Computer Vision Case Study: Deep Learning for Satellite Imagery and Land Cover Statistical Estimation

Francesco Pugliese, PhD
neural1977@gmail.com

Outline

Goal

Land Cover (LC) statistics and maps represent a critical statistical product, for Official Statistics, for example. Since they require a big effort to be created, the underlying idea is building an automatic and efficient system that processes satellite images in order to generate:

- Automatic Land Cover Maps
- Automatic Land Cover Estimates

Methodology

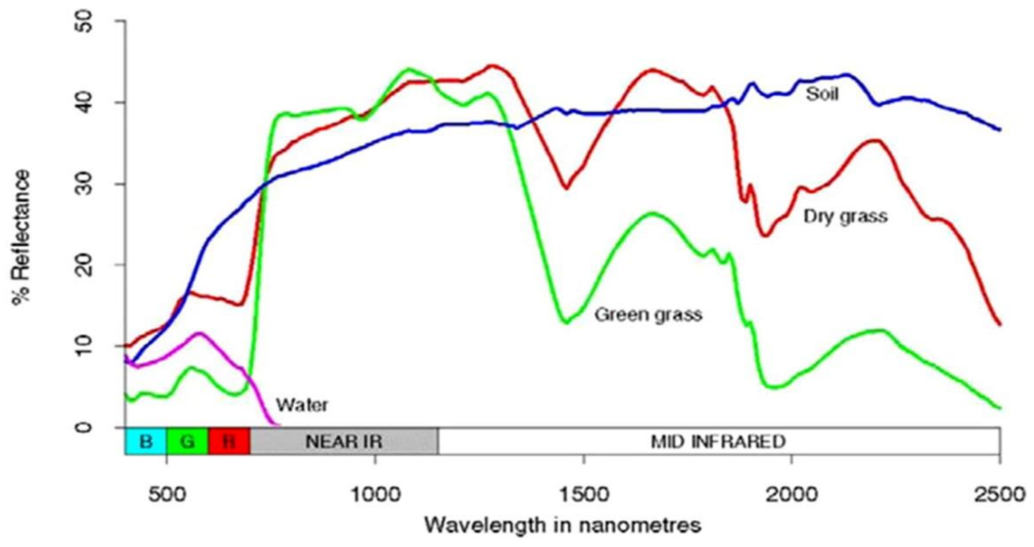
New Approach: using Deep Learning techniques to generate Land Cover maps (classification and segmentation through CNN and U-Net).

Results

A Deep Learning-based integrated (CNN + U-Net) and automated architecture that gives accurate results for all LC classes

Deep Learning for Land Cover from Satellite Imagery

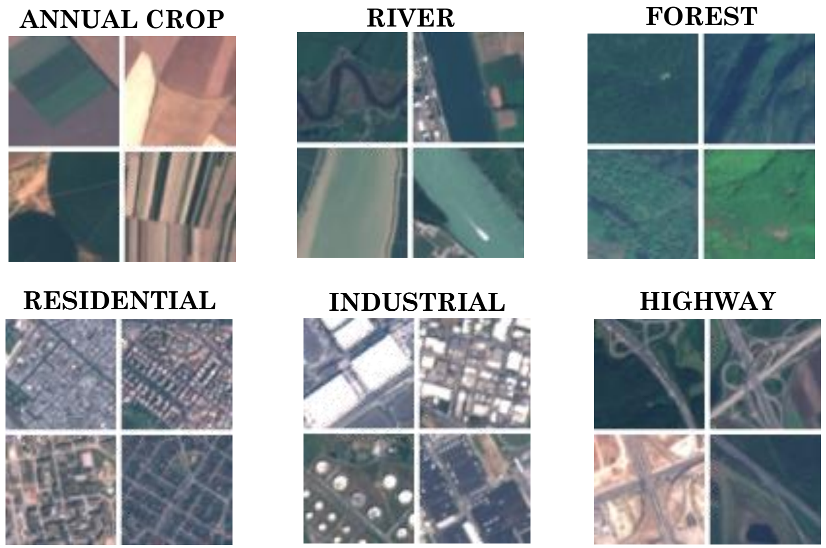
Standard approach: Spectral Signature



Different LC classes have different reflectance spectra

- Patterns of Reflectance Variation are used to predict LC classes
- A Trained ML algorithm predicts the LC class independently
- Decision on each data point does not depend on neighboring data points

New approach: Computer Vision (Deep Learning)



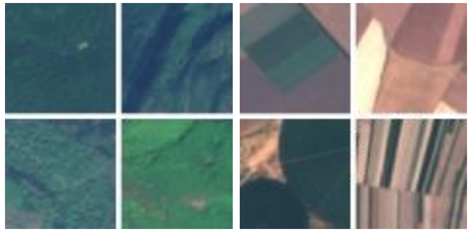
Different LC classes have different spatial/visual patterns

- The Variation of visual/spatial patterns is used to predict LC classes
- Trained ML algo (CNN) predicts the LC class of image pixels based on information from neighboring pixels
- Decision on each pixel depends on the whole sub-image (tile) the pixel belongs to

Machine Learning for LC from satellite imagery

TRAINING DATASET

The Training.set affects the classification and prediction accuracy of our machine learning algorithm. We have harnessed the **EuroSAT** dataset features and its EuroSAT Land-Cover classification. EuroSAT is made of satellite imagery which is carefully collected and selected from Sentinel 2 Satellite Copernicus Project.

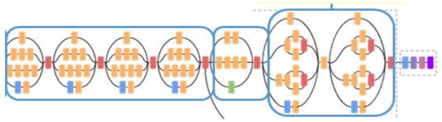


INPUT

Input images to be processed, must be in excellent agreement with training pictures. That’s why we take portions of Italy satellite imagery from Sentinel 2 itself.

