UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura

DEPARTMENT OF MATHEMATICS, COMPUTER SCIENCE AND PHYSICS

Master degree in Artificial Intelligence & Cybersecurity

Artificial Intelligence for Precision Agriculture: Semantic Segmentation of Crop Fields Using Deep Learning

SUPERVISOR Prof. Giuseppe Serra

CO-SUPERVISORS Prof. Kyamakya Kyandoghere
Dott. Alex Falcon
Dott.ssa Beatrice Portelli

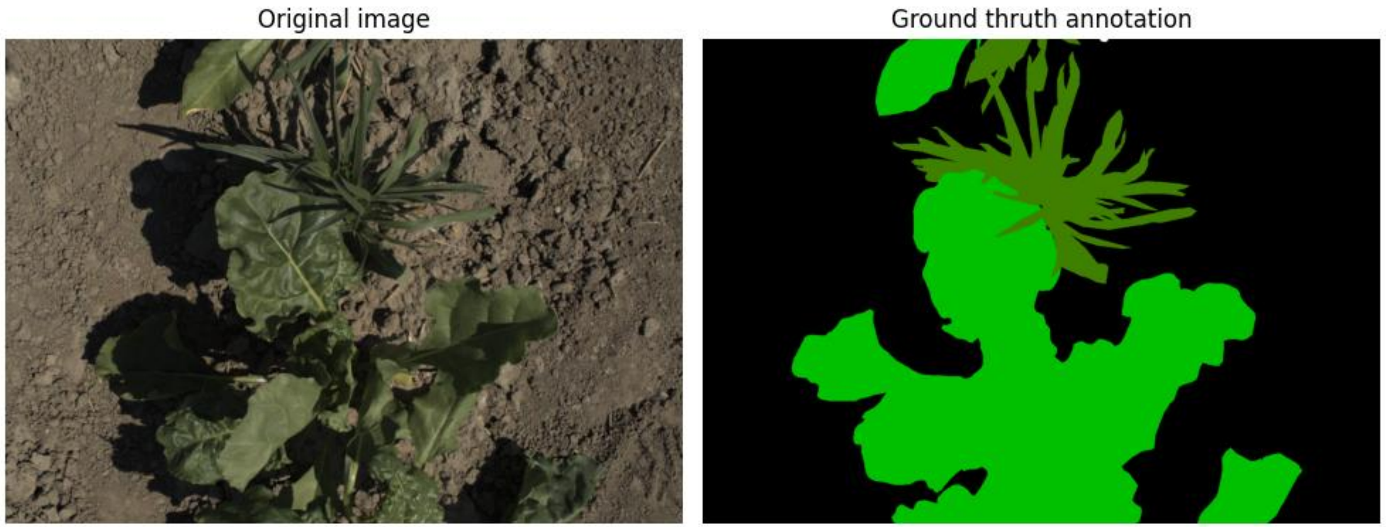
CANDIDATE Federico Dittaro

Academic Year 2024/2025

MOTIVATION AND OBJECTIVES

- Automate the estimation of vegetation cover (the proportion of ground area covered by vegetation) from images captured in real agricultural fields.

- Achieve a balance between accuracy and efficiency:
  - Accuracy ensures reliable estimation of vegetation classes.
  - Efficiency allows real-time or large-scale processing in practical field conditions.



