

## Computer vision for plant phenotyping: the Polyploidbreeding 4.0 project and the AGRI-VISION initiative

#### Filippo Biscarini (And the Polyploidbreeding Core Team)

CNR-IBBA, Milan, Italy













#### Overview

- 1. Neural Network models for **Computer Vision**:
  - CNN
  - Self-attention and transformers vs RNN
- 2. The Polyploidbreeding project
  - drone phenotyping
  - rhizotron
  - genotyping & sequencing
- 3. drone2report (software) + vegetation indices
- 4. **AGRI-VISION**: a proposal for a new COST Action (submitted)





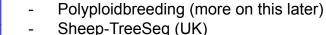






### Filippo in one slide

- Roma (born)
- Perugia (MSc degree)
- Cork, ICBF (Web-design & Database)
- Cremona, ANAFI (Quantitative Genetics)
- Guelph, CGIL (Visiting Scientist)
- Wageningen, WUR (PhD)
- Göttingen University (post-doctoral researcher)
- Lodi, PTP ('omics in animals, plants, humans)
- Caldes de Montbui, IRTA (Visiting Scientist)
- Milan CNR (tenured researcher)
- Cardiff University (biostatistician)
- Milan CNR (senior researcher)
- Bruxelles ERC (seconded national expert)
- Milan CNR (senior researcher) -



Currently coordinating:

DeepMicroCore (Serbia)