ARCHITECTURE

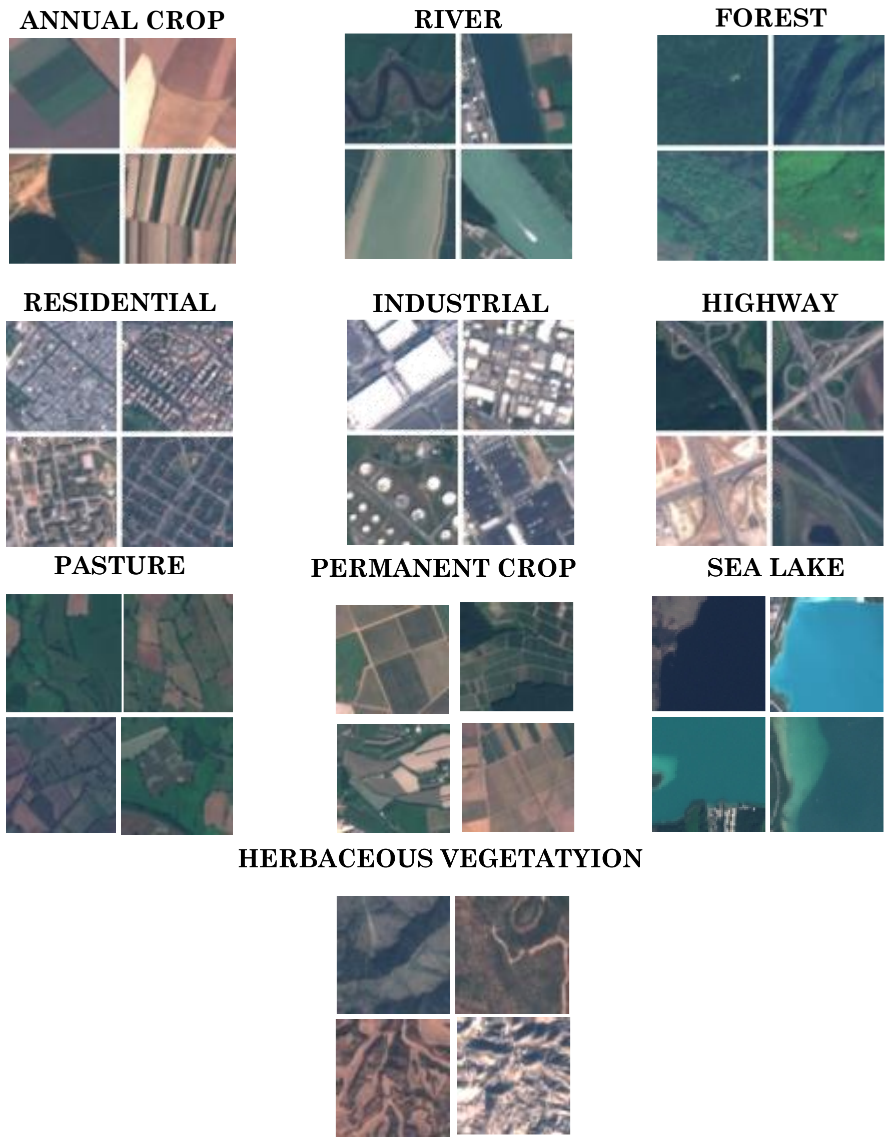
We chose our Algorithm Architecture accordingly: **Classification Convolutionary Neural Network (CNN) or Segmentation Convolutionary Neural Networke (U-Net).** We built the model by selecting the right hyper-parameters, optimizations and validations.



OUTPUT

The output ground truth is within the training dataset in order to validate the model. Then, the ouput is yielded by processing the Input dataset in order to perform classifications/predictions.

CNN: Satellite Imagery Dataset

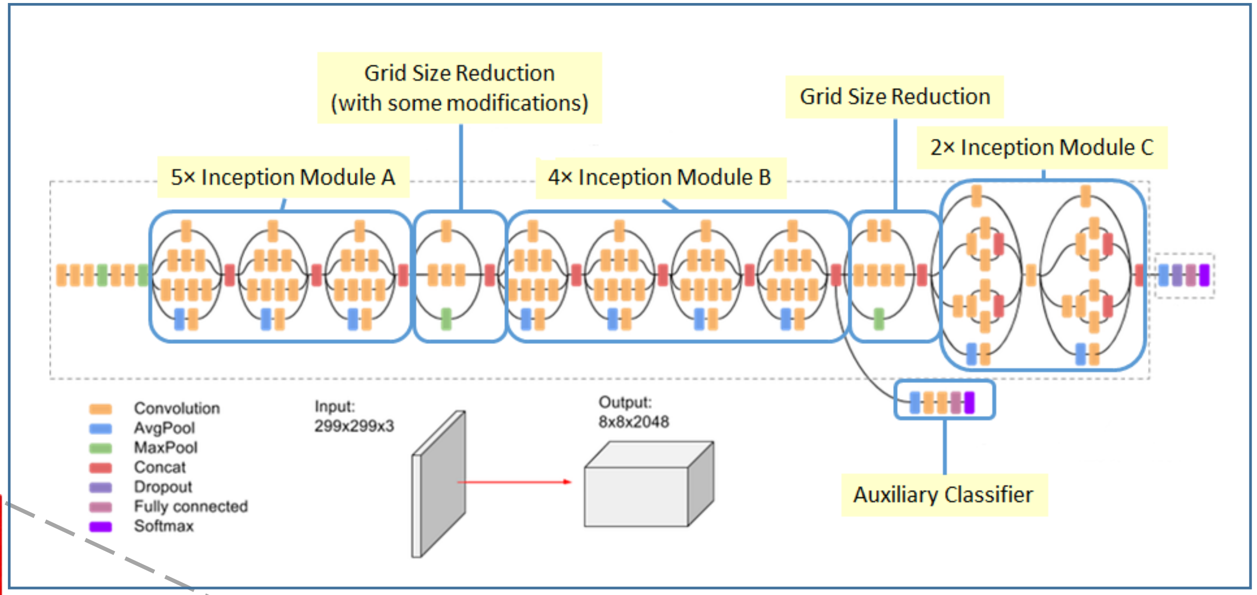


EuroSAT dataset
(https://github.com/phelber/eurosat):

