Università
degli Studi
di Ferrara

Implementing Deep Neural Networks for in-situ crop yield prediction

Supervisor

Prof Fabio Mantovani

Candidate

Nicola Macchioni

Outline

- Crop yield prediction challenge
- Rationale
- Experimental site and data taking
- Customization of the DNN pipeline
- Results and discussions

Wheat yield: from classical estimation to ML

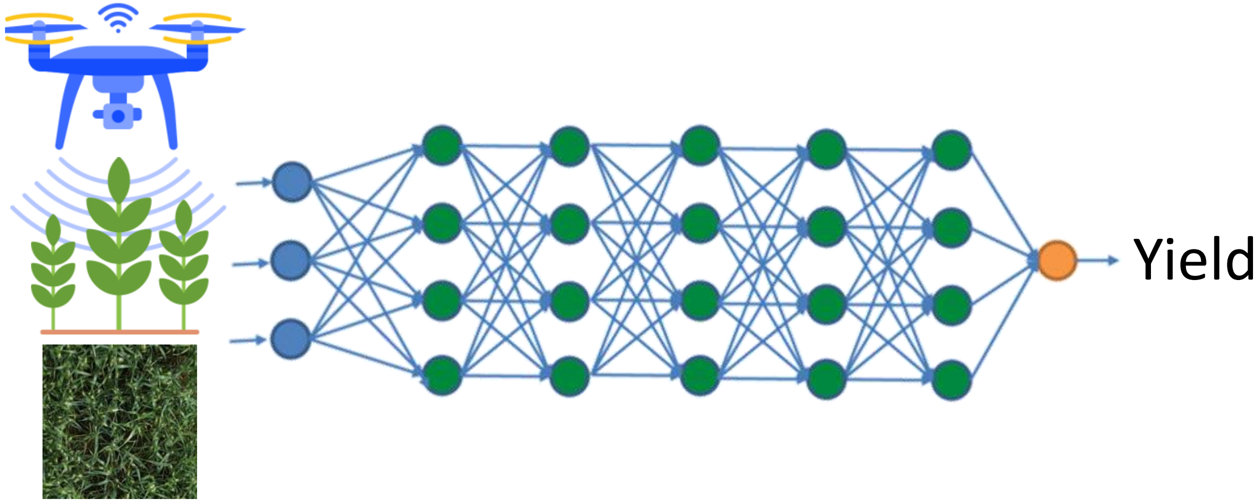
Wheat covers 15% of the world's arable land (**220 million ha**). 66% of global wheat production (**5 of 7.7 billion q per year**) is used for food. With population growth to 9.7 billion by 2050, demand for wheat will rise by **1.3 billion q per year**.

Yield $\left(\frac{\boldsymbol{q}}{\boldsymbol{ha}} = 0.01\,\frac{\boldsymbol{kg}}{\boldsymbol{m}^2}\right) = \mathcal{F}(\textit{soil conditions, temperature, genotype, environmental stress, fertilizer, …})$

Classical yield formula

$$\text{Yield Estimate}\left(\frac{q}{ha}\right) = \frac{\#\text{heads}}{ha}\cdot\frac{\#\text{grains}}{\text{heads}}\cdot\frac{q}{\text{grain}}$$

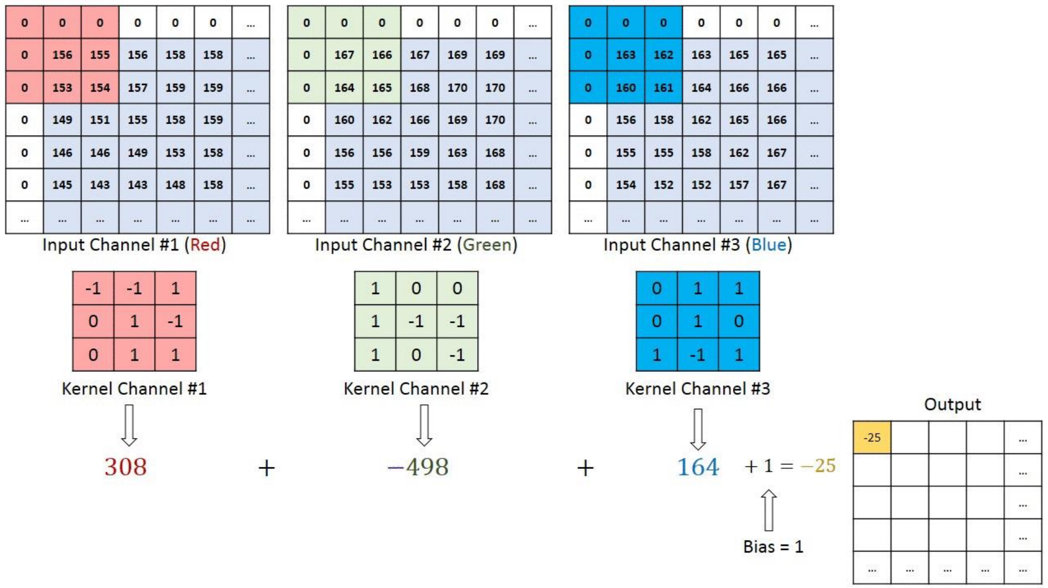
Automated phenotyping



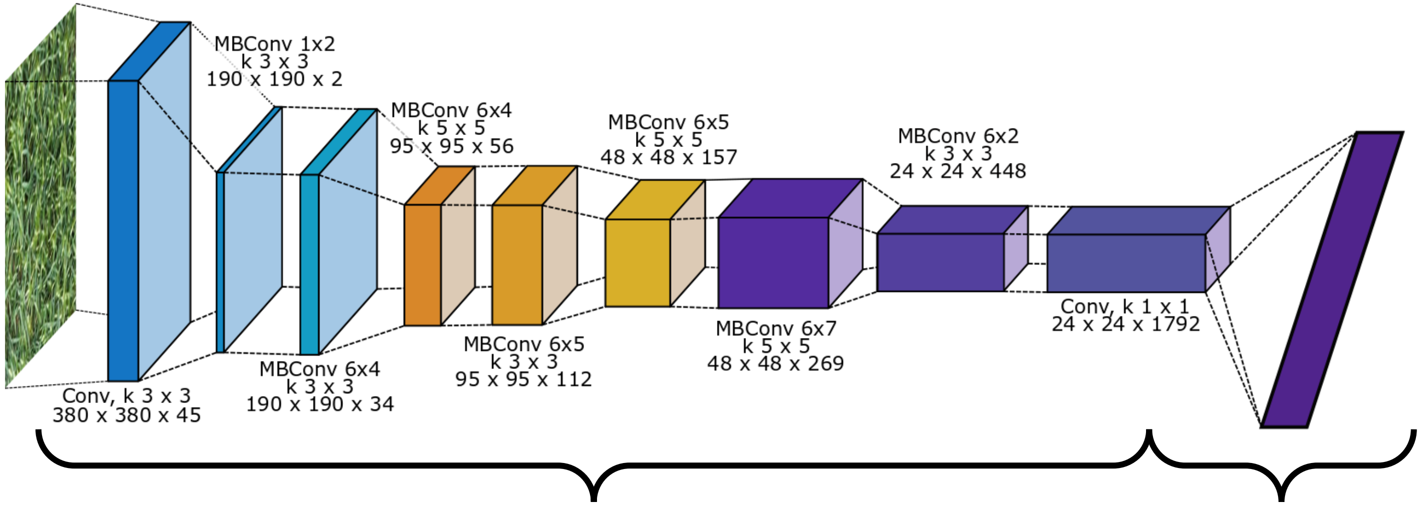
Automatic extraction of color, growth uniformity, spatial patterns, canopy texture, environmental damages.

Rationale: Convolutional + Feedforward Neural Network

- The CNN processes **quadratic RGB images** and condenses features (i.e. texture, borders) into **1D feature vectors** through convolutions
- I employed a **pre-trained CNN** (EfficientNetB4*, 17 673 823 parameters) for feature extraction



RGB image (380 px x 380 px x 3)



Convolutional layers

Pooling layers

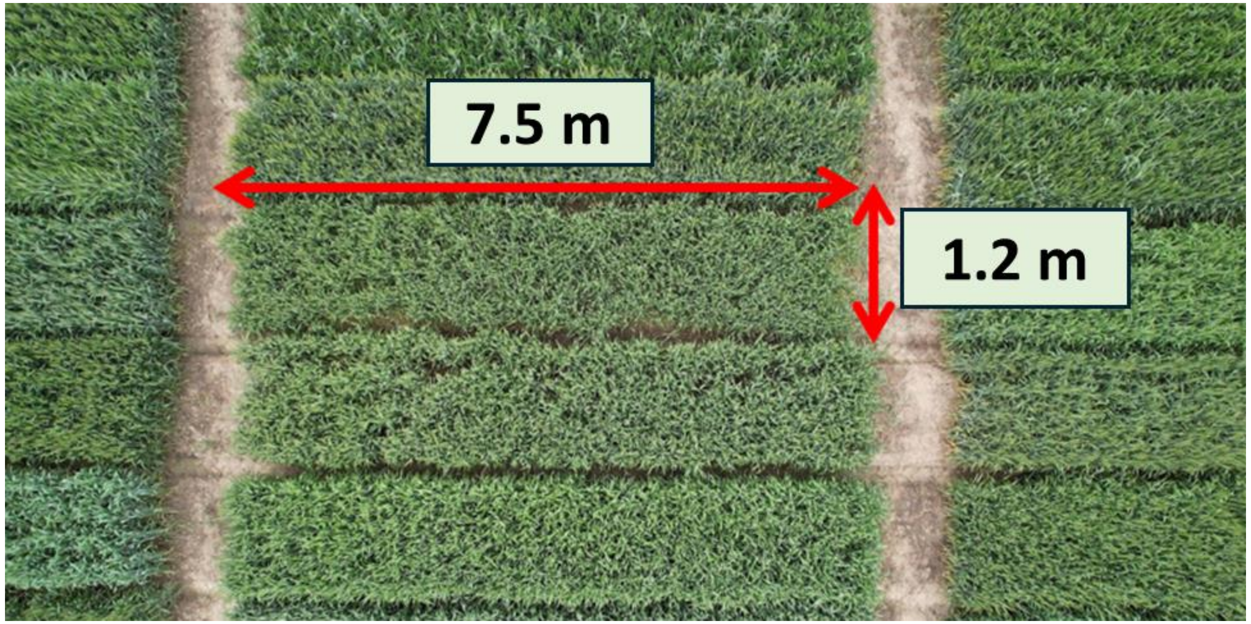
+ FNN

* M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019.

