DeepLearn 2019
Warsaw

Istat

Automatic Extraction of Official Statistics Indicators from Satellite Imagery by Deep Neural Networks

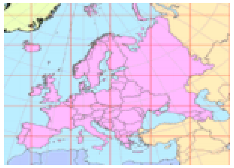
Fabrizio De Fausti, Francesco Pugliese, Diego Zardetto, Monica Scannapieco

Italian National Institute of Statistics, Division "Information and Application Architecture", Directorate for Methodology and Statistical Design

Email：defausti@istat.it, frpuglie@istat.it, zardetto@istat.it, scannapieco@istat.it

MOTIVATION OF THE WORK

- Eurostat has been carrying out the LUCAS survey every 3 years since 2006 (last round in 2018) to estimate Land Cover (LC) and Land Use in the European Union:



  - A 2-phase area sample survey of the whole EU territory
    - 1^{st} phase: Master Sample of ~1.1 million points in a square grid of (2 km x 2 km) cells
    - 2^{nd} phase: ~330,000 random points from the Master Sample
  - Direct data collection, mainly on the ground (~70% of 2^{nd} phase points), the rest by clerical photo-interpretation
  - Provides LC estimates for all Member States up to NUTS-2 territorial level

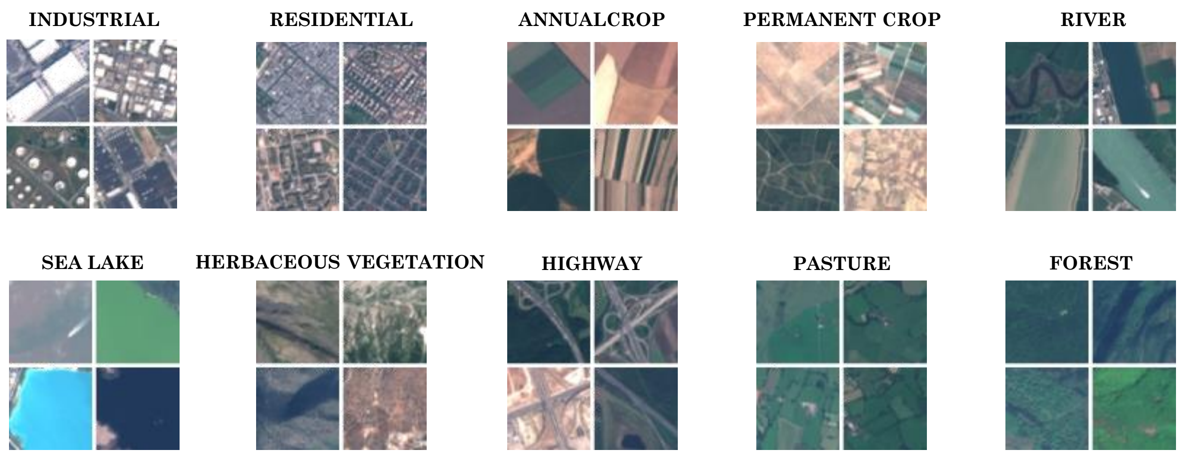
- **Computer Vision** methods (e.g. Deep Learning) **+** **Satellite Imagery** data (e.g. Sentinel-2) can be used for LC estimation:
  - **Classify-and-Count approach**
    - Train an image classification algorithm to predict the LC class of a satellite image tile
    - Divide the satellite images covering a target area into tiles and use the trained algorithm to predict LC classes
    - Obtain LC statistics for the target area by simply computing the relative frequencies of predicted LC classes

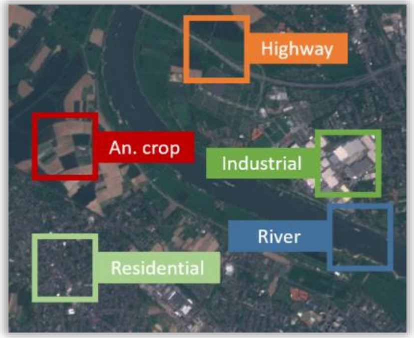
- **Pros**: (i) dramatically reduce data collection costs/burden, (ii) provide more timely statistics, (iii) provide LC statistics beyond the NUTS-2 level, (iv) produce moderate resolution maps of the whole territory

- **Challenge**: Can a fully automated approach provide LC estimates of satisfactory accuracy?

