is not working yet

WILDS: A Benchmark of in-the-Wild Distribution Shifts

	Domain generalization					Subpopulation shift	Domain generalization + subpopulation shift			
Dataset	iWildCam	Camelyon17	RxRx1	OGB-MolPCBA	GlobalWheat	CivilComments	FMoW	PovertyMap	Amazon	Py150
Input (x)	camera trap photo	tissue slide	cell image	molecular graph	wheat image	online comment	satellite image	satellite image	product review	code
Prediction (y)	animal species	tumor	perturbed gene	bioassays	wheat head bbox	toxicity	land use	asset wealth	sentiment	autocomplete
Domain (d)	camera	hospital	batch	scaffold	location, time	demographic	time, region	country, rural-urban	user	git repository
# domains	323	5	51	120,084	47	16	16 x 5	23 x 2	2,586	8,421
# examples	203,029	455,954	125,510	437,929	6,515	448,000	523,846	19,669	539,502	150,000
Train example						What do Black and LGBT people have to do with bicycle licensing?			Overall a solid package that has a good quality of construction for the price.	`import numpy as np … norm=np.___`
Test example						As a Christian, I will not be patronizing any of those businesses.			I *loved* my French press, it's so perfect and came with all this fun stuff!	`import subprocess as sp p=sp.Popen() stdout=p.___`
Adapted from	Beery et al. 2020	Bandi et al. 2018	Taylor et al. 2019	Hu et al. 2020	David et al. 2021	Borkan et al. 2019	Christie et al. 2018	Yeh et al. 2020	Ni et al. 2019	Raychev et al. 2016

hiphen

How to solve the domain shift problem ?

Improve size of training dataset

Use engineering best practices

Explicitly train for robustness

Global WHEAT Dataset

kaggle

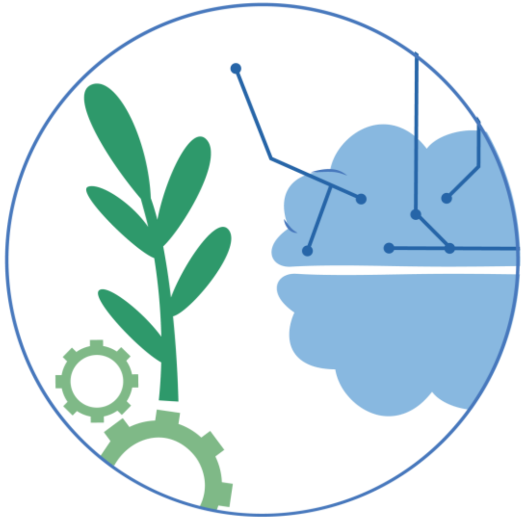
AIcrowd

Stanford University

WILDS

Trait evaluation over large range of conditions is difficult

Environmental conditions



The higher the expected diversity, the more difficult it is to measure a trait accurately.

Trait evaluation error

Genetic variations

ML for phenotyping is more than a script

ML method for trait estimation

Training

1. What are the limit of the algorithm ?
2. Is my training dataset representative of my expected diversity ?
3. Do I use most of the engineering best practices?

Initial evaluation

1. Do I evaluate on different conditions than training ?
2. Do I use more than one session of acquisition ?
3. Is my performance the same over different sessions of acquisition ?

Production

1. How do I check the accuracy of my algorithm ?
2. How do I improve the model over time ?

Hiphen's Human in the loop approach

Human-in-the-loop approach

Data acquisition → Data upload → Data processing → Data validation → Data visualization & analytics

Data acquisition

Acquisition protocol for drones

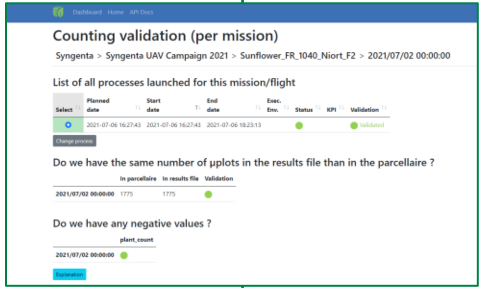
Data upload

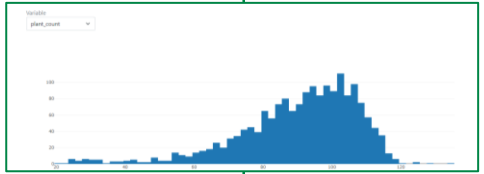
Automated quality acquisition reports

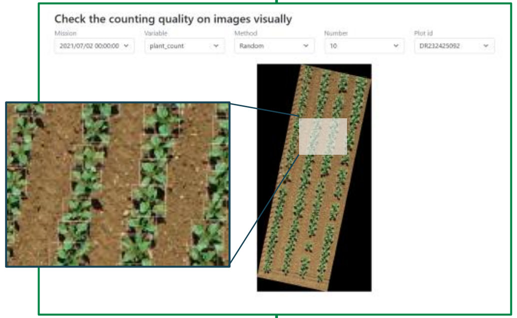
We give you feedback

Data processing

Data validation



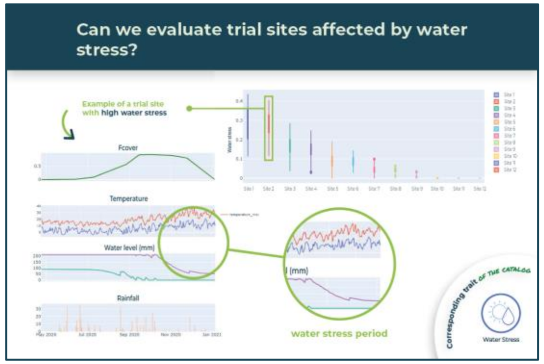




Auto-generated data validation report

Data visualization & analytics

Browse and visualize all the traits computed



We help you to understand your data with analytics

hiphen

You want to discover more?

Discover more about our solutions with the catalogs

Discover Now

Discover Now

Discover Now

Coming Soon!

Welcome to Avignon !

Hiphen

120 rue Jean Dausset

Agroparc – Bâtiment technicité

84140 Avignon

France

+33 (0)7 84 14 31 63

hiphen-plant.com | contact@hiphen-plant.com