Experimental site

- **Airborne surveys** were run on a **wheat** field in **Argelato (BO)**, Italy, during the final two months before harvest (**May and June 2024**) to acquire **RGB images**.

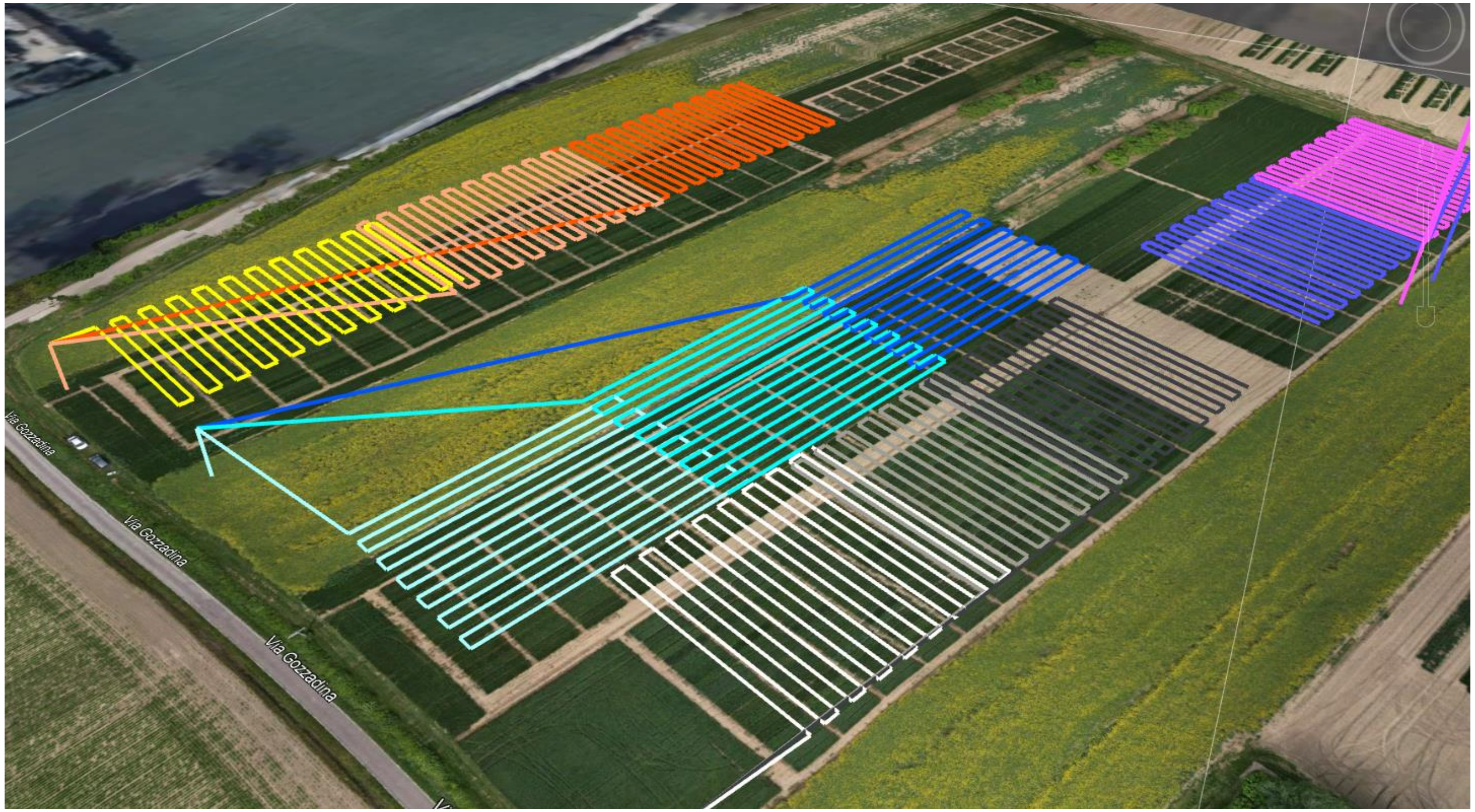
- The surveyed area (**1.6 ha**) is composed of **941 plots** (60% durum, 40% common). The plot dimension is **1.2 m x 7.5 m**, while the canopy height is 0.9 m.



Flight altitude [m]	Photo capture interval [Hz]	Speed [m/s]	Sensor dimension
10	0.5	1	1" (13.1 x 8.8 mm)

Survey planning and execution

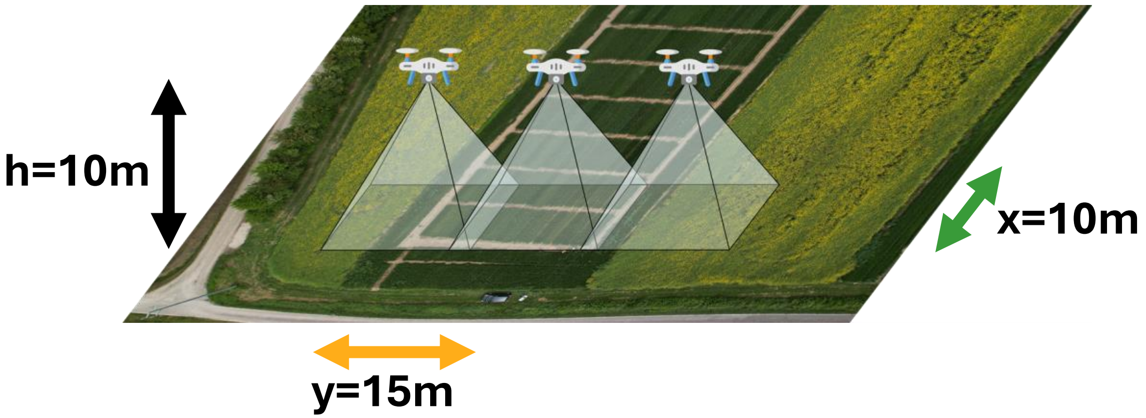
- An **altitude** of **10 m** was chosen to avoid disturbances to the wheat canopy.
- Each survey required **11 flights** due to battery limits (~20-minute duration).
- Across **11 surveys**, a total of **48 675 images** (649 GB) were collected.

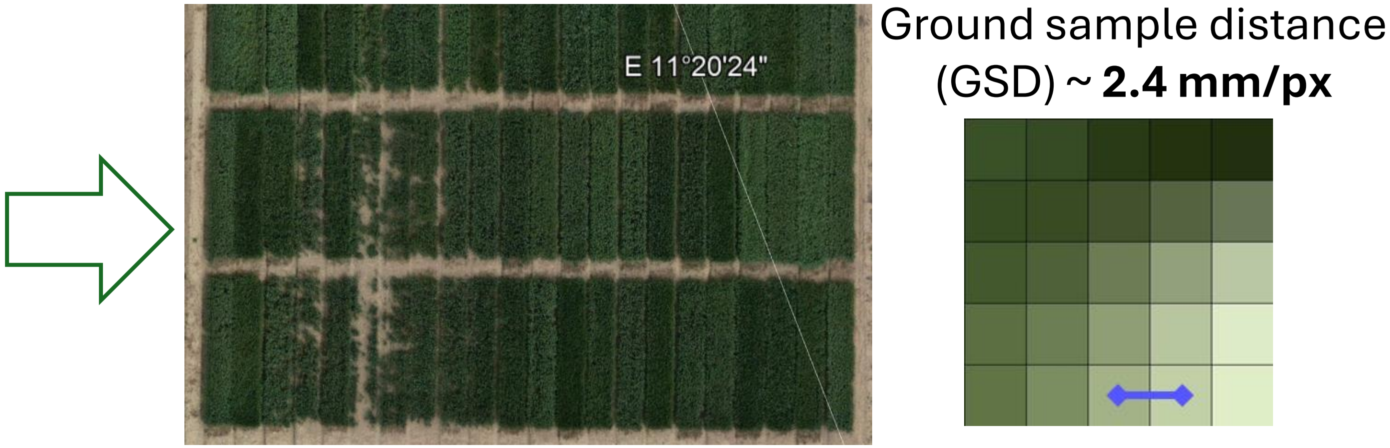


Survey	S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
Date	29/04	06/05	09/05	13/05	19/05	22/05	27/05	04/06	07/06	10/06	14/06	20/06

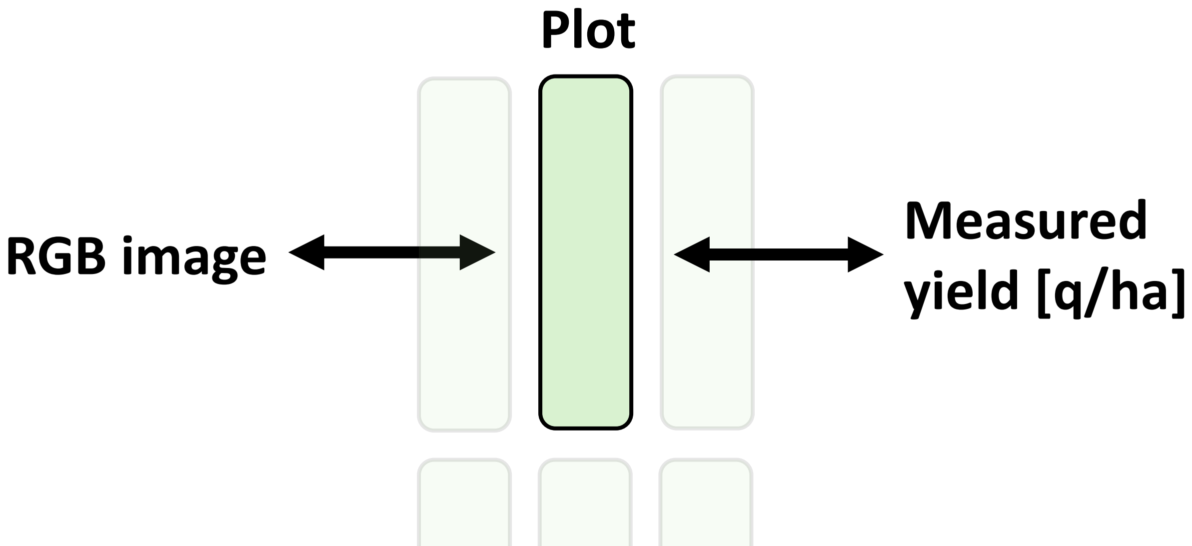
Why georeferenced orthomosaics?