- Based on Sentinel-2 satellite images
- 27000 geo-referenced and labeled image patches (each one of 64x64 pixels)
- 10 different Land Use and Land Cover classes, with 2000-3000 images per class
- RGB (8-bit) and Multi-Spectral (13 spectral bands, 16-bit) versions available

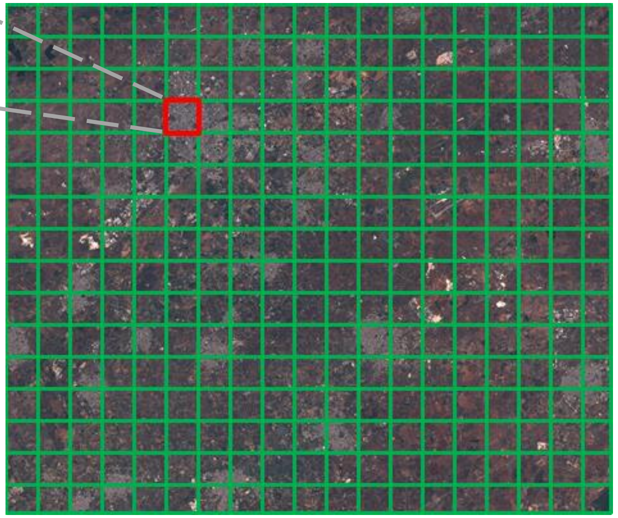
CNN: Inception-V3 Architecture

Classify-and-Count Approach



RESIDENTIAL

Input satellite image



LAND COVER CLASS	AREA SHARE
...	...
RESIDENTIAL	$\frac{45}{16 * 19} \cong 15\%$
...	...

GoogleNet – Google (Szegedy, C., et al., 2015)

Critical Feautures (Szegedy, C., et al., 2015):

Computationally Effective Deep architecture: 22 layers

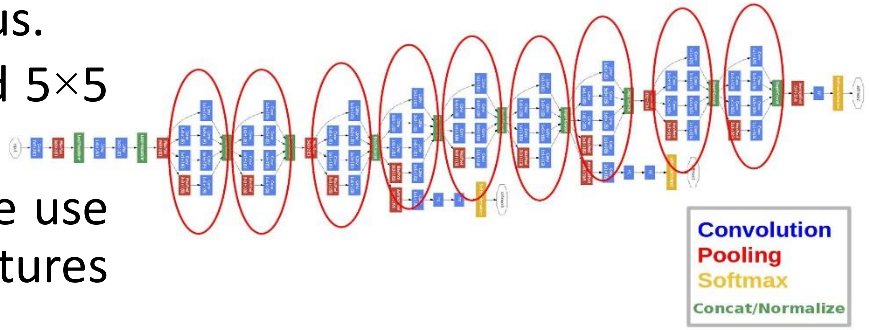
Why the name inception, you ask? Because the module represents a network within a network. If you don't get the reference, go watch Christopher Nolan's ***“INCEPTION”***, computer scientists are hilarious.

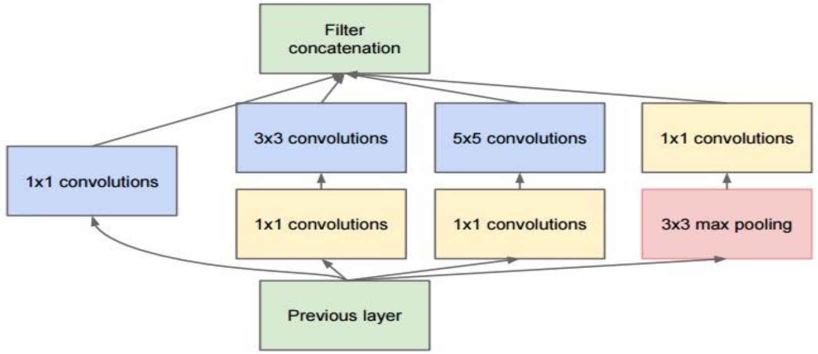
Inception: it isbasically the parallel combination of 1×1, 3×3, and 5×5 convolutional filters.

Bottleneck layer: The great insight of the inception module is the use of 1×1 convolutional blocks (NiN) to reduce the number of features before the expensive parallel blocks.

Upside: 4 millions parameters!

Downside: Not scalable!