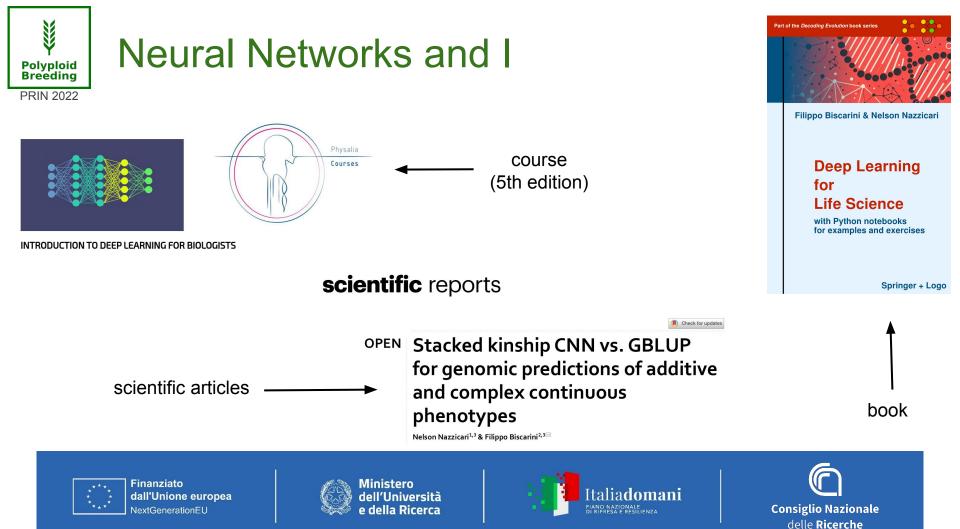


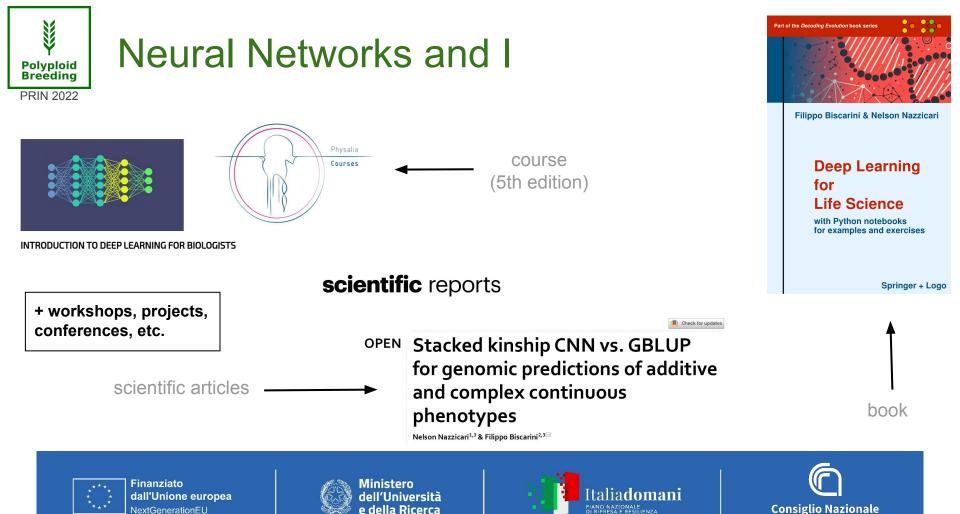












delle Ricerche



#### Part 1:

#### Neural Network models for Computer Vision









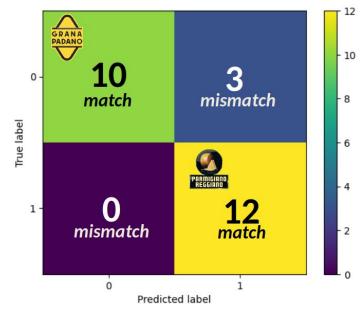


### **Computer Vision**

Automatic reading, processing, analysing and "understanding" of digital images by computers:

- image recognition (classification)

Largely based on CNN















### **Computer Vision**

Automatic reading, processing, analysing and "understanding" of digital images by computers:

image recognition (classification)

Largely based on CNN

- confocal microscopy images of GP and PR
- N = 130 images (69 GP + 61 PR)
- CNN model (85,089 parameters, 200 epochs)
- 25 validation images (13 GP + 12 PR)
- validation accuracy = 0.84 (but highly variable!)

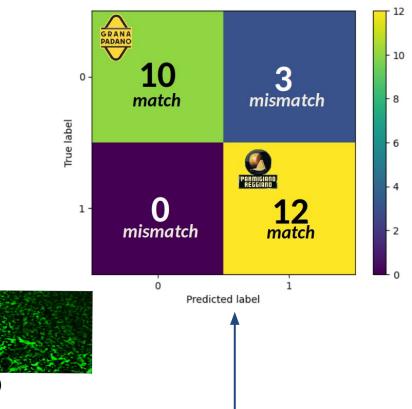






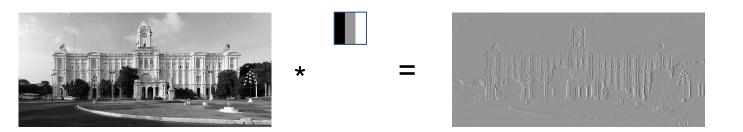








#### **CNN: Cable News Network**



#### **CNN: Convolutional Neural Networks**



Finanziato dall'Unione europea NextGenerationEU



dell'Università e della Ricerca



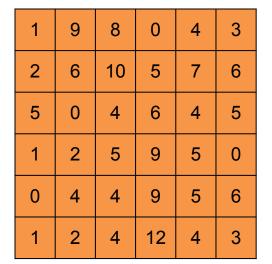


# CNN: how the computer sees an image

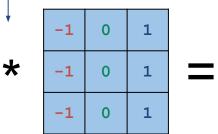
Breeding PRIN 2022

\*

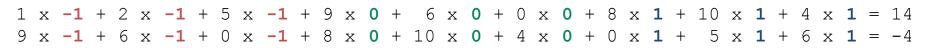
Polyploid



Hadamard product



14	-4	-7	3





. . .