EUROSAT DATASET

P. Helber, B. Bischke, A. Dengel, and D. Borth, Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification (2017), arXiv preprint arXiv:1709.00029.





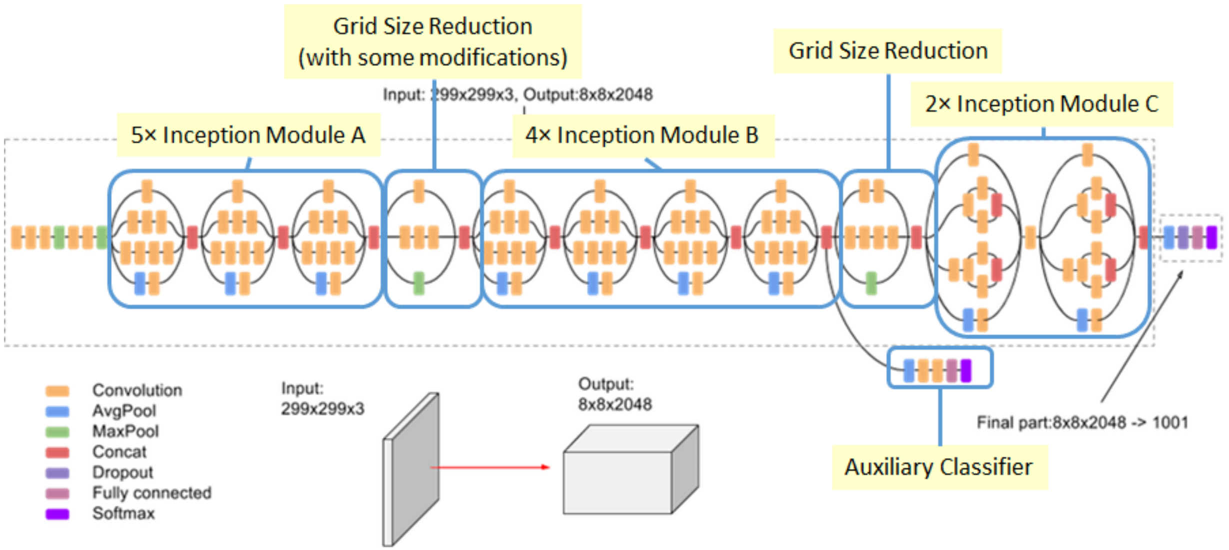
The images have a size of 64x64 pixels.

Each class contains 2,000 to 3,000 images.

The dataset contains 27,000 images

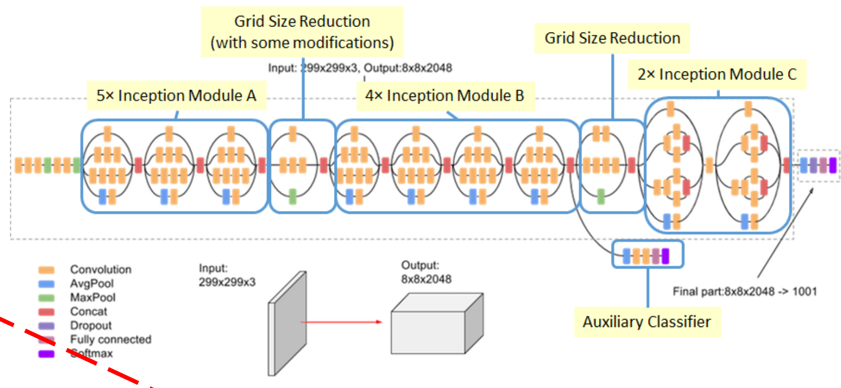
Manually labelled Sentinel-2 images covering 34 European countries

INCEPTION V3

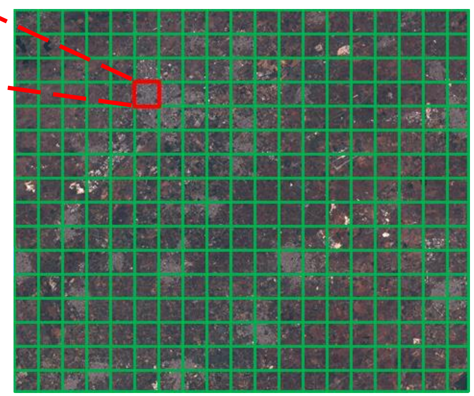


- Number of parameter: 23,851,784
- Implementation on python library Keras
- Transfer learning on Imagenet