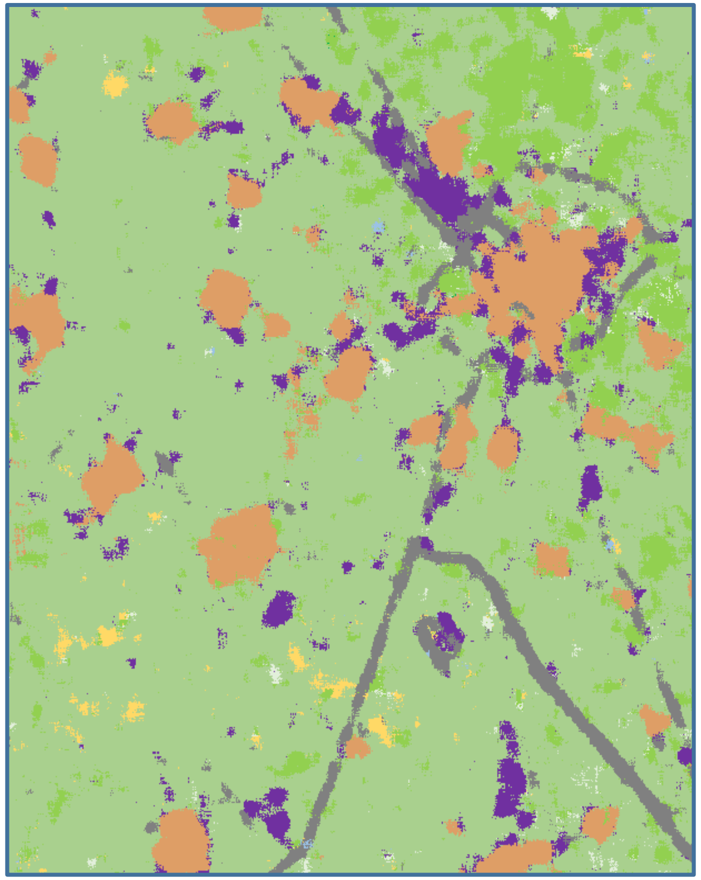


Results:

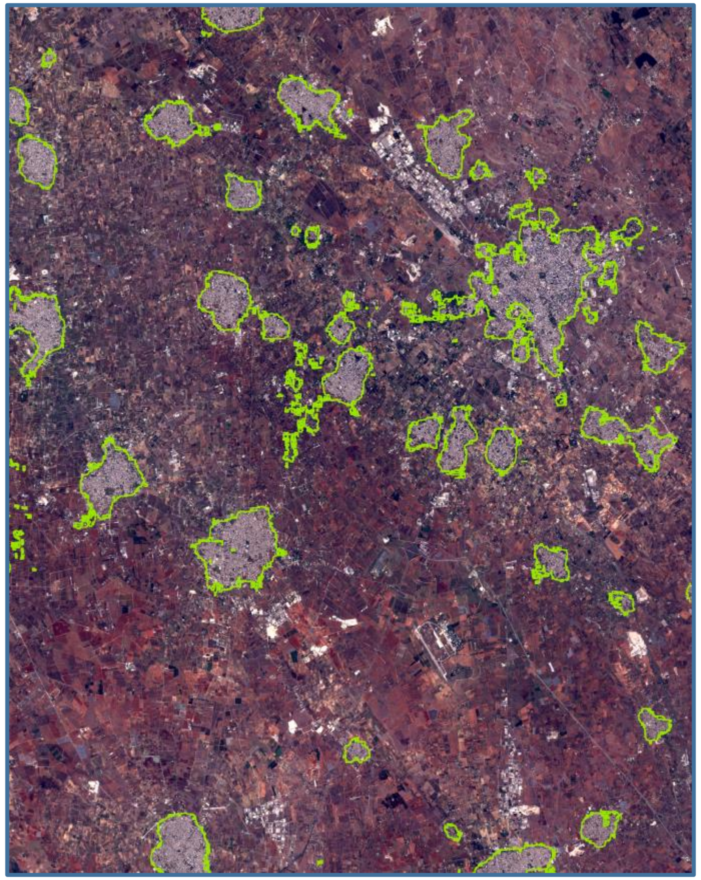
7 Models Ensemble : 6.67% Top-5 Error.

CNN: Example of Automatic Land-Cover Map

[A] Input Lecce’s picture (751 km^2)



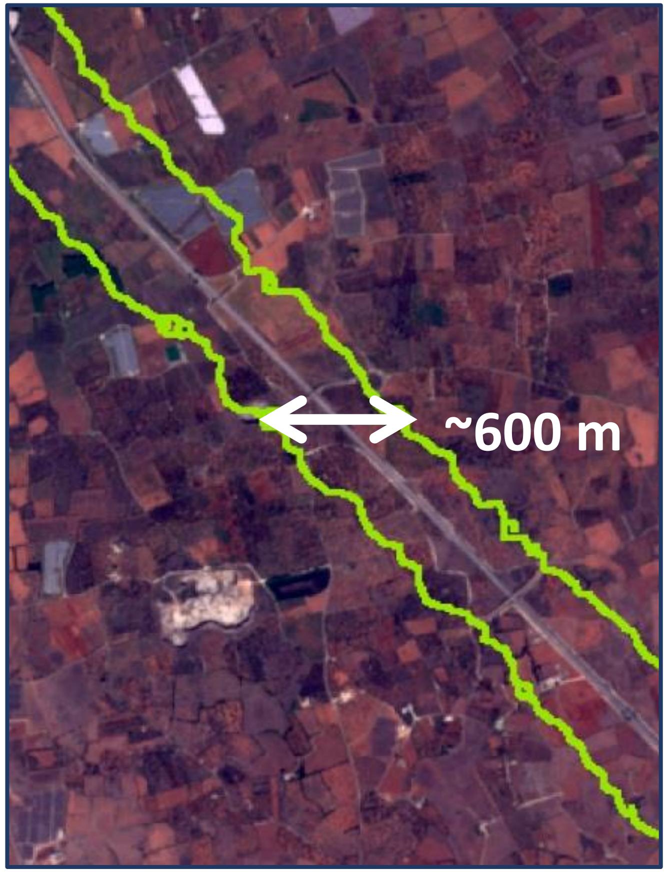
[B] Related Lecce LC Map (output)



[C] Depiction of the *Residential* class generated by overlapping [A] and [B]

CNN: Over-estimation effect of *River* and *Highway* classes

[D] Detailed overview of the **Arno River** pathway (**Pisa**, 443 km^2) on top of a semi-transparent version of the corresponding LC map.