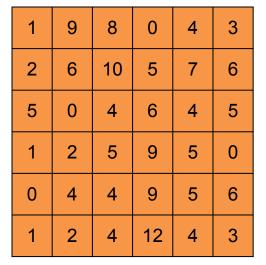




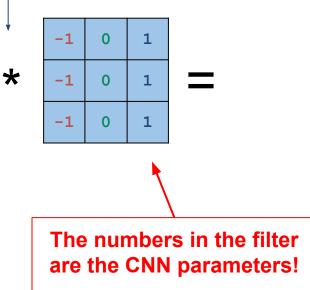
#### \* Polyploid Breeding

### CNN: how the computer sees an image

**PRIN 2022** 



Hadamard product



14	-4	-7	3



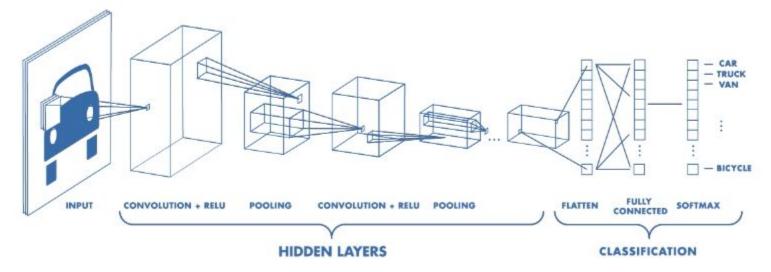








### CNN: image recognition



Source: https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html









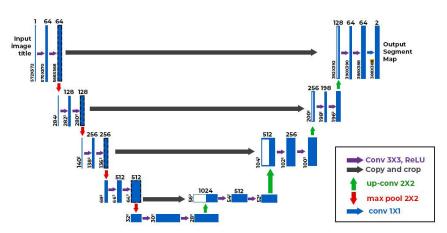


### **CNN: image segmentation**



https://indiaai.gov.in/article/image-segmentation-the-deep-learning-approach

**U-NET** 



https://www.geeksforgeeks.org/u-net-architecture-explained/



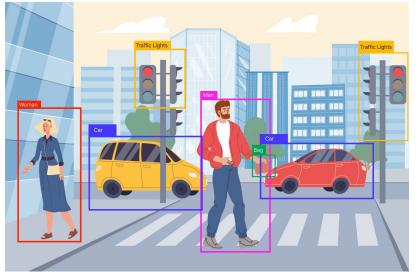






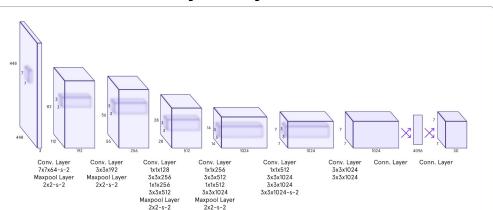


#### **CNN:** object detection



https://deeplobe.ai/exploring-object-detection-applications-and-benefits/

#### YOLO: you only look once



The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1x1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224x224 input image) and then double the resolution for detection.

#### https://pareto.ai/blog/yolo-object-detection



Finanziato dall'Unione europea **NextGenerationEU** 

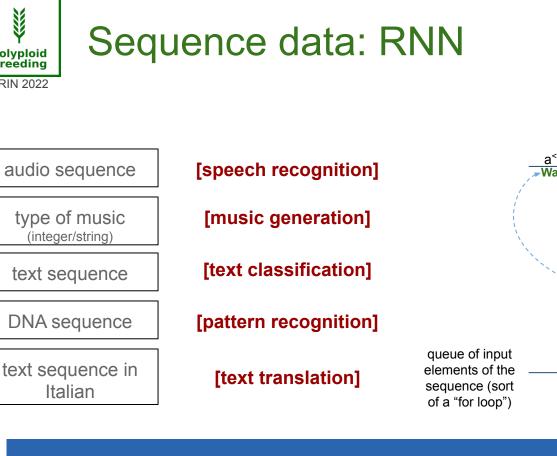


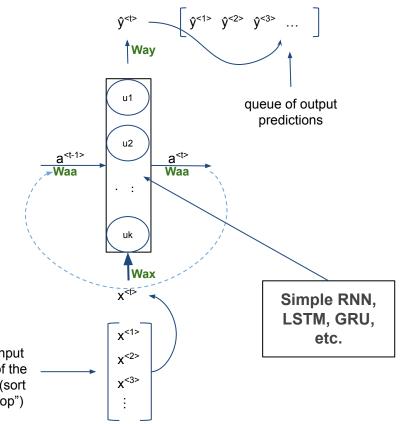
dell'Università e della Ricerca











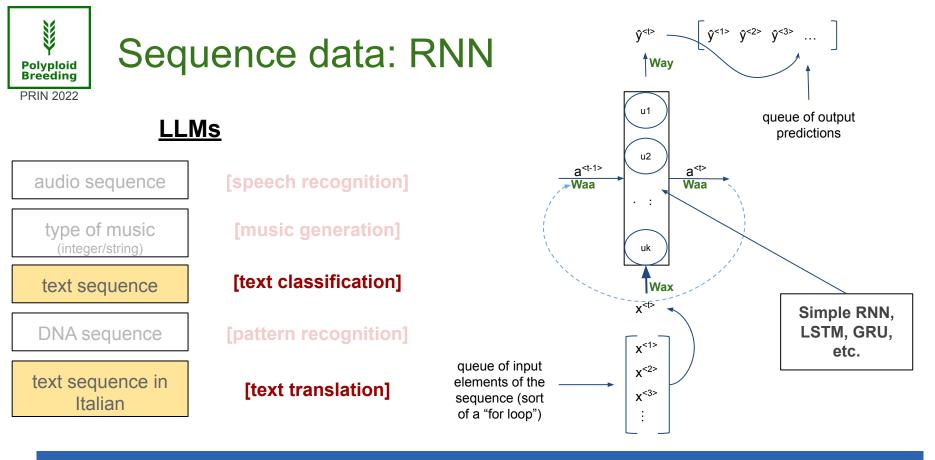
























#### Self-attention

#### **Attention Is All You Need**

Ashish Vaswani\* Google Brain avaswani@google.com Noam Shazeer\* Google Brain noam@google.com