TRAINED DEEP CNN



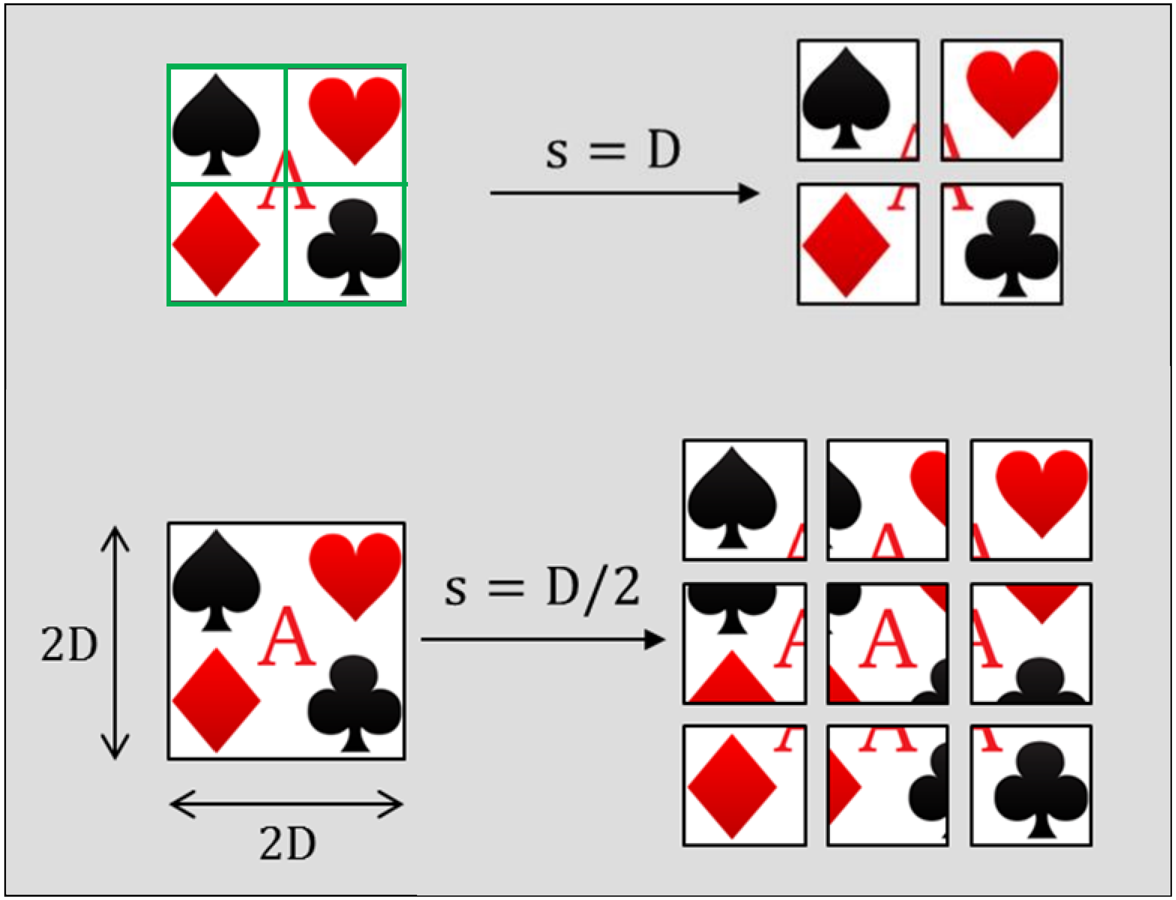
RESIDENTIAL

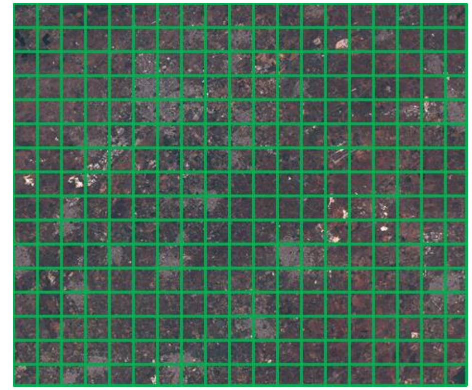


Satellite image (Puglia crop)