An orthomosaic is a geometrically corrected aerial image obtained by merging several overlapping frames. It provides a georeferenced overview image of the entire field.



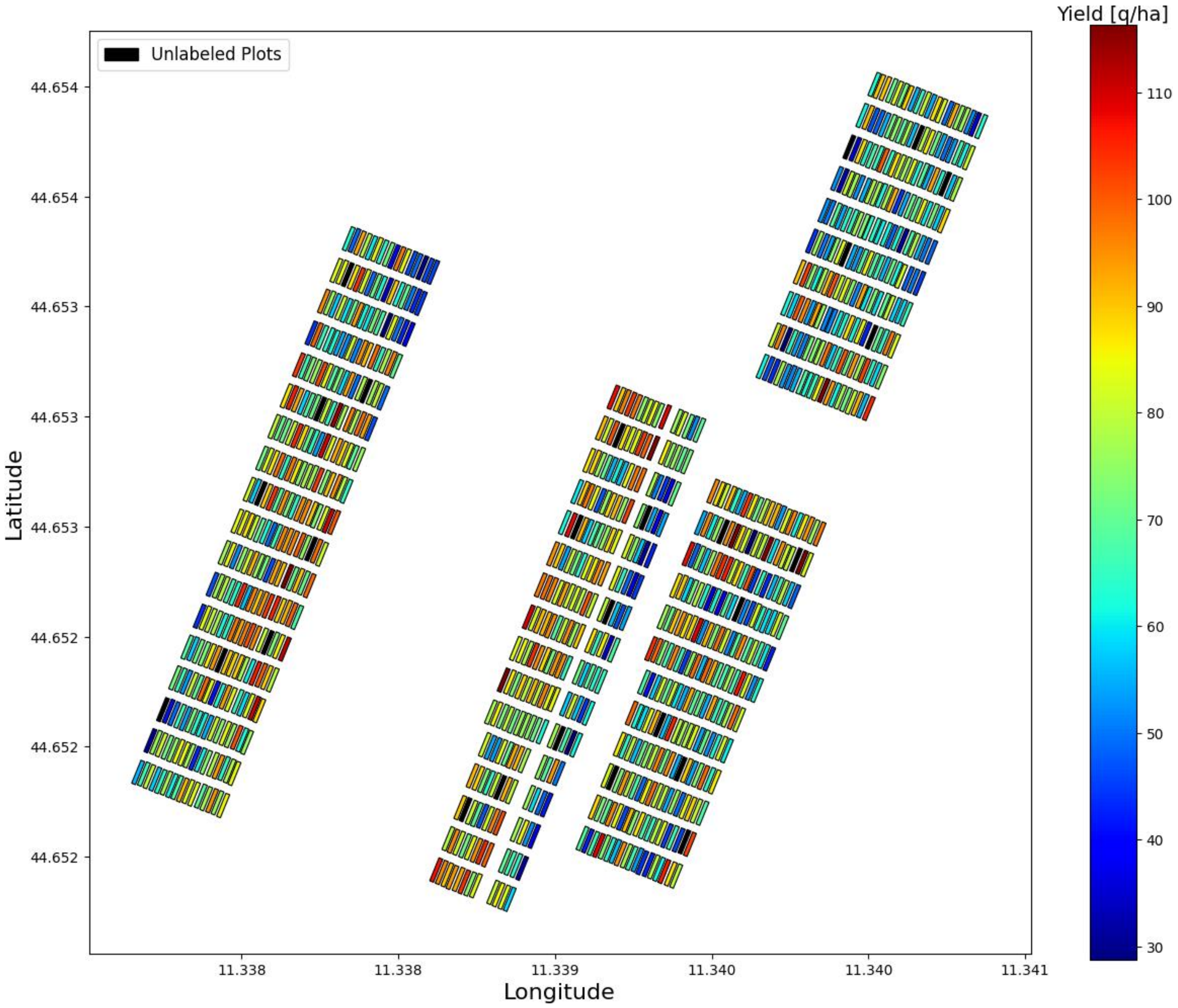


- An orthomosaic is build for each survey, for a total of **11 orthomosaics (194 GB)**
- Individual **plots** are **georeferenced**, enabling precise alignment of RGB data with measured yield

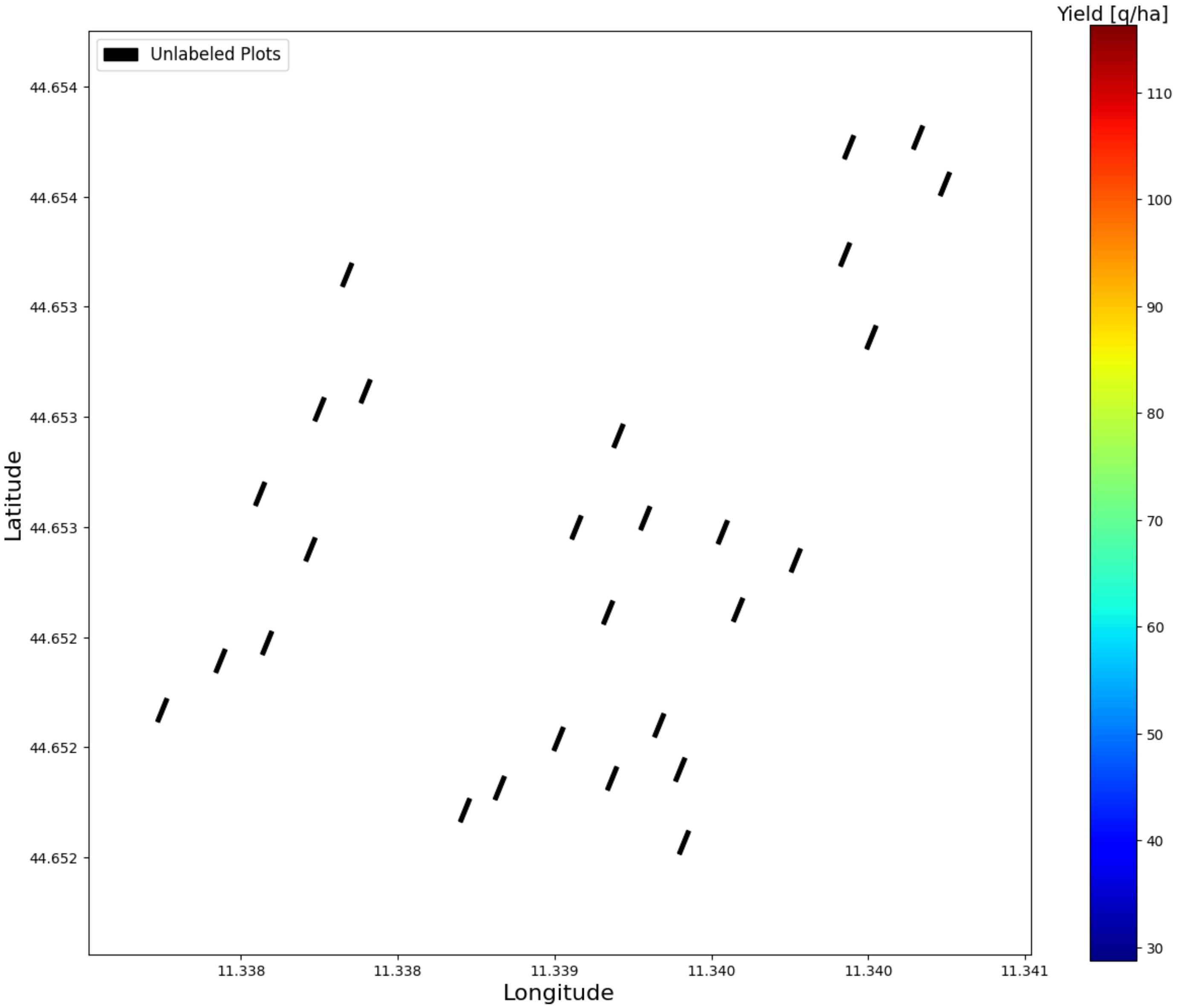



C96 06/05/24

Measured yield spatial distribution



Measured yield spatial distribution



- For **27/941** plots, **no measured yield** was given
- They were arbitrarily selected by the company for blind testing
- The **DNN** was **trained** upon the RGB images of the plots with measured yield, and used to **predict** the **yield** for the 27 plots

Image DB for DNN training

Survey	Data
S0	29/04
S1	06/05
S2	09/05
S3	13/05
S4	19/05
S5	22/05
S6	27/05
S7	04/06
S8	07/06
S9	10/06
S10	14/06
S11	20/06