[E] Fragment of Lecce highlighting the border of the *Highway* class predicted by the model. We can see the extent of the over-estimation

U-Net: Dataset creation

EuroSAT Imagery *River* | Segmentation Masks *River*

- The dataset creation to train the **U-Net** segmentation model was performed by exploting satellite pictures from *River* class by ***EuroSAT*** and data from the *Copernicus* ***High Resolution Layer****.*

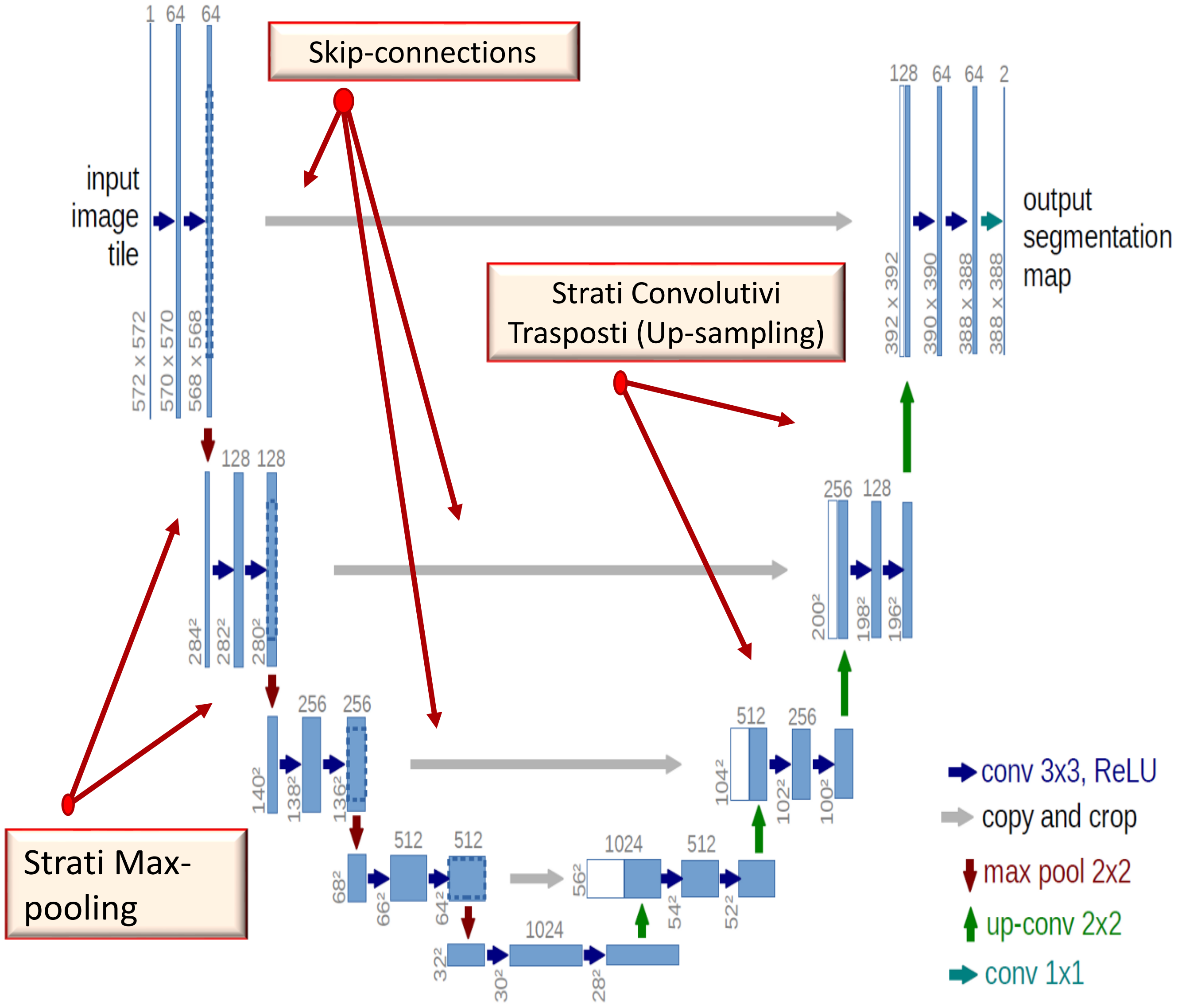
- For each image we determined a label consisting in a ***mask*** keeping track of the class information for every pixel:
  0 means the pixel belongs to *River* class (**black**)
  1 means the pixel does not belong to *River* class (**white**)

- Main difference with the Standard (Classification) CNN Dataset: in classification, the class label refers to the whole image and not to each single pixel such as in segmentation stage.