CLASS	LAND COVER
…	…
Residential	14.8 %
…	…

HOW TO GENERATE TILES (STRIDING)





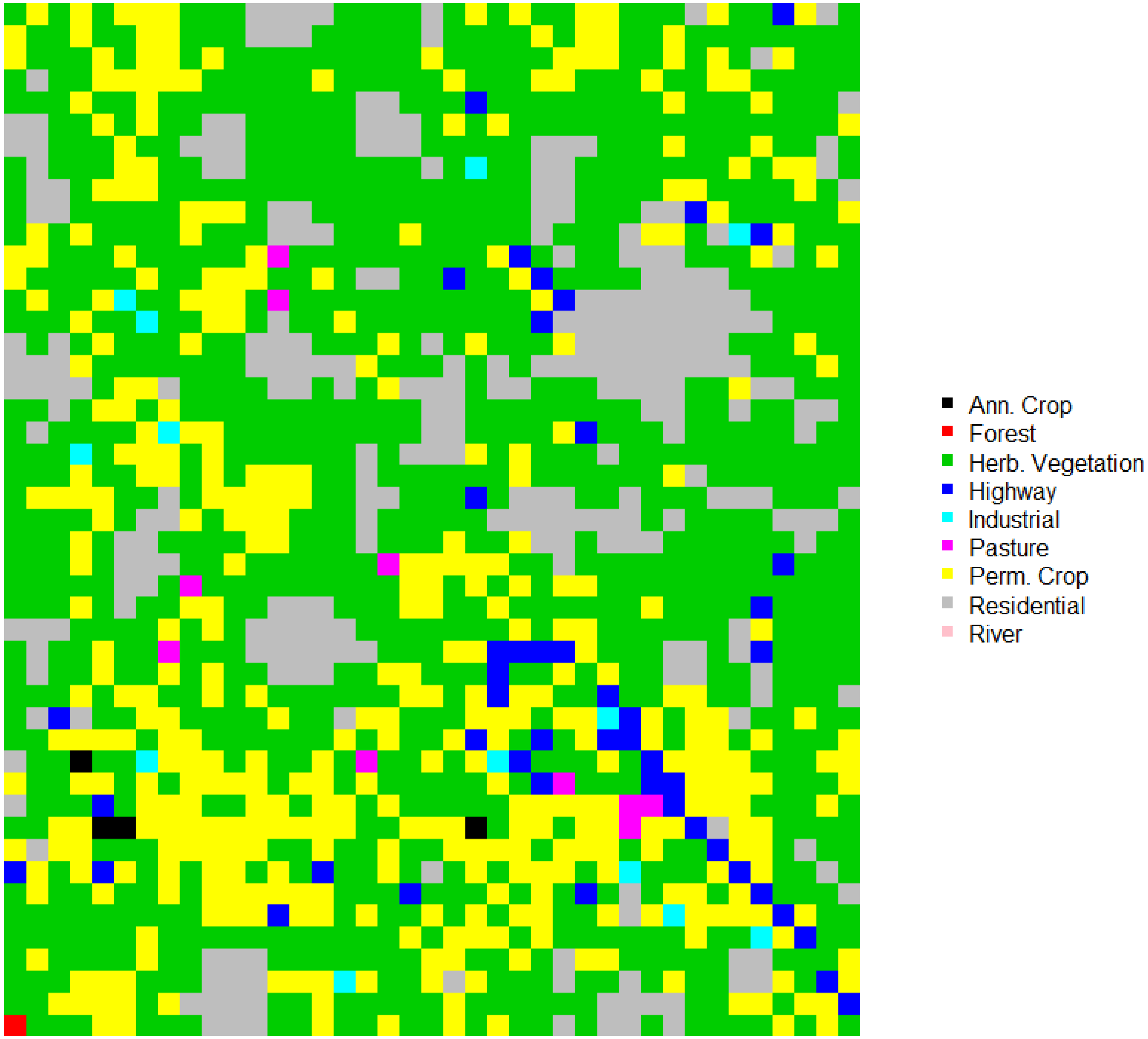
$$N_{rows} = \mathbf{1} + \left\lfloor \frac{D}{s} \right\rfloor \left(\left\lfloor \frac{H}{D} \right\rfloor - \mathbf{1} \right)$$