Original

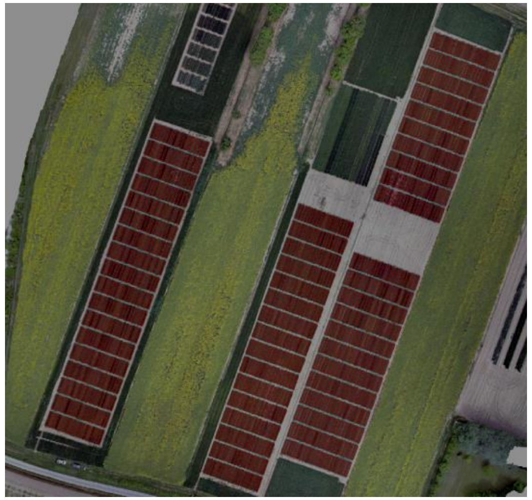
Rotated 90°

Rotated -90°

Horizontal flip

Vertical Flip

$$914 \times 11 \times 7 \times 5 = \mathbf{351\,890}\ (13\ \text{GB})$$



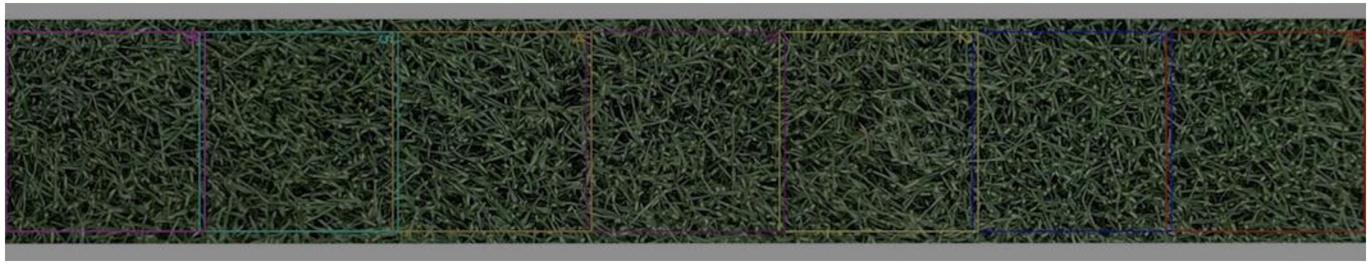
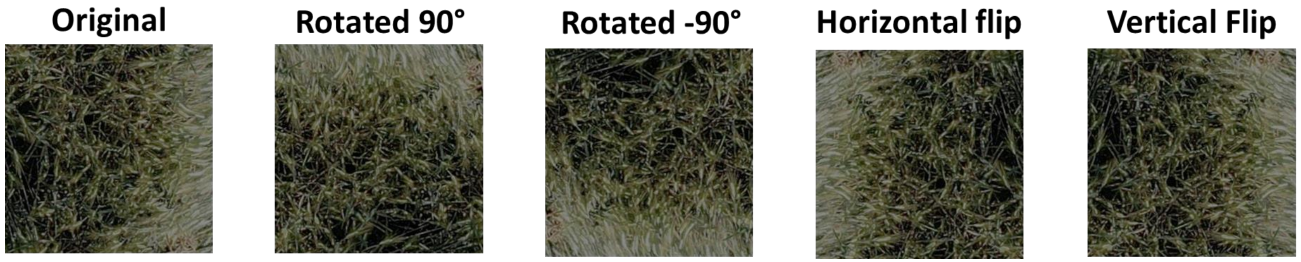


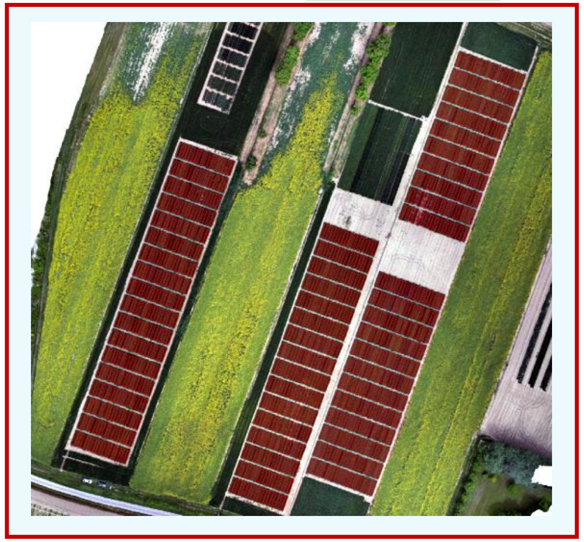
Image DB for DNN training

Rectangular masks are applied to orthomosaics to extract plot images…

Survey	Data
S0	29/04
S1	06/05
S2	09/05
S3	13/05
S4	19/05
S5	22/05
S6	27/05
S7	04/06
S8	07/06
S9	10/06
S10	14/06
S11	20/06



914 x 11 x 7 x 5 = 351 890 (13 GB)



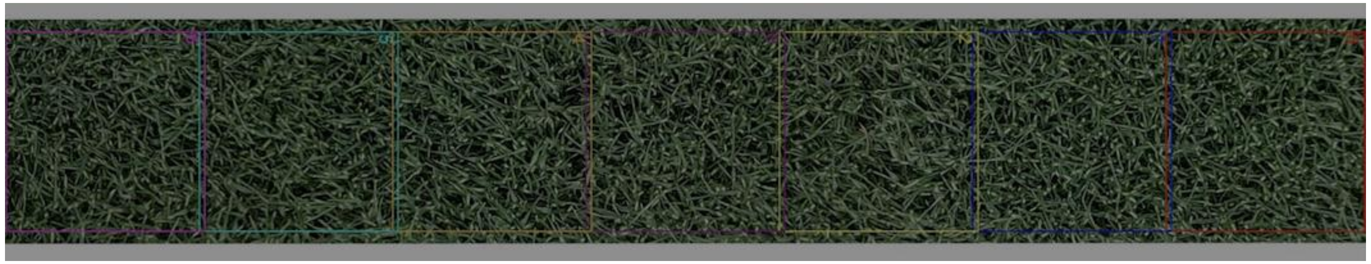
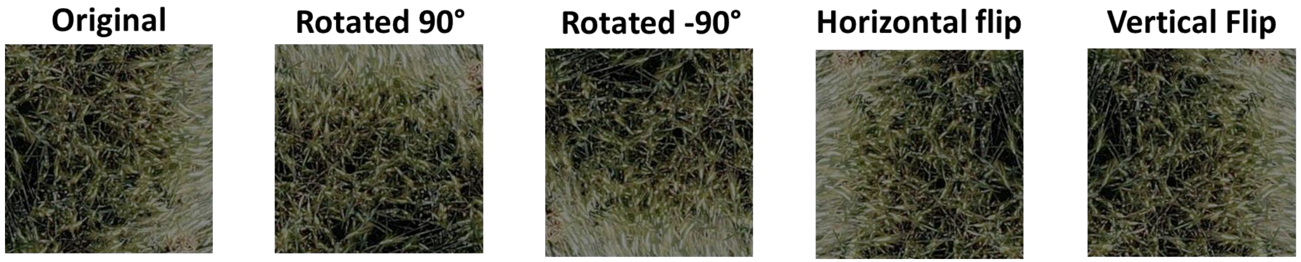


Image DB for DNN training

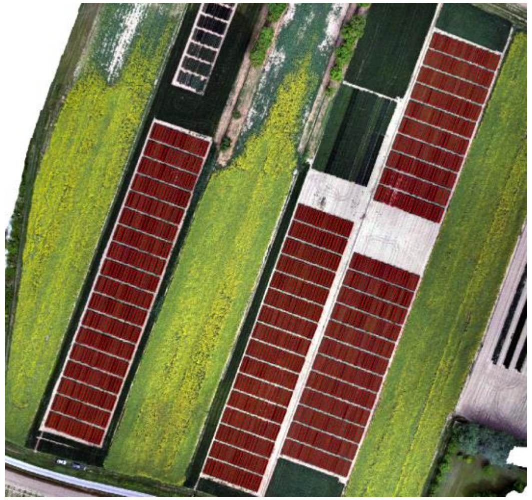
Rectangular masks are applied to orthomosaics to extract plot images…

… across 11 surveys

Survey	Data
S0	29/04
S1	06/05
S2	09/05
S3	13/05
S4	19/05
S5	22/05
S6	27/05
S7	04/06
S8	07/06
S9	10/06
S10	14/06
S11	20/06



$$914 \times \mathbf{11} \times 7 \times 5 = 351\,890\ (13\ \text{GB})$$



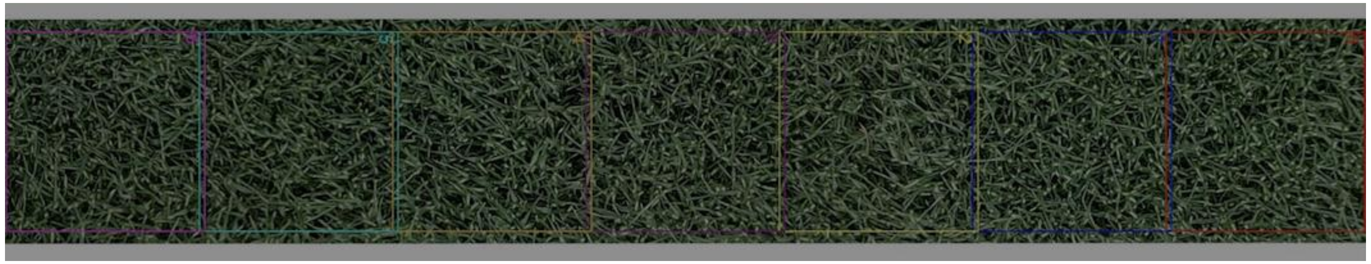
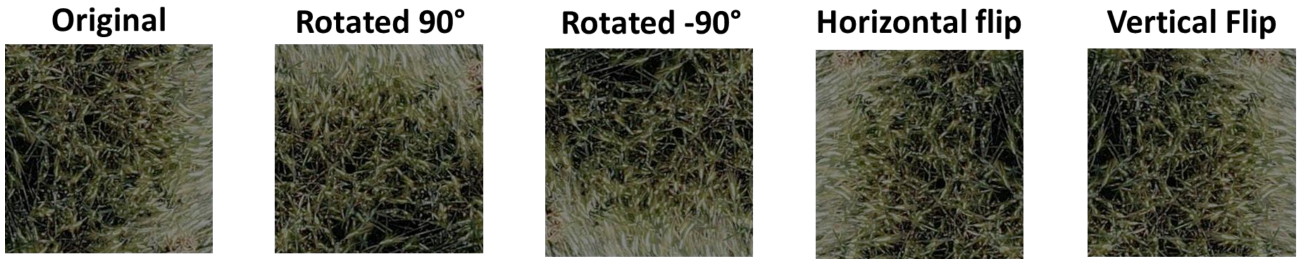


Image DB for DNN training

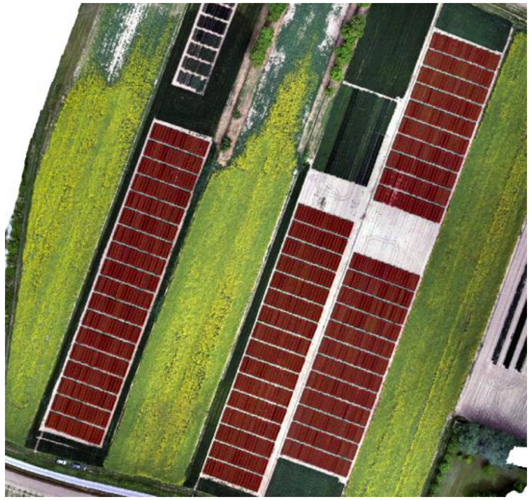
Rectangular masks are applied to orthomosaics to extract plot images…