Niki Parmar\* Google Research nikip@google.com

Jakob Uszkoreit\* Google Research usz@google.com

Llion Jones\* Google Research llion@google.com Aidan N. Gomez<sup>\* †</sup> University of Toronto aidan@cs.toronto.edu Łukasz Kaiser\* Google Brain lukaszkaiser@google.com

Illia Polosukhin\* <sup>‡</sup> illia.polosukhin@gmail.com

12/06/2017











#### Transformers

- (self)attention: ability of the model to automatically, dynamically and independently highlight and use the salient parts of the input data
- attention is the engine of transformer models
- RNN, LSTM, GRU etc.: sequential calculations, no parallelization possible (severe computational limit)
- transformers capture long-range dependencies in the data and at the same time are amenable to parallelization
- transformers are a new network architecture that dispenses with recurrence and convolutions entirely (no CNN, no RNN)
- transformers are successfully <u>applied also to image data</u> (computer vision)

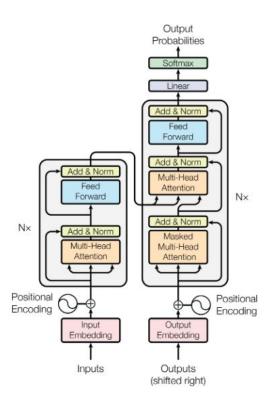


Figure 1: The Transformer - model architecture.













#### Were RNNs All We Needed?

Leo Feng Mila – Université de Montréal & Borealis AI leo.feng@mila.quebec Frederick Tung Borealis AI frederick.tung@borealisai.com

#### **Mohamed Osama Ahmed**

Borealis AI mohamed.o.ahmed@borealisai.com

Yoshua Bengio Mila – Université de Montréal yoshua.bengio@mila.quebec

#### Hossein Hajimirsadeghi

Borealis AI hossein.hajimirsadeghi@borealisai.com

04/10/2024











#### Part 2:

### The Polyploidbreeding project











### Polyploidbreeding

"Expanding the toolbox for cereal breeding: high-throughput genomics, 2D-3D phenomics and artificial intelligence for breeding with increasing genome complexity, from **barley** to **durum** and **bread wheat**"

- PRIN (EU/ITA funds) research project
- 09/2024 09/2025
- ~€240k
- project website: https://polyploidbreeding.ibba.cnr.it/
- barley (diploid), durum wheat (tetraploid), bread wheat (hexaploid)





#### PRIN 2022 (Settore LS2)

- The project
- The people
  List of posts by category
- List of posts by tag
- Knowledge base

Start date: 28 September 2023

End date: 27 September 2025



Hosted on GitHub Pages — Theme by orderedlist

#### Z5 October 2023 Kick-off meeting: Bologna

by Filippo Biscarini

The In-presence kick-off meeting of the project **Polyploidbreeding** will take place in <u>Bologna on 4-5 December 2023</u>.

- social dinner, December 4th
- project meeting at S.I.S. headquarters on December 5th. Two
   sessions
  - morning: 9:00 13:00
  - afternoon: 14:00 16:00

#### Programme

1. Initial address from the Director General of S.I.S.

- 2. General presentation of the project
- Scientific lecture: "Genomics of polyploid wheats: durum (4x) and bread (6x)"
- 4. Discussion of ongoing activities:
  - field trials: barley, durum wheat, bread wheat
  - drone phenotyping
  - genotyping strategy
  - climate variables

 presentation of breeding programmes at S.I.S.
 presentation of project website and discussion on social media platforms:
 xerao-up discussion