$$N_{cols} = \mathbf{1} + \left\lfloor \frac{D}{s} \right\rfloor \left(\left\lfloor \frac{W}{D} \right\rfloor - \mathbf{1} \right)$$

Best stride for land cover statistics s = 1

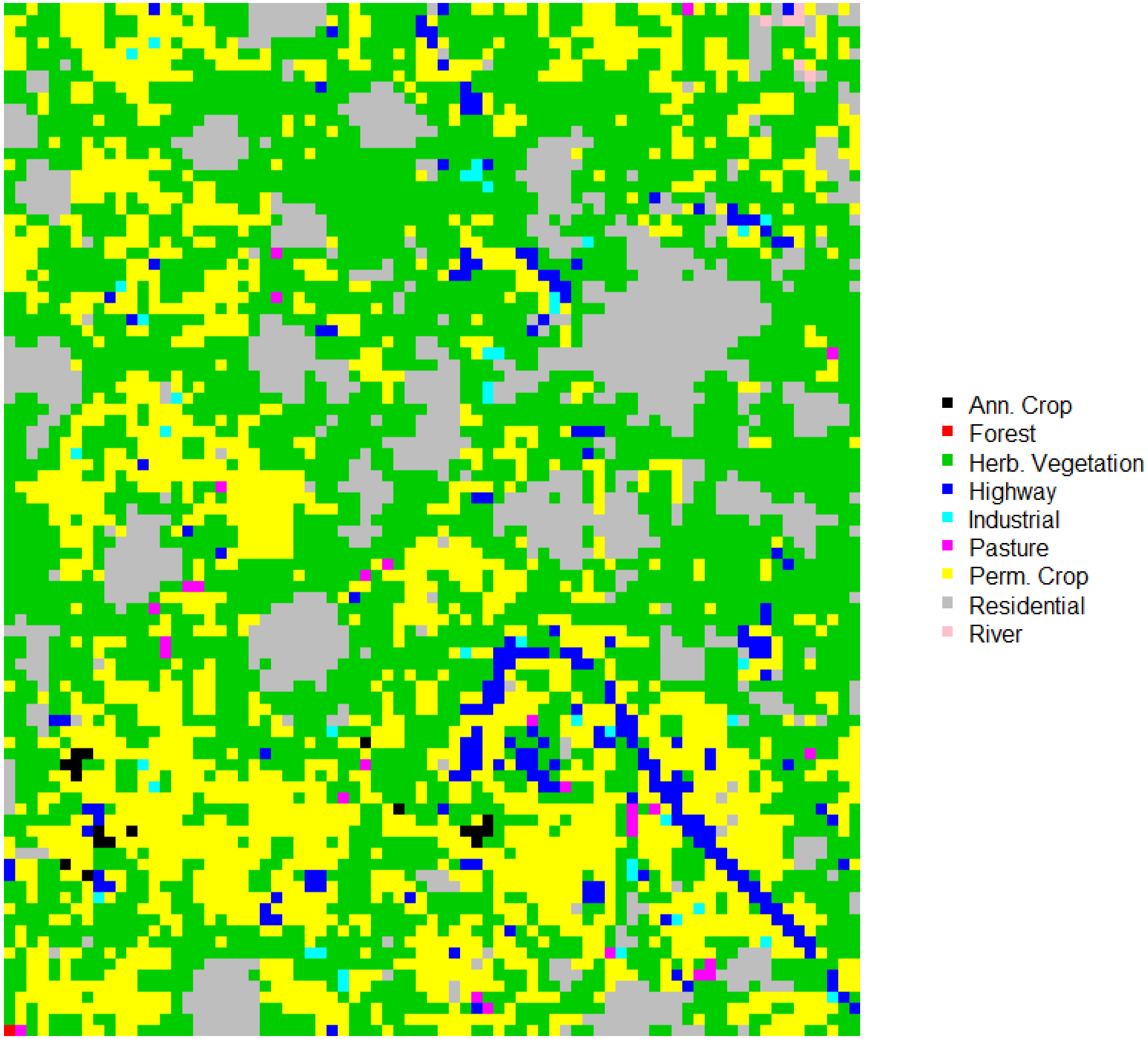
STRIDE

64

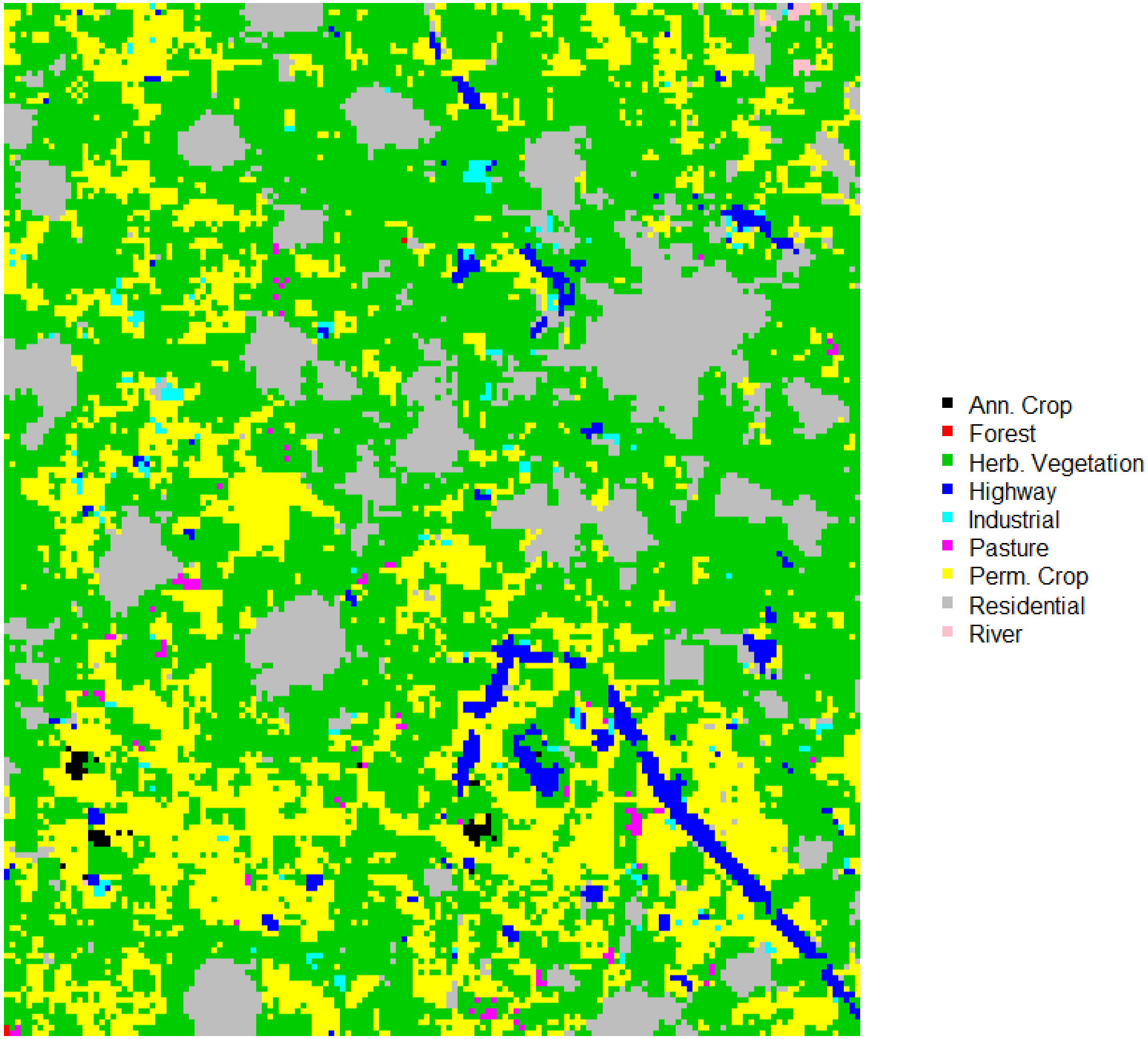

Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

STRIDE