… across 11 surveys

Survey	Data
S0	29/04
S1	06/05
S2	09/05
S3	13/05
S4	19/05
S5	22/05
S6	27/05
S7	04/06
S8	07/06
S9	10/06
S10	14/06
S11	20/06



$$914 \times 11 \times \mathbf{7} \times 5 = 351\,890 \text{ (13 GB)}$$



From each plot, 7 square images (**380 px** x **380 px**) were extracted

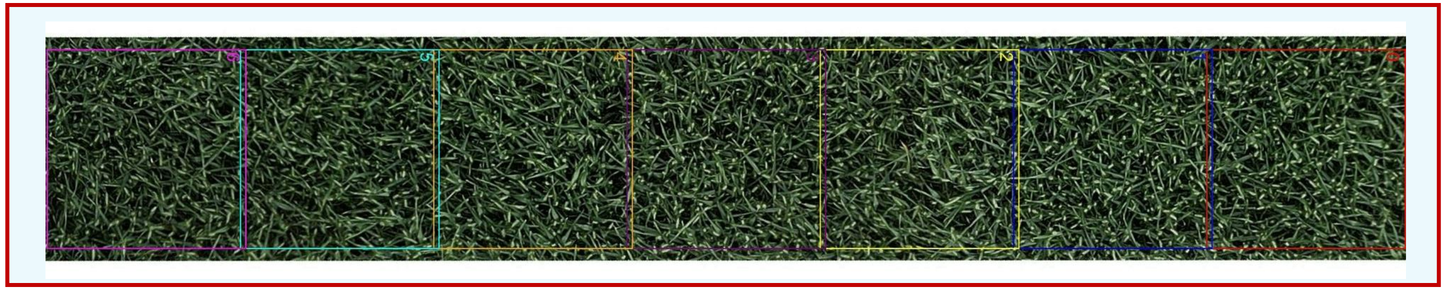


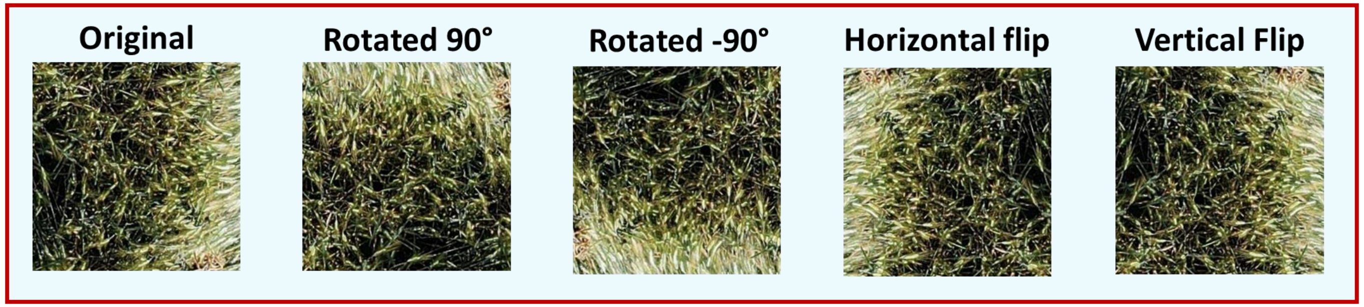
Image DB for DNN training

Rectangular masks are applied to orthomosaics to extract plot images…

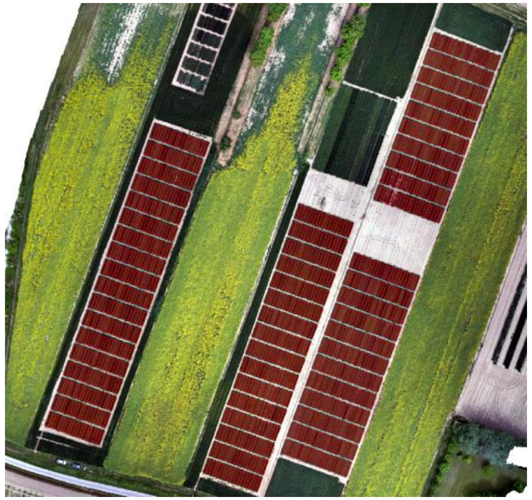
… across 11 surveys

Survey	Data
S0	29/04
S1	06/05
S2	09/05
S3	13/05
S4	19/05
S5	22/05
S6	27/05
S7	04/06
S8	07/06
S9	10/06
S10	14/06
S11	20/06

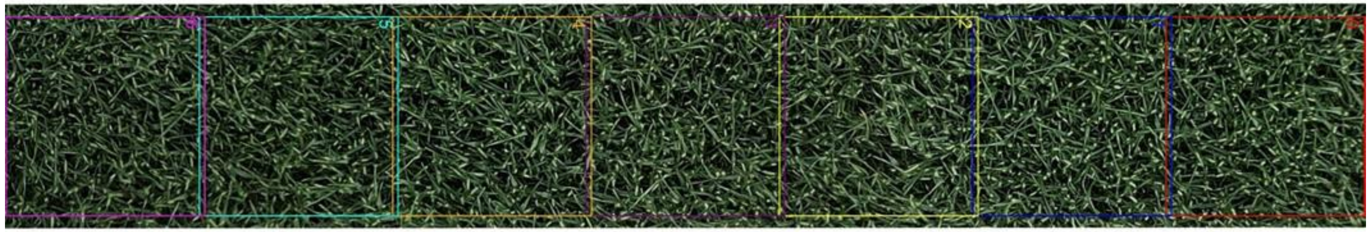
Augmentation is a process that manipulates (i.e., rotation, flip) the images to avoid biases related to the nature of the acquisition.



914 x 11 x 7 x **5** = 351 890 (13 GB)

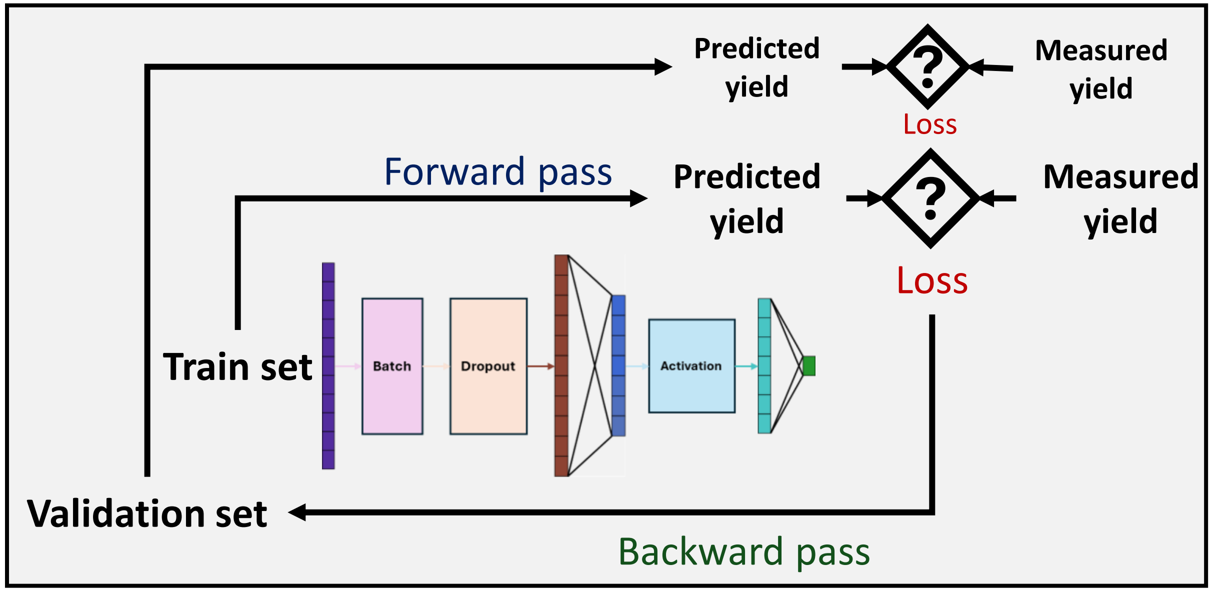


From each plot, 7 square images (**380 px** x **380 px**) were extracted



DB organization for FNN training

Set	Number of plots	Number of images [10^3]
Train (70%)	3200	247
Validation (20%)	915	70
Test (10%)	455	35
Complete DB	4570	352



(Epoch: cycle in which the FNN processes the train set and validation set.)

- **Train Set**: plots used to train the model to learn relationships between plots and yield.
- **Validation Set**: separate from the train set, used to monitor performance during training.
- **Test Set**: plots not used in training, employed to evaluate the model's generalization to new data.