## Polyploidbreeding - the team

- 1. CNR (National Research Council) (www.cnr.it): project coordinator
  - a. Filippo Biscarini (coordinator)
  - b. Giulia Moscatelli (post-doctoral researcher)
- 2. University of Bologna (<u>www.unibo.it</u>):
  - a. Elisabetta Frascaroli
  - b. Marco Maccaferri, Matteo Bozzoli + new trainee (to be appointed)
- 3. CREA (Council for Research in Agriculture) (www.crea.gov.it):
  - a. Nelson Nazzicari
  - b. Agostino Fricano
- 4. SIS (Società Italiana Sementi): breeding company (www.sisonweb.com/):
  - a. Eder Groli, Paolo De Franceschi
- 5. Forschungszentrum Jülich (<u>www.fz-juelich.de</u>):
  - a. Fabio Fiorani, Kerstin Nagel















## Polyploidbreeding - the team

- 1. CNR (National Research Council) (www.cnr.it): project coordinator
  - a. Filippo Biscarini (coordinator)
  - b. Giulia Moscatelli (post-doctoral researcher) -
- 2. University of Bologna (<u>www.unibo.it</u>):
  - a. Elisabetta Frascaroli
  - b. Marco Maccaferri, Matteo Bozzoli + new trainee (to be appointed)
- 3. **CREA** (<u>Council for Research in Agriculture</u>) (<u>www.crea.gov.it</u>):
  - a. Nelson Nazzicari
  - b. Agostino Fricano
- 4. **SIS** (Società Italiana Sementi): breeding company (www.sisonweb.com/):
  - a. Eder Groli, Paolo De Franceschi
- 5. **Forschungszentrum Jülich** (<u>www.fz-juelich.de</u>):
  - a. Fabio Fiorani, Kerstin Nagel



#### (vegetation indices)



(drone2report)





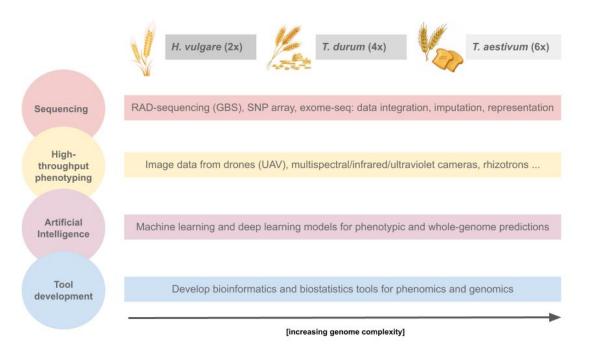
Ministero dell'Università e della Ricerca







### Polyploidbreeding - overview













### The experiments

- 1. <u>drone phenotyping experiment</u> (led by CNR, CREA): completed
  - a. 10 flights, 3 cameras, 3 crop species:
    - i. barley (Fiorenzuola, CREA): ~260 varieties
    - ii. durum and bread wheat (S. Lazzaro Bologna, SIS): ~247 (durum) and 230 (bread) varieties

Сгор	Number of individuals	Traits	Genotypic data	Connected project	Notes
Barley (H. vulgare)	403	DtH,PH,TKW, AL	Exome sequencing (~300K SNPs)	WHEALBI, FP7, 2019	multiple seasons and sites
Bread wheat (T. aestivum)	483	DtH,PH,TKW, AL, PL + other categorical traits (cold tolerance, biotic stress resistance)	Exome sequencing (~100K SNPs)	WHEALBI, FP7, 2019	
Barley ( <i>H. vulgar</i> e)	~200	DtH, GY + phenological adaptation, yield stability, and resilience	50k SNP array	BARISTA. SusCrop-ERA-NE T, 2022	multiple seasons and sites
Barley (H. vulgare)	~1000 (diverse landraces)	DtH	GbS (~100K SNPs)	GENDIBAR, EU-PRIMA, 2022	multiple seasons and sites
Durum ( <i>T. durum</i> )	250 (cultivar, landraces, emmer)	DtH,PH,TKW + other categorical spike fertility grain characterzation	Exome sequencing (~800K SNPs)	UNIBO (background material)	multiple years
Durum (T. durum)	250 (cultivar, world)	HTP vegetation indexs DtH,PH, + osmotic adj, drought index	illumina (~90K SNPs array)	InnoVar, EU-H2020, 2024	multiple years
Durum ( <i>T. durum</i> )	250	DtH,PH,TKW + GY and components NDVI, SPAD	illumina (~90K SNPs array)	InnoVar, EU-H2020, 2024	multiple sites
Durum ( <i>T. durum</i> )	300	HTP vegetation indexs, DtH,PH,TKW, GY and components, NDVI, SPAD, disease rust fusarium	illumina (~90K SNPs array)	InnoVar, EU-H2020, 2024	
Durum (T. durum)	300 (landraces)	phenological adaptation, yield stability, resilience and resistance to diseases	illumina (~90K SNPs array)	CerealMed, EU-PRIMA, 2022	multiple sites
Durum (T. durum)	300 (cultivars)	phenological adaptation, yield stability, resilience and resistance to diseases	illumina (~90K SNPs array)	CerealMed, EU-PRIMA, 2022	multiple sites