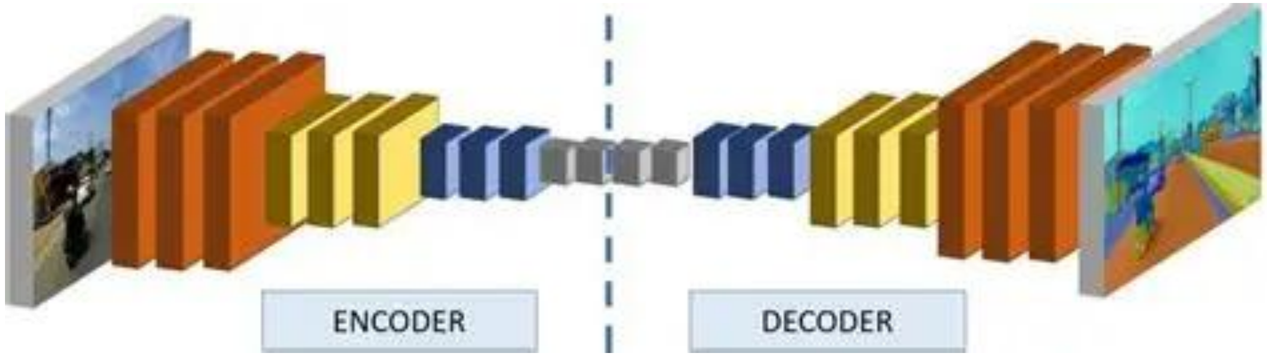
BACKGROUND – CNNs, ENCODER & DECODERS

- Convolutional Neural Networks (CNNs) are a class of deep neural networks designed to extract spatial patterns from images, making them ideal for pixel-level semantic segmentation.

- Within the semantic segmentation, we distinguish two main components:

1. Encoders: extracts high-level features from the input image
    - EfficientNet-B0
    - MobileNetV2

2. Decoders: restores the information into a full-resolution segmentation map
    - UNet
    - UNet++
    - DeepLabV3



BACKGROUND – LOSS FUNCTIONS

- We explored different loss functions to deal with imbalanced datasets where some vegetation classes are underrepresented.

1. **Cross Entropy Loss**: measures the dissimilarity between the predicted probability distribution and the true class labels.

2. **Weighted Cross Entropy Loss:** uses class-specific weights into the loss formulation.

3. **Dice Loss:** is based on a region-overlap measure between the predicted segmentation and the ground truth.

4. **Focal Loss:** dynamically down weights the contribution of easy examples and focus the training on hard examples.

DATASETS – WE3DS, CROPANDWEED AND POST-EVALUATION

	WE3DS	CropAndWeed	UNIUD
N° of images	2.568	8.034	40
Perspective	Top-down	Top-down	Top-down
Annotation type	Pixel-level	Pixel-level	Pixel-level
N° of classes	18	101	3
Usage	Train, val, test	Train, val, test	Test

EXPERIMENTAL SETUP - EVALUATION METRICS

- The evaluation was conducted on the test sets by computing the following metrics:

  1. **Intersection over Union (IoU)**: quantifies the overlap between the predicted segmentation mask and the ground truth mask.
  2. **Precision**: evaluates how many of the predicted vegetation pixels are actually correct.
  3. **Recall**: evaluates how many of the true vegetation pixels were correctly identified.
  4. **F1-Score**: is the harmonic mean of precision and recall.

- Metrics were computed per class and then averaged to highlight global performance.

RESULTS – BEST MODEL

- To identify the most suitable segmentation model, a two–stage procedure was adopted:

  1. First stage: each model was trained for 25 epochs with a batch size of 8, employing early stopping with a patience of five epochs.
  2. Second phase: the goal was a careful examination of batch-size effects. Training was extended to 50 epochs and early stopping was disabled.

BEST MODEL: EfficientNet-B0 + UNet with a batch size of 32.

QUANTITATIVE RESULTS

WE3DS

Dice Loss achieved the highest precision (0.6474) and F1-score (0.6120), coupled with a recall value of 0.6011. These results indicate that Dice Loss provides a superior performance for the WE3DS dataset compared to the other loss functions.

Loss Function	IoU	Precision	Recall	F1-score
CE	**0.4694**	0.6067	0.5566	0.5690
WCE	0.3996	0.5055	**0.6131**	0.5420
Focal	0.4056	0.5413	0.4886	0.5077
Dice	0.4484	**0.6474**	0.6011	**0.6120**

CropAndWeed

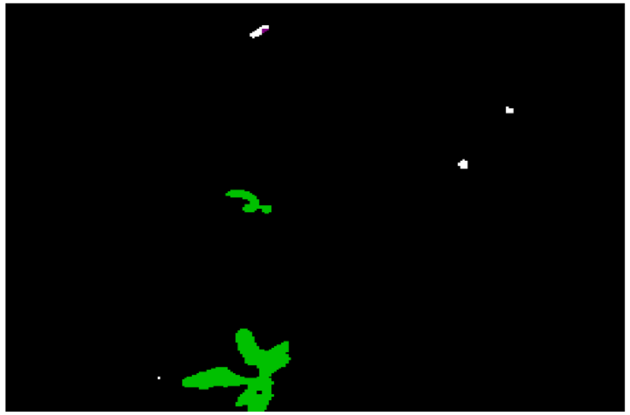
The CropAndWeed dataset presents a substantially more complex challenge compared to WE3DS. This complexity is reflected in the lower overall performance metrics across all loss functions. In this case WCE stands out as the best loss function.