Loss function

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

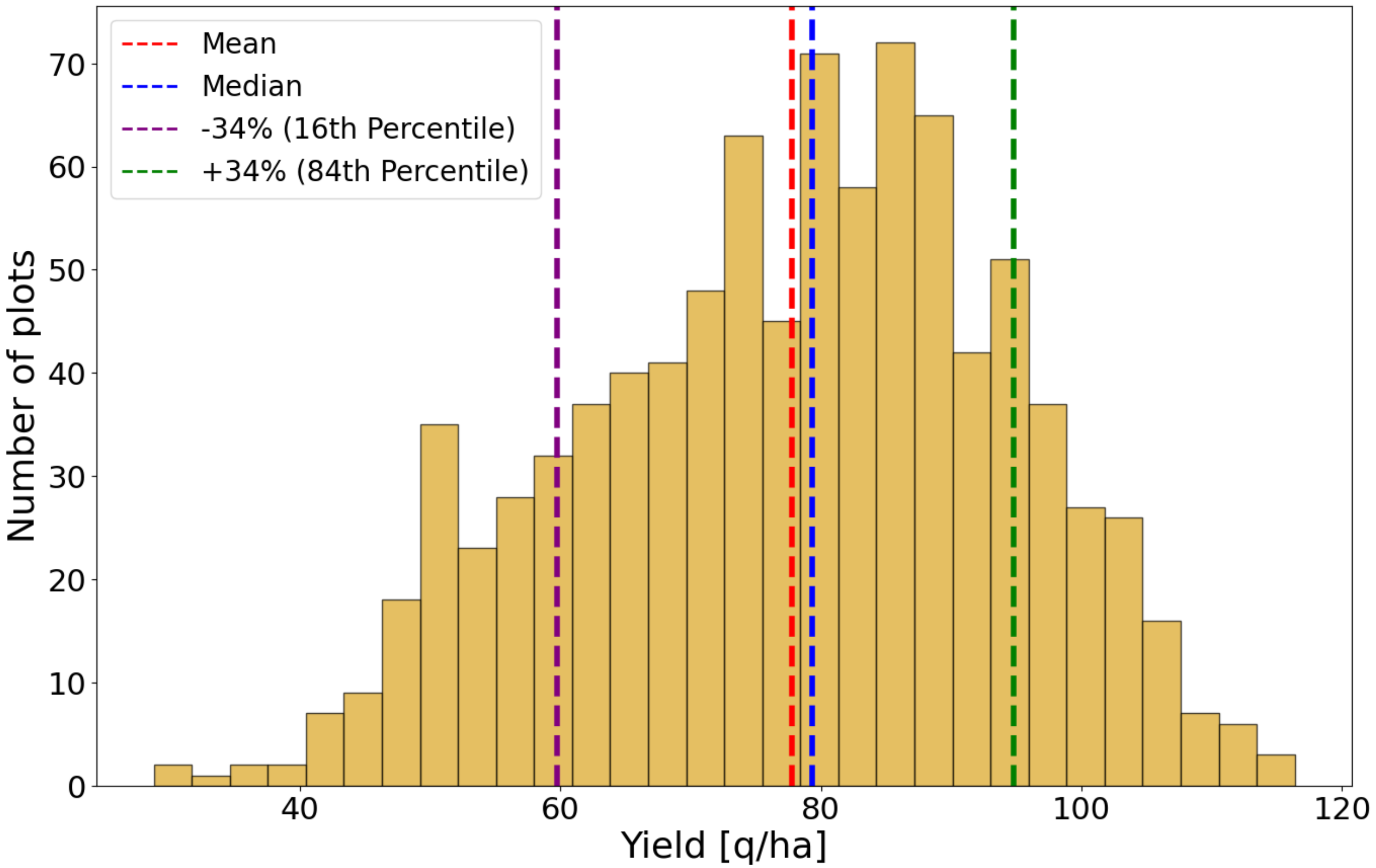
Plot yield statistics

15/07
Measured yield for 914/941 plots are given

15/07 – 23/10
FNN training on 914/941 plots

23/10
Blind yield prediction on 27/941 plots and measured yield unveiling

Mean [q/ha]	77.8
Standard deviation [q/ha]	16.4
Min [q/ha]	28.8
Max [q/ha]	116.4
Median [q/ha]	79.3
16th Percentile [q/ha]	59.7
84th Percentile [q/ha]	94.8
Percentile distance [q/ha]	[-19.6 ; +15.5]
Skewness	-0.2



Yield prediction strategy

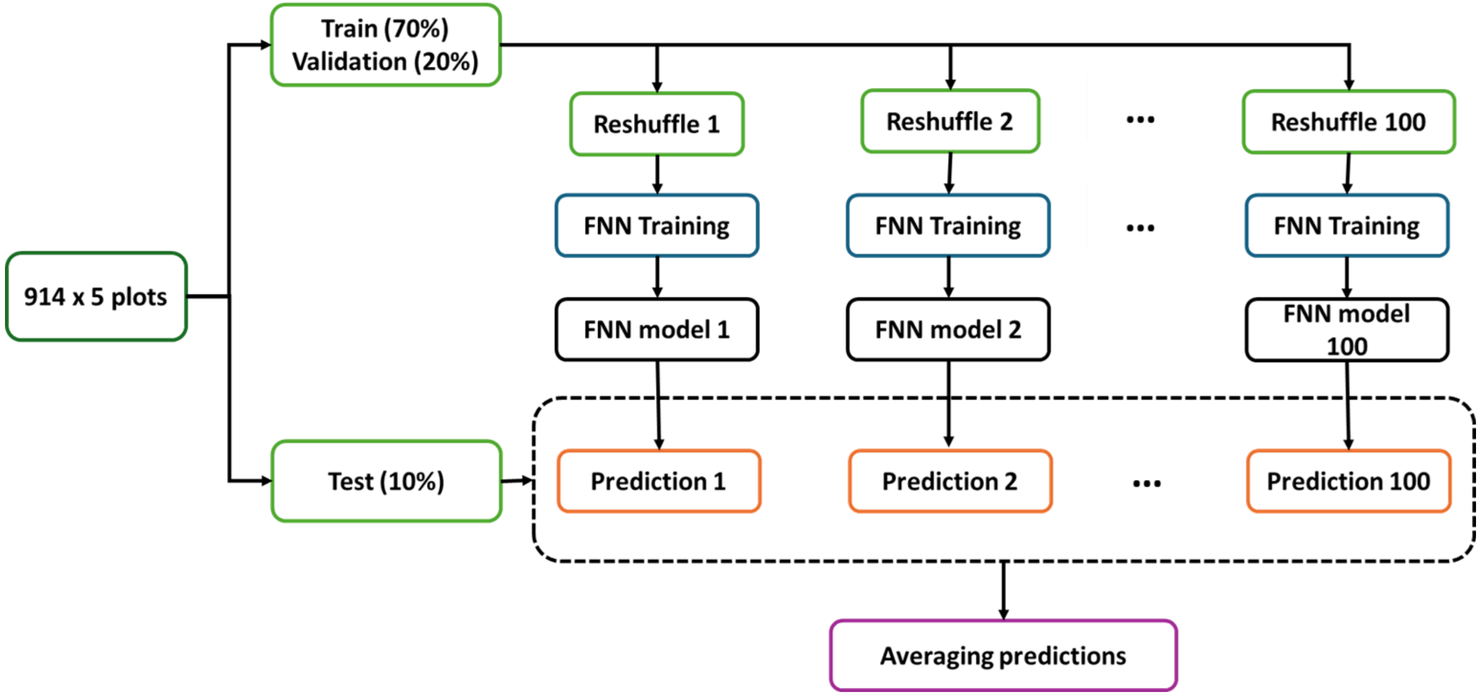
15/07
Measured yield for 914/941 plots are given

15/07 – 23/10
FNN training on 914/941 plots

23/10
Blind yield prediction on 27/941 plots and measured yield unveiling

1. The FNN is trained **100** times from scratch, with **train and validation sets randomly reshuffled** for each iteration

2. Each trained FNN model predicts yields for the **test set**, and the **final prediction** is obtained by **averaging** results from all 100 models

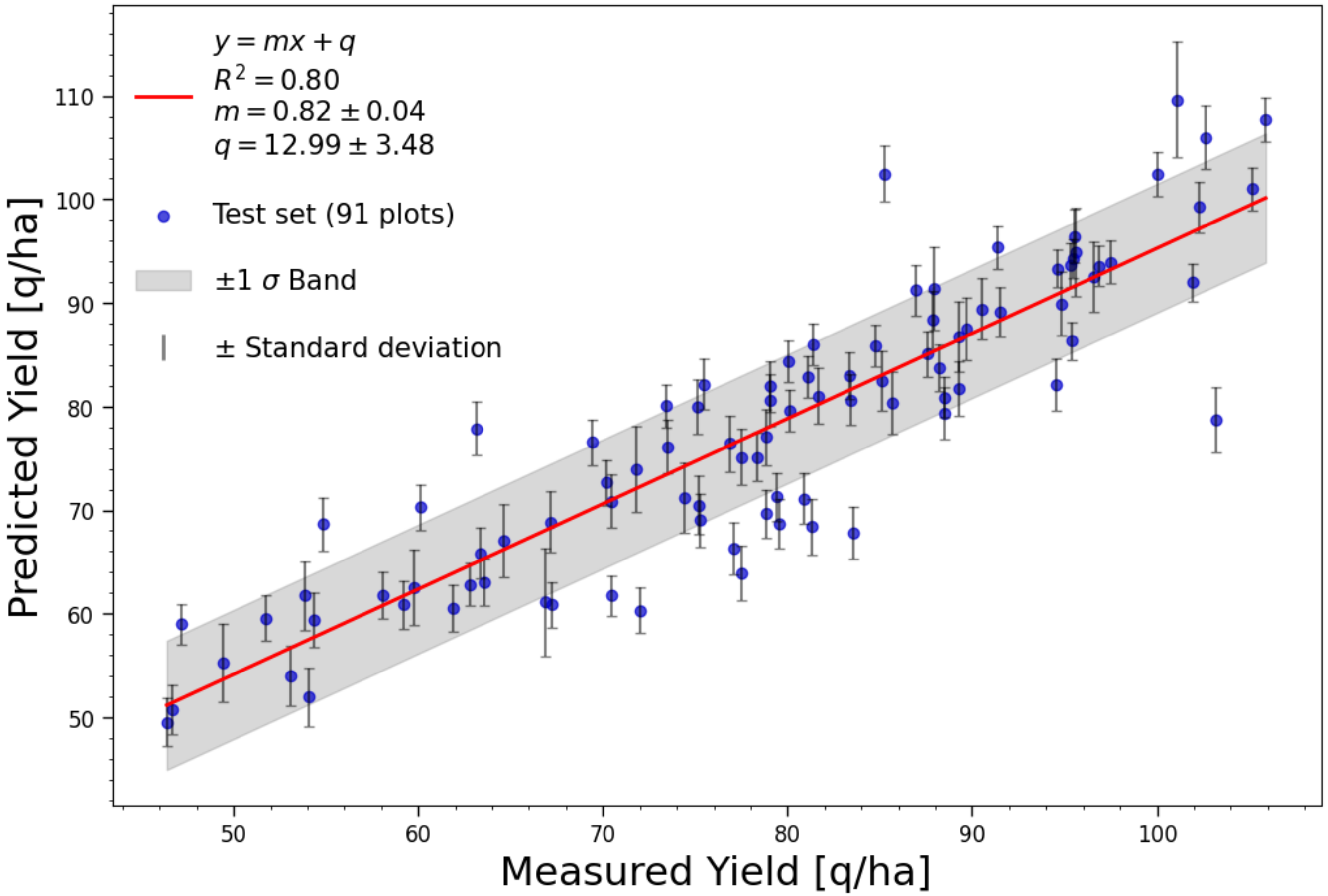


Yield prediction on 91 test plots

15/07
Measured yield for 914/941 plots are given

15/07 – 23/10
FNN training on 914/941 plots

23/10
Blind yield prediction on 27/941 plots and measured yield unveiling



$$\mathrm{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - \widehat{y_i}}{y_i} \right|$$