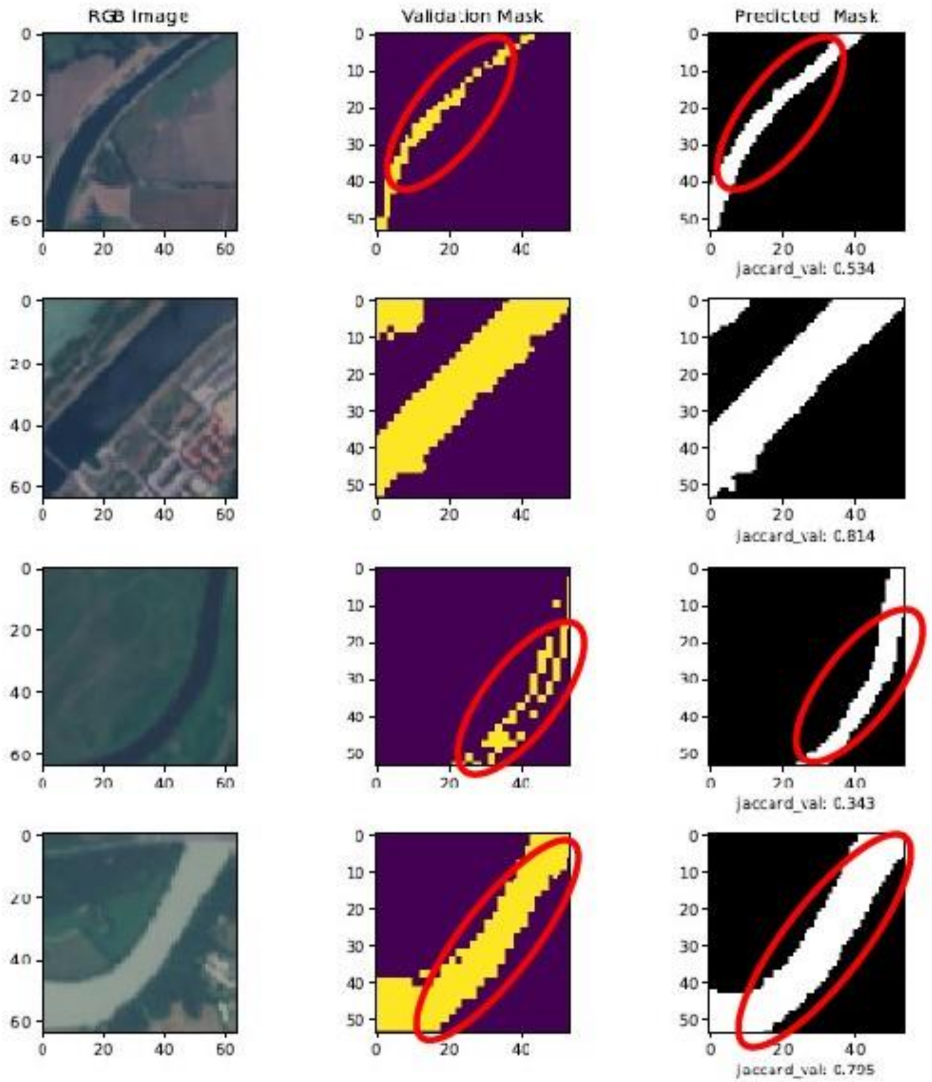
- Final dataset: 1500 segmentartion masks validated.

- Implementation of a similar training dataset for *Highway* thanks to ***Open Street Maps*** data.

U-Net: Architecture



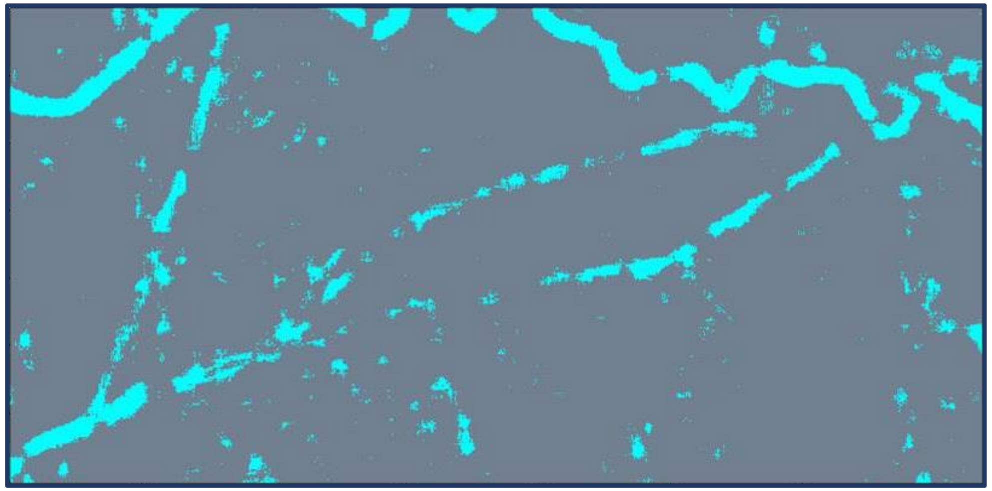
Completion abilities of U-Net

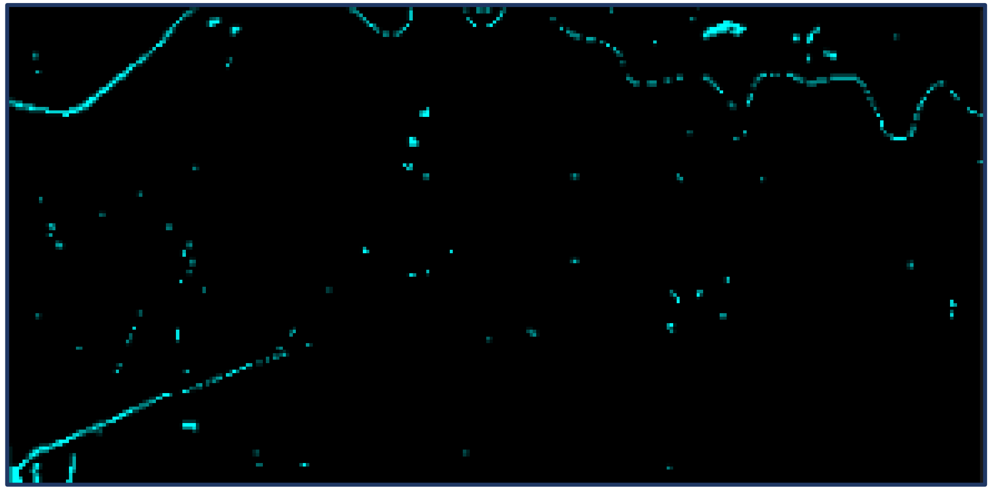


U-Net: adjustment of the over-estimation issue in “thread-like” classes such as *River*