- 2. rhizotron experiment (led by Jülich): 15 jan 15 feb 2025
  - a. durum wheat, ~60 varieties
- 3. <u>sequencing/genotyping experiment</u> (led by UNIBO): in progress
  - a. durum and bread wheat:
    - i. ~300 varieties (150/150): genotyping (25k SNP array)
    - ii. sequencing (high-coverage):  $\sim$ 50-100 varieties  $\rightarrow$ **polyploidy**

#### new data + historical data











### Polyploidbreeding - drone phenotyping

Campo	Sensore	Altezza volo (m)	GSD (cm/pixel)	Overlap	Sidelap
1	RGB	12	0.33	80%	80%
1	Termico	12	1.57	68%*	90%
1	Multispettrale	12	0.83	80%	80%
2	RGB	12	0.33	80%	80%
2	Termico	12	1.57	68%*	90%
2	Multispettrale	12	0.83	80%	80%

Location

- Camera
- Height
- Ground Sampling Distance (~resolution)



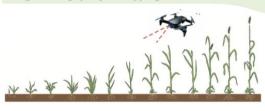






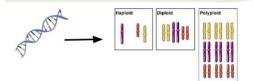
### Polyploidbreeding - drone phenotyping

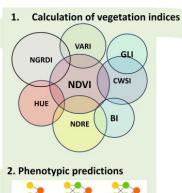
UAV (unmanned aerial vehicle: drone)high-throughput phenotyping:



Multi-camera image acquisition:

**Genomic information** 

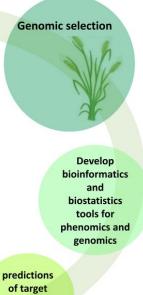




DNN models to predict: i) future growth; ii) high-resolution phenotypes from lower-resolution data

3. Genomic predictions

Use genomic data to predict image phenotypes (besides traditional phenotypes)



\*\*\*\* \* \* \*\*\*

ÿ

Polyploid Breeding

**PRIN 2022** 

Finanziato dall'Unione europea NextGenerationEU







phenotypes





### Tool development: drone2report

#### https://github.com/ne1s0n/dro ne2report

Python-based software to read drone-captured images (RGB, multispectral, thermal) and calculate vegetation indices

README   MIT license	0	) :=
drone2report		
From orthophoto to summary statistics, indexes, and more.		
Usage		
Clone the repository:		
git clone https://github.com/neis0n/drone2report.git		
Create the conda enviroment with:		
cd drone2report conda env createfile environment.yml	d	p
Activate the environment:		
conda activate drone2report		
then fill an appropriate .ini file and run:		
python3 drone2report.py <your .ini="" config="" file=""></your>		
The idea is that you don't write any python code, but just fill your appropriate .ini file.		
Then, if you really really want, you can write some python code (i.e. you can easily expand t your own custom analyses).	he pipeline with	

The sample\_config.ini file is self documented and should be easy enough to understand to let you start.

















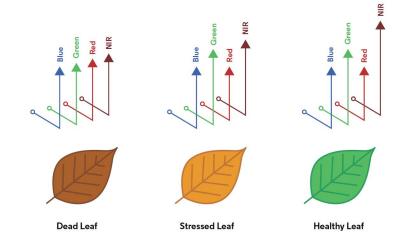


### **Vegetation Indices**

Vegetation Indices are **metrics / summary stats** obtained from spectral images of any kind, that are intended to capture some properties of the crops

<u>https://www.indexdatabase.de/db/i.php</u> : > 500 different vegetation indices

- $RGB \rightarrow VARI$
- Multispectral  $\rightarrow$  NDVI
- Thermal  $\rightarrow$ TSI
- DEM  $\rightarrow$  height, volume



https://support.insights.granular.ag/hc/en-us/articles/360034834012-How-does-the-Vegetation-Index-work