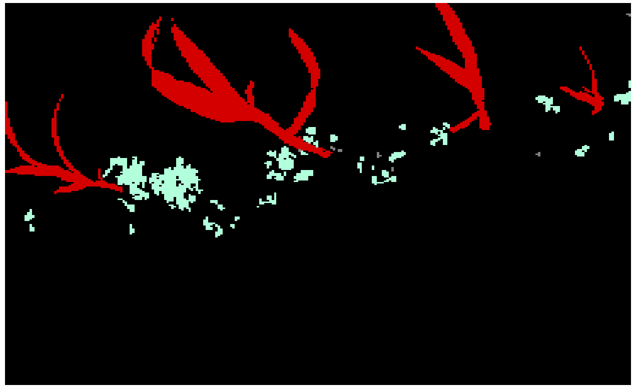
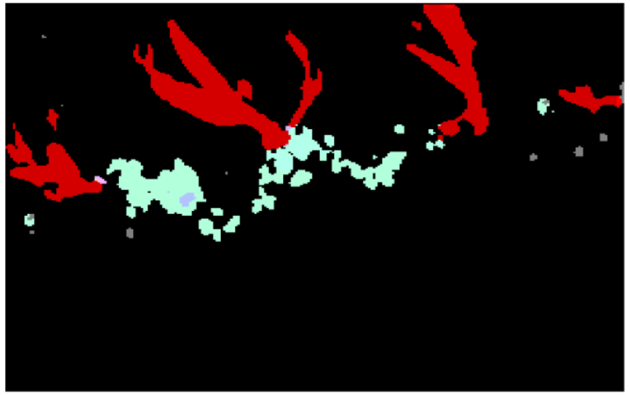
Loss Function	IoU	Precision	Recall	F1-score
CE	0.1329	**0.3017**	0.2348	0.2452
WCE	**0.1406**	0.2722	**0.3229**	**0.2803**
Focal	0.1087	0.2738	0.2029	0.2067
Dice	0.1007	0.1710	0.1983	0.1752

QUALITATIVE RESULTS

Original image | **Original mask** | **Predicted mask**





RESULTS ON THE POST-EVALUATION DATASET

- Given the high native resolution of the images, different prediction strategies were explored to balance computational efficiency and segmentation fidelity:
    1. Global resizing
    2. Patch-based inference

Method	Average Inference time
1-Patch	317 ms/img
6-Patches	567.5 ms/img
12-Patches	775.8 ms/img
35-Patches	1726.3 ms/img

Class	1-Patch	6-Patches	12-Patches	35-Patches
Soil	0.9641	0.9654	0.9762	0.9949
Buckwheat	0.4221	0.5029	0.5210	0.5315
Weeds	0.0384	0.0394	0.0494	0.0572

Precision metrics

Class	1-Patch	6-Patches	12-Patches	35-Patches
Soil	0.8655	0.9220	0.9240	0.9376
Buckwheat	0.7628	0.7639	0.7805	0.7794
Weeds	0.0032	0.0042	0.0054	0.0061

Recall metrics

Class	1-Patch	6-Patches	12-Patches	35-Patches
Soil	0.8610	0.9140	0.9288	0.9330
Buckwheat	0.4165	0.4959	0.5210	0.5214
Weeds	0.0030	0.0045	0.0061	0.0072

IoU metrics

Class	1-Patch	6-Patches	12-Patches	35-Patches
Soil	0.9164	0.9471	0.9529	0.9652
Buckwheat	0.5392	0.6070	0.6107	0.6253
Weeds	0.0056	0.0063	0.0071	0.0087