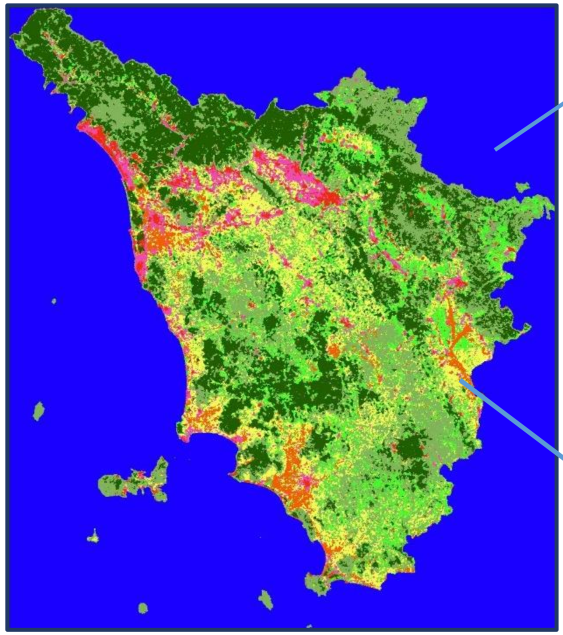
Classification by means of CNN

U-Net Segmentation

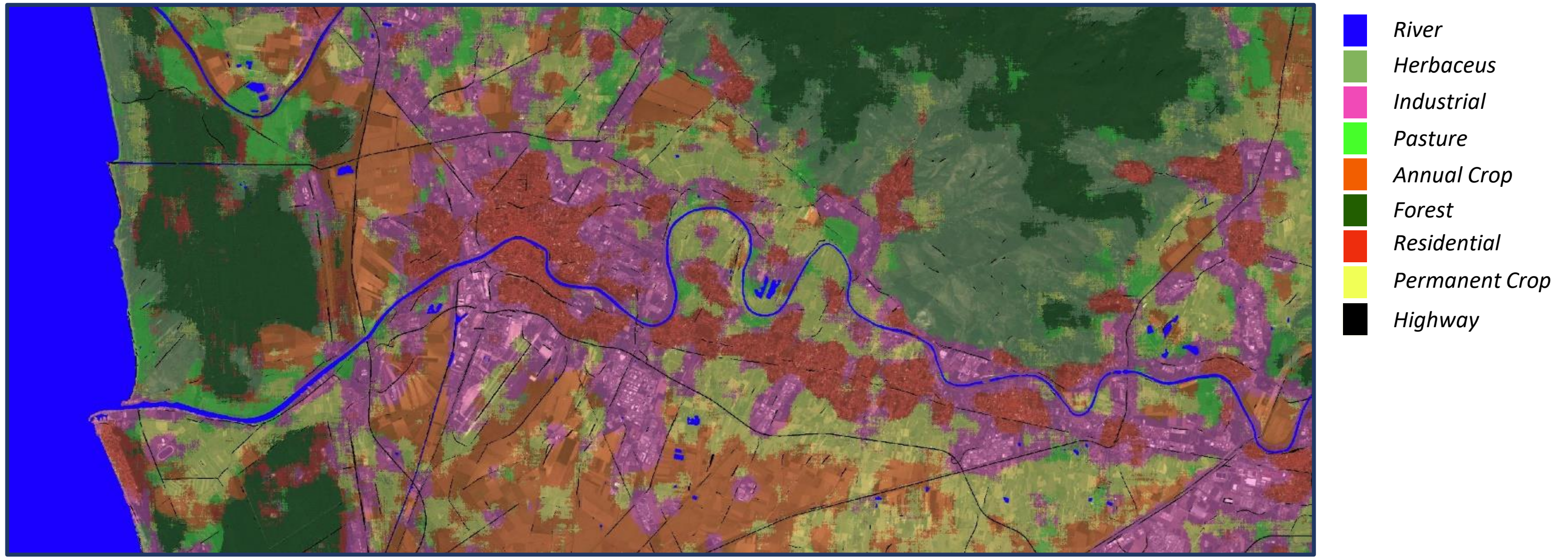




Integrated Map CNN + U-Net



Tuscany LC Map



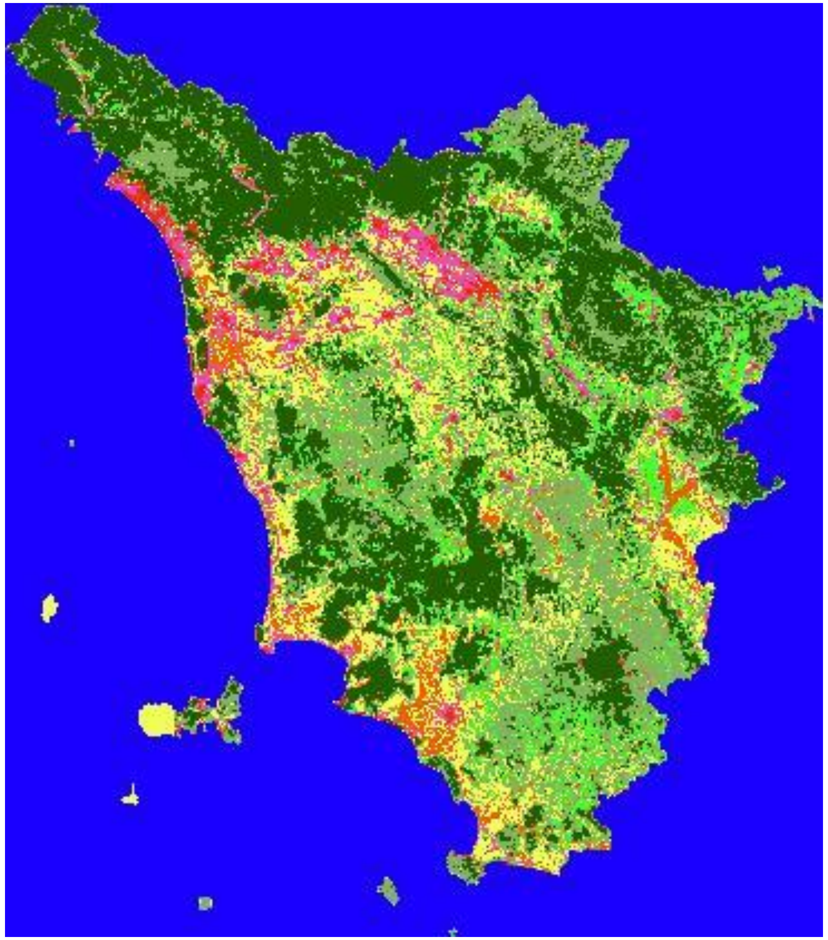
Detail of Arno river view (Pisa, 443 km^2) overlapped to a semi-transparent version of the corresponding map coming from the CNN and U-Net integration process. Different colours refer to different LC classes.