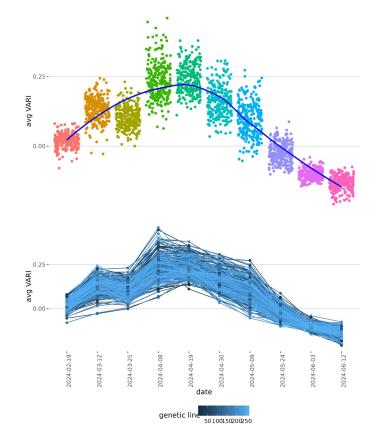


#### VARI: Visible Atmospherically Resistant Index

- **RGB** camera (M3E): 400~700nm.
- VARI emphasizes vegetation in the visible portion of the spectrum, while mitigating illumination differences (reflectance, scattering) and atmospheric effects

VARI can be used to detect <u>changes in biomass</u> accumulation and responds to the amount of chlorophyll in the leaves

$$VARI = \frac{(\text{green} - \text{red})}{(\text{green} + \text{red} - \text{blue})}$$









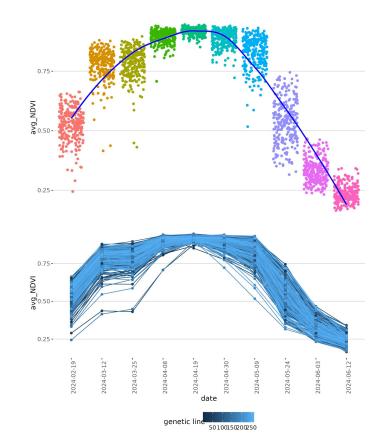




#### NDVI: Normalized Difference Vegetation Index

- **multispectral** camera (MICA): measure the plant canopy reflectance in both the visible (VIS) and the near-infrared (NIR) spectra
- normalized difference between NIR (correlated to leaf structure) and red (correlated to chlorophyll content)
- visible and NIR vegetation reflectance intensities are, respectively, negatively and positively correlated to leaf N content and biomass
- greater leaf area and green plant biomass translate into higher NDVI values

$$NDVI = \frac{(NIR - red)}{(NIR + red)}$$













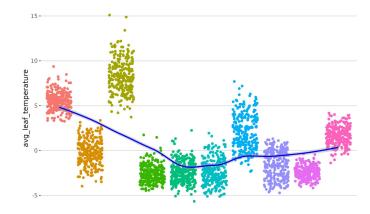
r

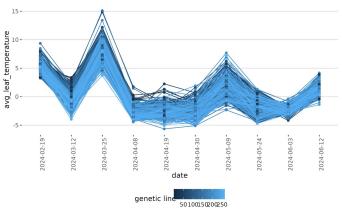
#### **TSI: Thermal Stress Index**

- Thermal camera (M3T)
- canopy to air temperature difference: canopy temperature considered as a proxy for monitoring crop water status

$$TSI = (T_{foliage} - T_{air})$$

DATE and HOUR	Tair
19/02/2024 h 13.00	14,8 °C
12/03/2024 h 12.00	16,7 °C
25/03/2024 h 12.30	15,5 °C
08/04/2024 h 13.00	23,5 °C
19/04/2024 h 12.00	17 °C
30/04/2024 h 12.45	23,7 °C
09/05/2024 h 12.30	22,5 °C
24/05/2024 h 14.30	20,7 °C
03/06/2024 h 11.30	21 °C
12/06/2024 h 11.30	24 °C







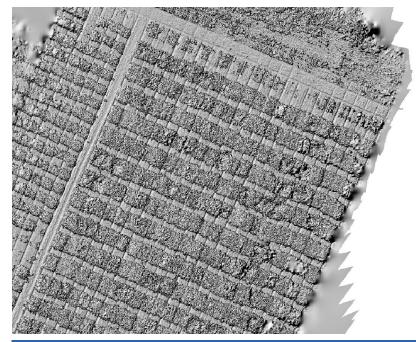








### **Digital Elevation Model**



- .dem file
- hillshade view
- (same barley field)