MAPE [%]	
Test (91 plots)	Blind test (27 plots)
6.9 ± 0.6	**???**

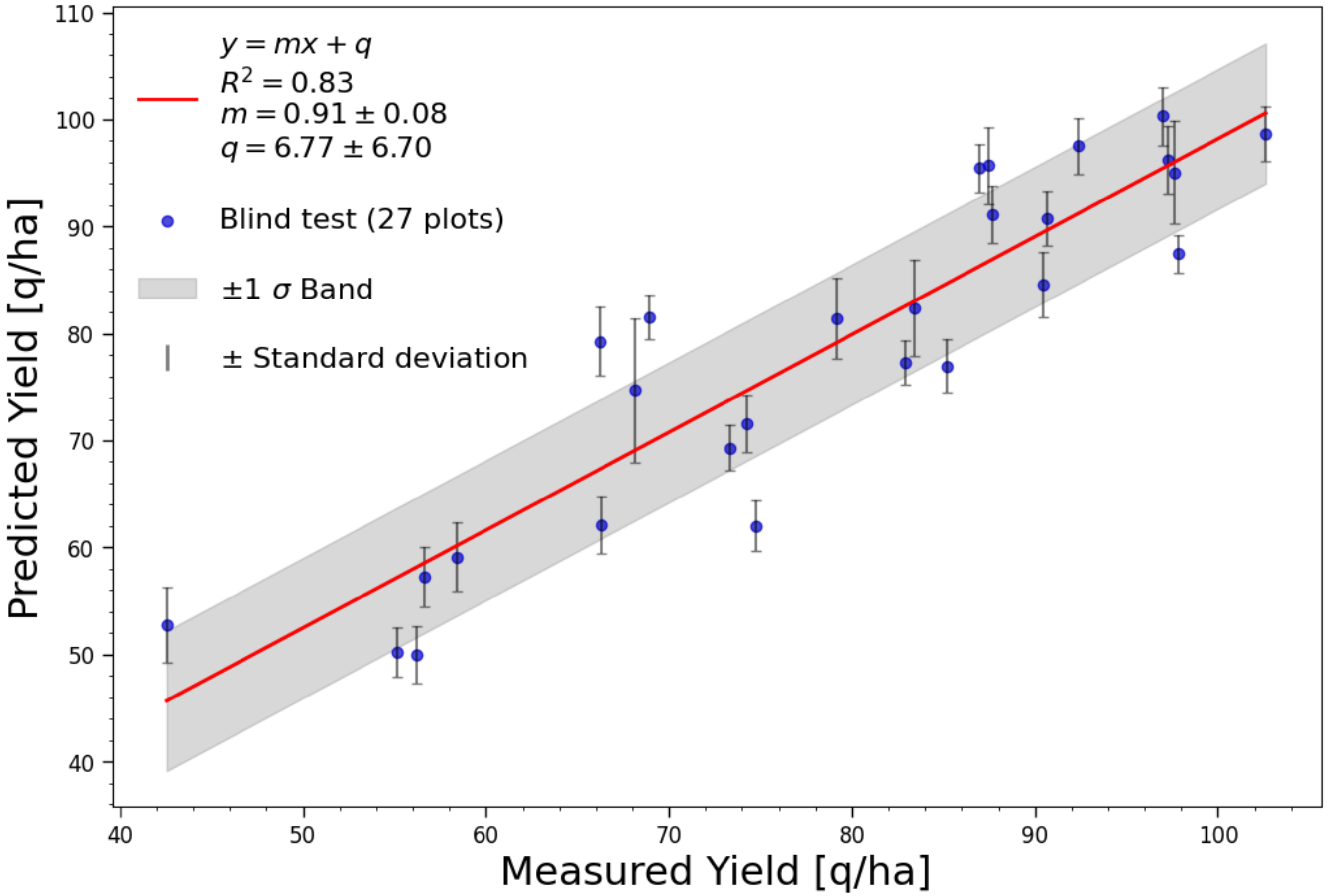
Expectation: it should mirror the predictive accuracy on the **27 (out of 914) plots** used for blind testing

Blind testing on 27 unknown plots

15/07
Measured yield for 914/941 plots are given

15/07 – 23/10
FNN training on 914/941 plots

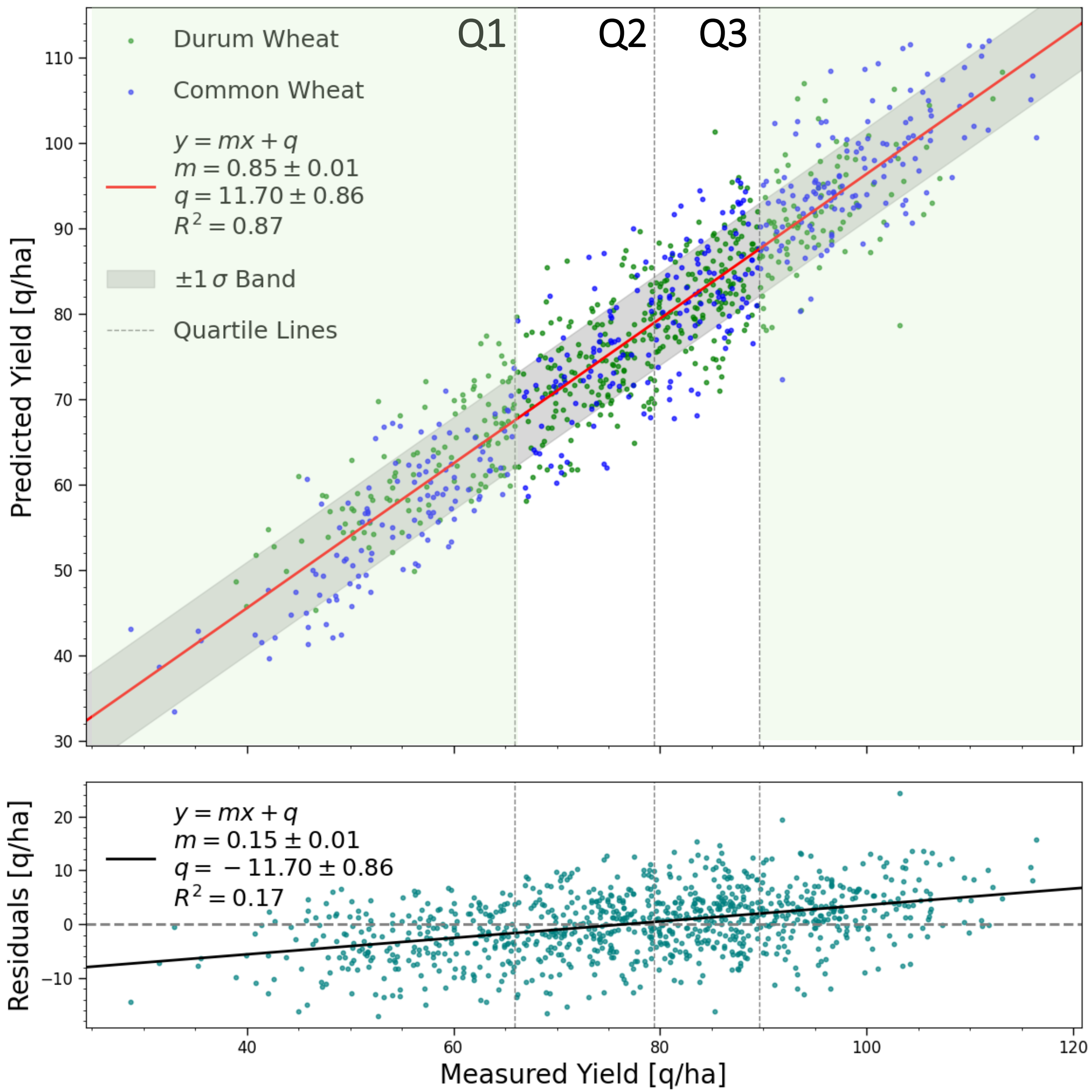
23/10
Blind yield prediction on 27/941 plots and measured yield unveiling



MAPE [%]	
Test (91 plots)	**Blind test (27 plots)**
6.9 ± 0.6	**7.5 ± 1.2**

- **MAPE**: increased by 15% relative to the test
- Good predictive accuracy with R^2=0.83
- **m**=0.91±0.08 and **q**=6.77±6.70 close to 1 and 0, but not robust statistically.

Yield prediction across the entire field



MAPE [%]		
Test (91 plots)	Blind test (27 plots)	**Entire field (941 plots)**
6.9 ± 0.6	7.5 ± 1.2	**6.5** ± 0.2

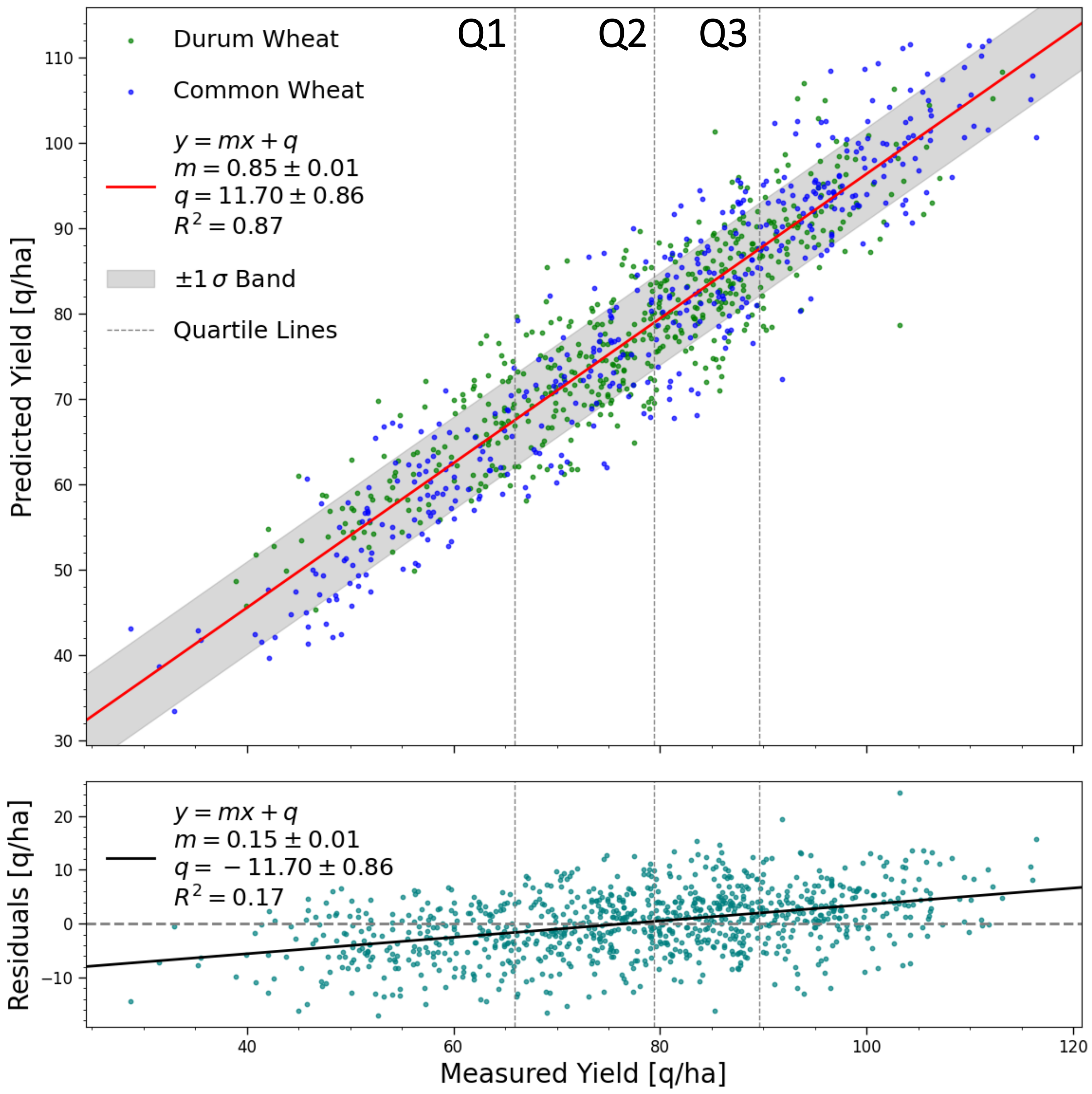
Residual analysis shows **underestimation and overestimation** of low and high yields