Integrated system (CNN + U-Net) works perfectly with all the LC classes

- U-Net predicts *River* and *Highway.*
- Whereas CNN classifies the other classes with are left over.

Example of integrated (CNN+U-Net) LC map and statistics

CNN + U-Net:

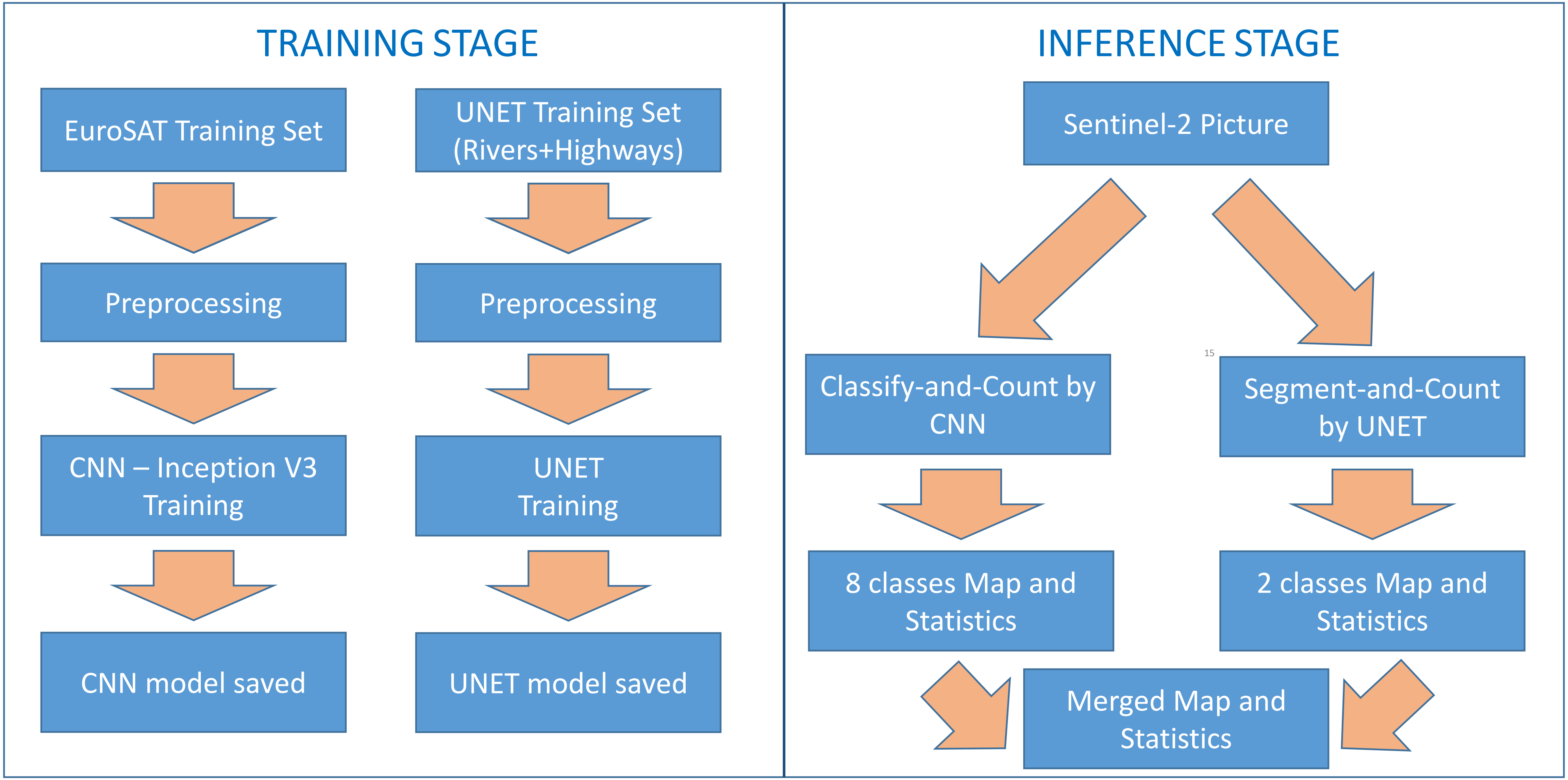


Land Cover map of Tuscany

⇨

LC CLASS	COVERAGE %
Annual Crop	4.1
Forest	34.1
Herb. Vegetation	27.2
Highway	0.7
Industrial	5.2
Pasture	10.6
Permanent Crop	14.6
Residential	2.4
River	1.6

Land Cover System – Pipeline Architecture

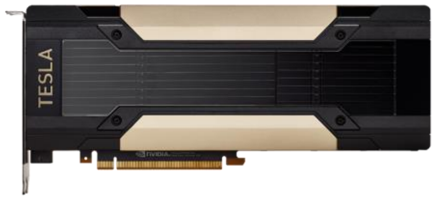


References

Azure Wishlist

1) Hadoop and **Spark** to fasten pre-processing and post-processing with different cloud nodes

2) More **GPUs** like **Tesla V100** or better to optimize **Deep Neural Networks** parallelization and speed up **Training** and **Testing**



3) Backup systems to save data, models and code

4) Cognitive services for Computer Vision and NLP tasks

Thank you for your attention

Francesco Pugliese