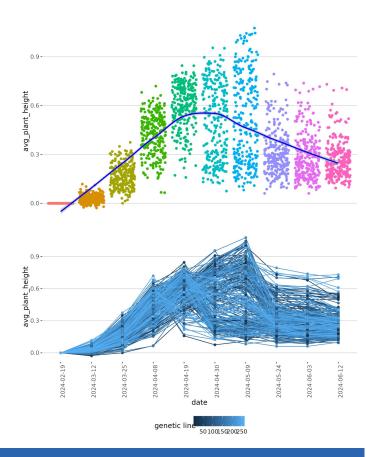






### $DEM \rightarrow plant height$

- First flight taken as reference
- meters









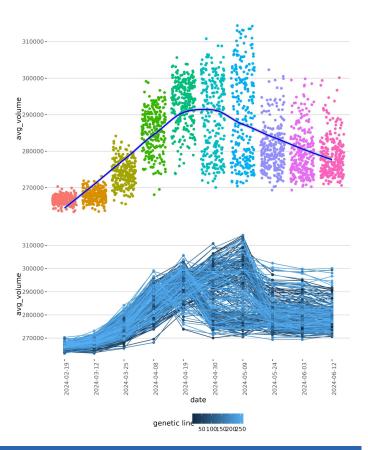






#### $\mathsf{DEM} \rightarrow \mathsf{volume}$

- First flight taken as reference
- By summing up all plant heights we can get the volume
- Volume is a good proxy for crop biomass
- Still refinements are needed: i) check normalization; ii) thresholding; iii) other?















#### Part 3: The AGRI-VISION COST-Action proposal











## AGRI-VISION

"Computer vision applications for sustainable, productive and efficient food & agriculture sectors"

- COST Action proposal [submitted]: interdisciplinary research network
- food & agriculture (livestock, plants, aquaculture, insects, food science), computer science, statistics, law
- € 575k (total), 4 years





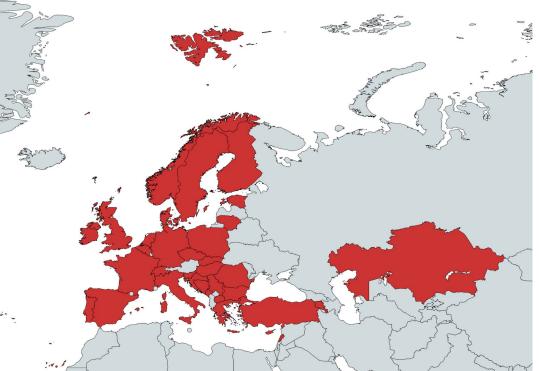






### AGRI-VISION - network

- 80 partners
- 35 countries and international organizations
- 60% ITCs
- academia, government, SME
- coordinator: CNR















#### WG1: Technology, data and infrastructure for Computer Vision

- survey on <u>devices</u>, <u>hardware</u> and <u>data</u> requirements for computer vision: current landscape and future developments
- training school(s) on <u>big data processing</u>, <u>scalable infrastructure</u> and <u>ML methods for computer</u> <u>vision</u> for food & agriculture (jointly with WG2)
- guidelines and recommendations on the <u>needed infrastructure</u> for computer vision in food and agriculture at the European scale, including scaling up from national to EU level











#### WG2: Methods and models for Computer Vision

- report on state-of-the-art and cutting-edge <u>ML methods and models for computer vision</u> to be used in food & agriculture
- guidelines on standardised practices for the application of computer vision to food & agriculture, including use of <u>software tools</u>, <u>model implementation</u> and <u>data annotation</u>
- tutorials on the practical <u>use of computer vision</u> tools and models, that will cover the most common scenarios and challenges











#### WG3: Case-studies from food & agriculture

- report on current and future (foreseen, desired) <u>computer vision applications</u> in various areas of food & agriculture: livestock, insect, plant farming and breeding, aquaculture, food processing
- videos of expert talks on the <u>impact of computer vision</u> on sustainable agriculture and food production in different areas











#### WG4: Societal impact, policies and the future frontier

- critical review of computer vision applications in food & agriculture in relation with the 2024 <u>EU</u> <u>AI Act</u>
- workshop on the drafting of <u>data sharing agreements</u>, <u>intellectual property</u> and <u>licensing</u> issues, and the <u>compliance with GDPR</u>
- report on the <u>future directions</u> of computer vision developments and the relationship with similar technologies
- collection of available <u>surveys on the benefits and potential risks of AI</u> applications in food and agriculture as perceived by relevant stakeholders











#### WG5: Dissemination, training and public engagement

• white paper on current and future <u>computer vision developments</u> in food and agriculture (jointly with all other WGs)

COST Actions are usually open and it is typically possible to recruit additional partners also after funding: in case, get in touch!













#### COST Actions are usually open and it is typically possible to **recruit additional partners also after funding**: in case, get in touch!











# Tak!