$$\text{Residual} = \text{Measured yield} - \text{Predicted yield}$$

Mean residual < Q1
-3.3 ± 5.2 q/ha

Mean residual > Q3
3.0 ± 5.6 q/ha

Yield prediction across the entire field



MAPE [%]		
Test (91 plots)	Blind test (27 plots)	**Entire field (941 plots)**
6.9 ± 0.6	7.5 ± 1.2	**6.5** ± 0.2

Residual analysis shows **underestimation and overestimation** of low and high yields

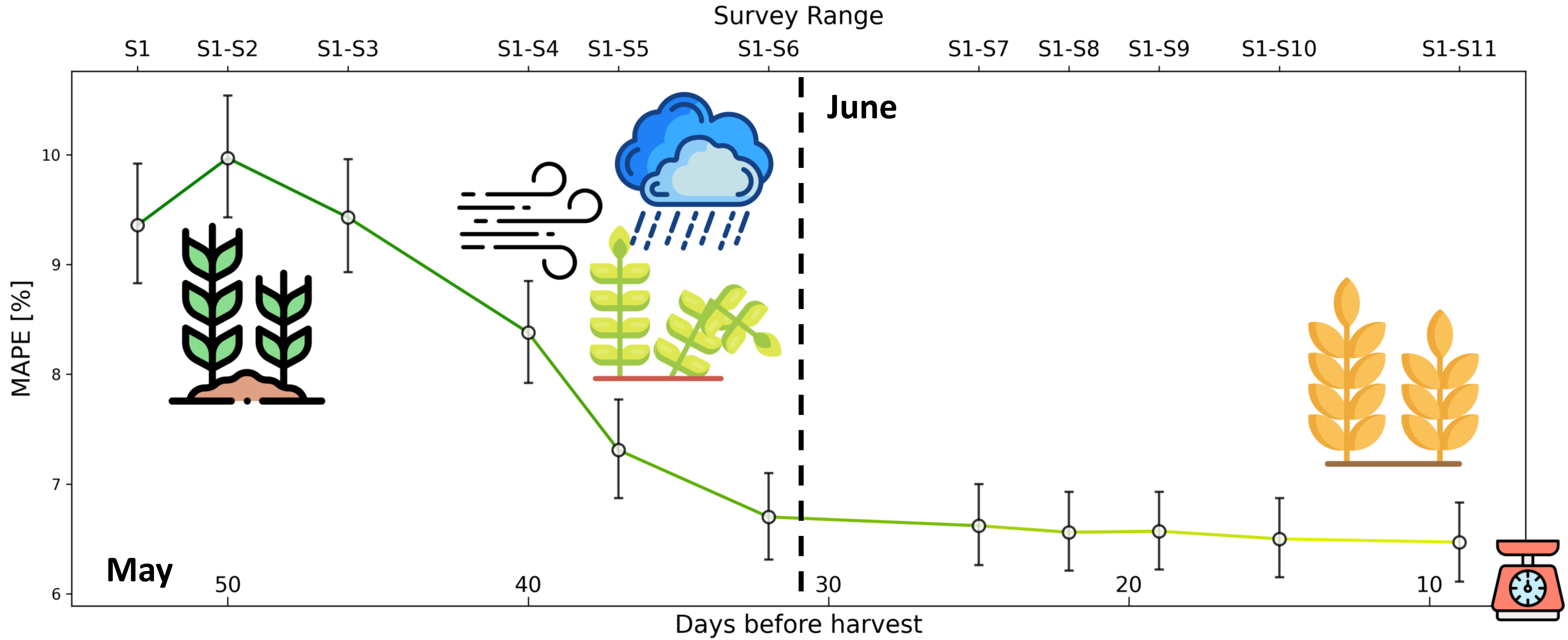
$$\text{Residual} = \text{Measured yield} - \text{Predicted yield}$$

Mean residual < Q1
-3.3 ± 5.2 q/ha

Mean residual > Q3
3.0 ± 5.6 q/ha

Despite a weak systematic bias, systematic discrepancy at high and low yields were corrected using linear, quadratic, and cubic functions, without significant improvement.

Are surveys from all growth stages necessary?



In **early May**, plots exhibit high uniformity in phenotypic traits.

In **late May**, phenotypic variability increases due to environmental stress

In **June**, the field structure stabilizes with minimal changes.

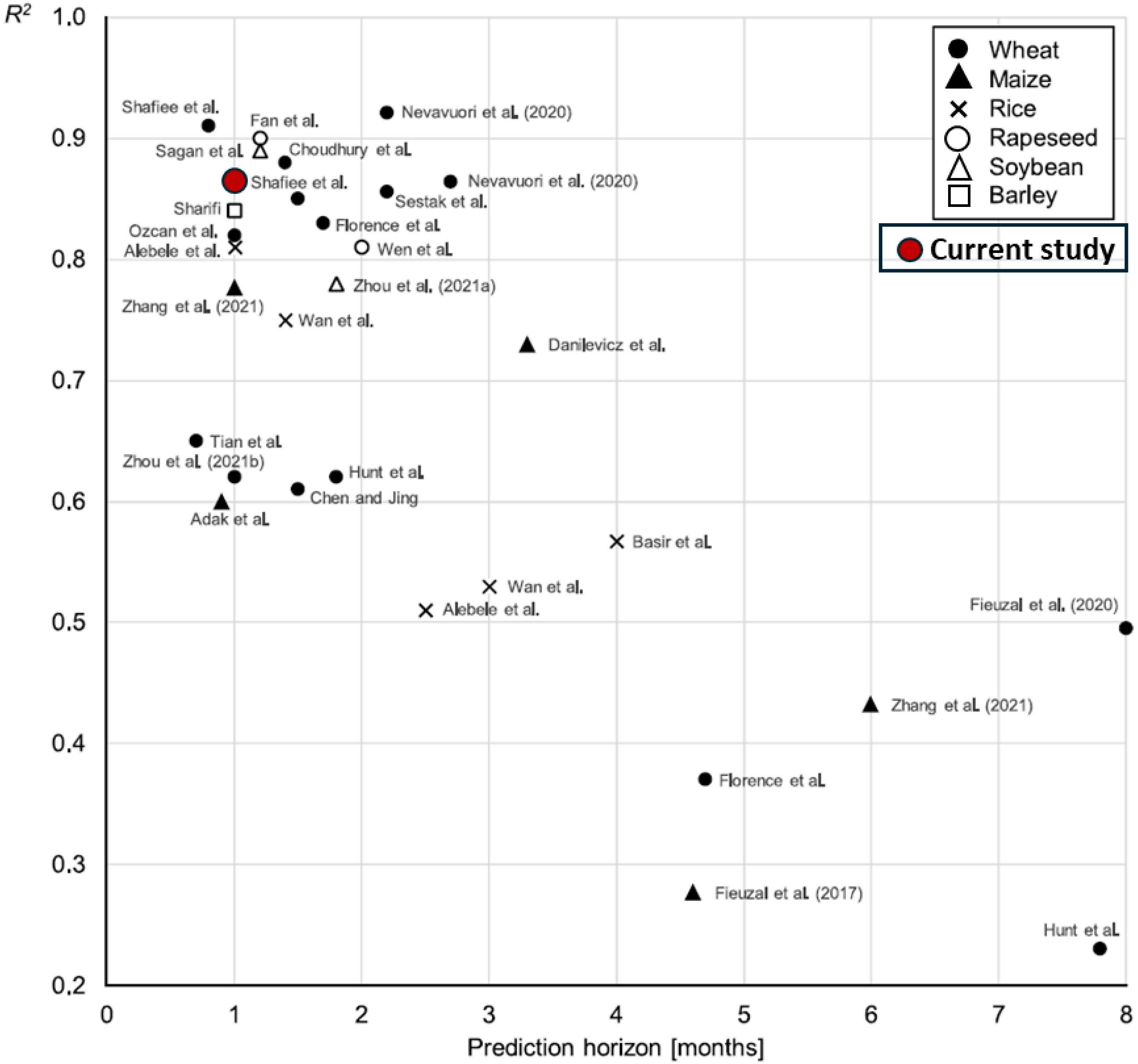
Conclusions

- Precise UAV survey planning combined with orthomosaic georeferencing ensured to extract automatically **350 000 RGB images** for DNN training.

- Yield prediction on 27 blind test plots achieved a **MAPE = 7.5%** and a **R^2 = 0.87** without any spatial and varietal bias.

- The inclusion of surveys across the **entire wheat growth cycle** enabled the DNN to account for **environmental stress** and variations in **crop phenotype**.

- Despite this study is the only one leveraging only RGB images, the performance in terms of **R^2** is comparable to the best results reported in the literature.



Thank you for your attention!