F1-score metrics

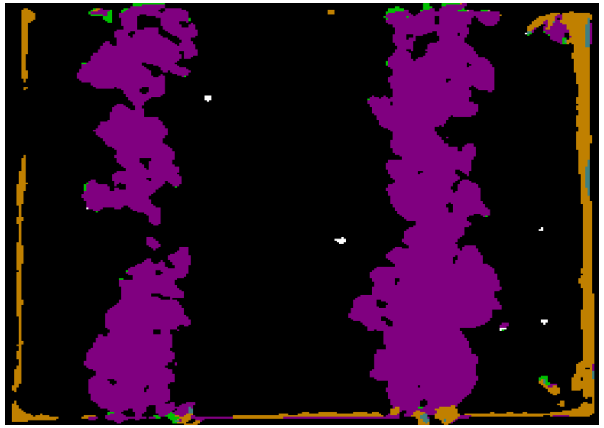
UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura

RESULTS ON THE POST-EVALUATION DATASET

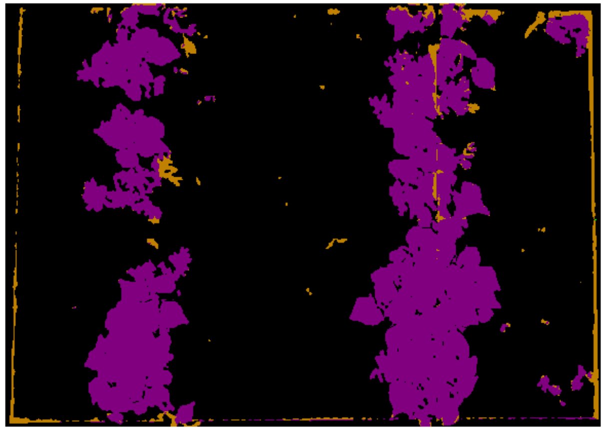
Original image

Original mask

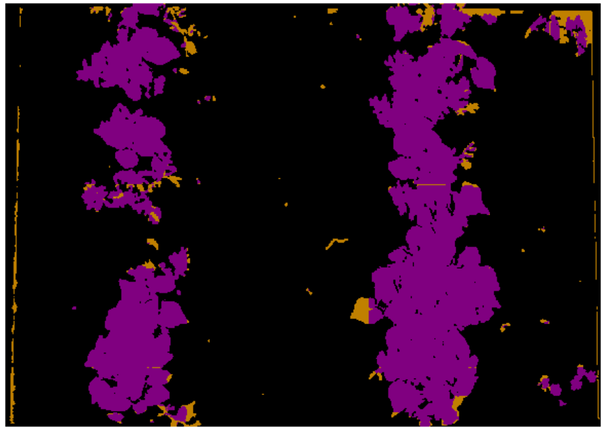
1-patch prediction



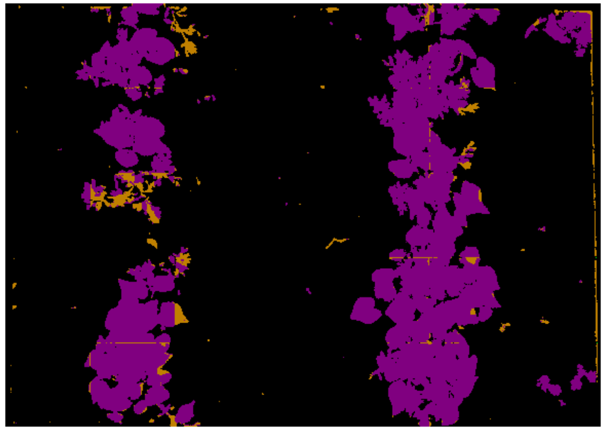
6-patches prediction



12-patches prediction



35-patches prediction



CONCLUSIONS

Main achievements

- Developed a deep learning framework to automate vegetation cover estimation from RGB images, addressing limitations of manual assessment.
- Ensured reproducibility and scalability through systematic model comparison and robust evaluation.

Future works

- Expand datasets with more species, cropping systems, and seasonal variations.
- Integrate multispectral and LiDAR data for richer vegetation representation.
- Apply temporal analysis for tracking crop development over time.