32

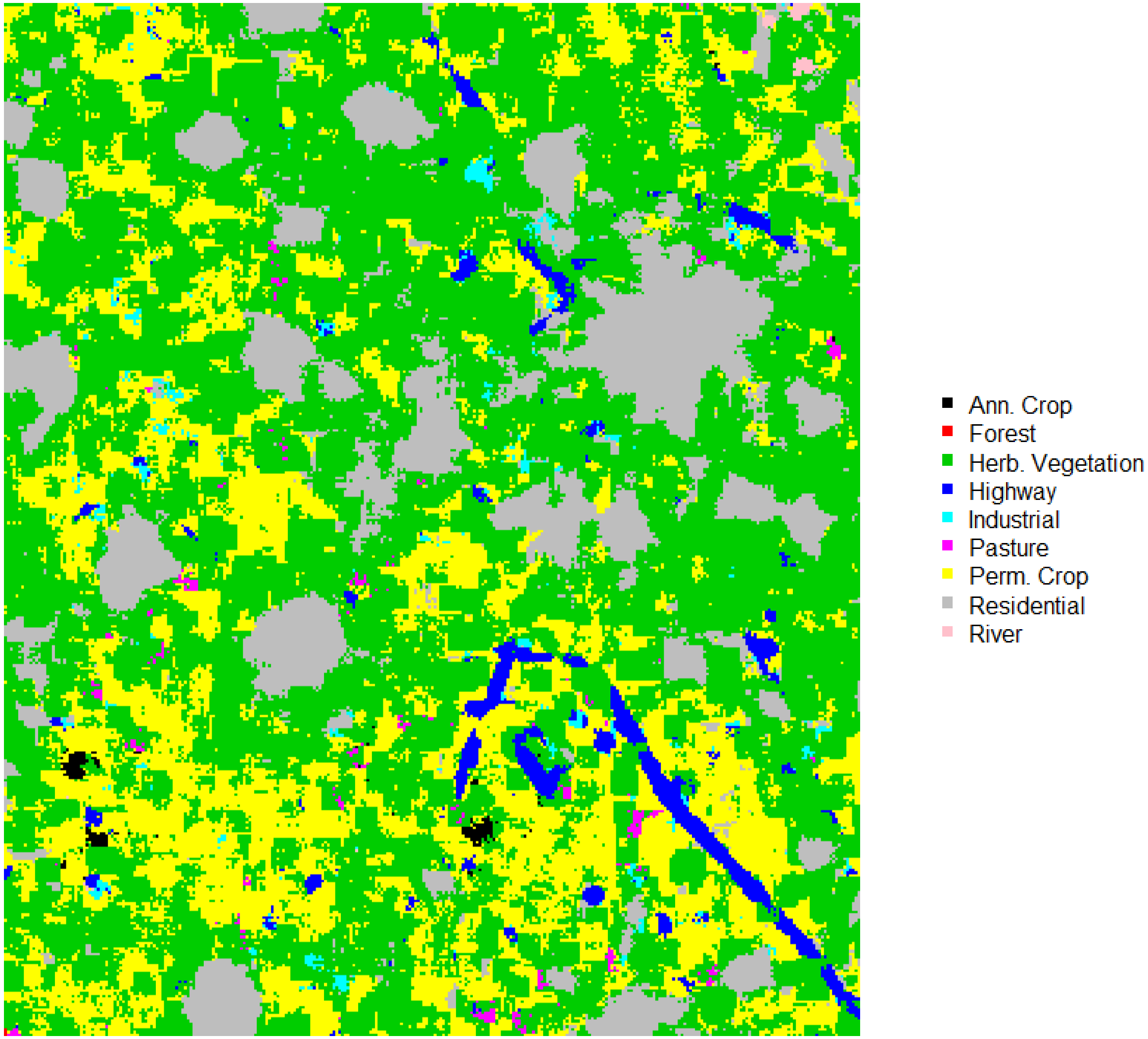

Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

STRIDE
16


Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

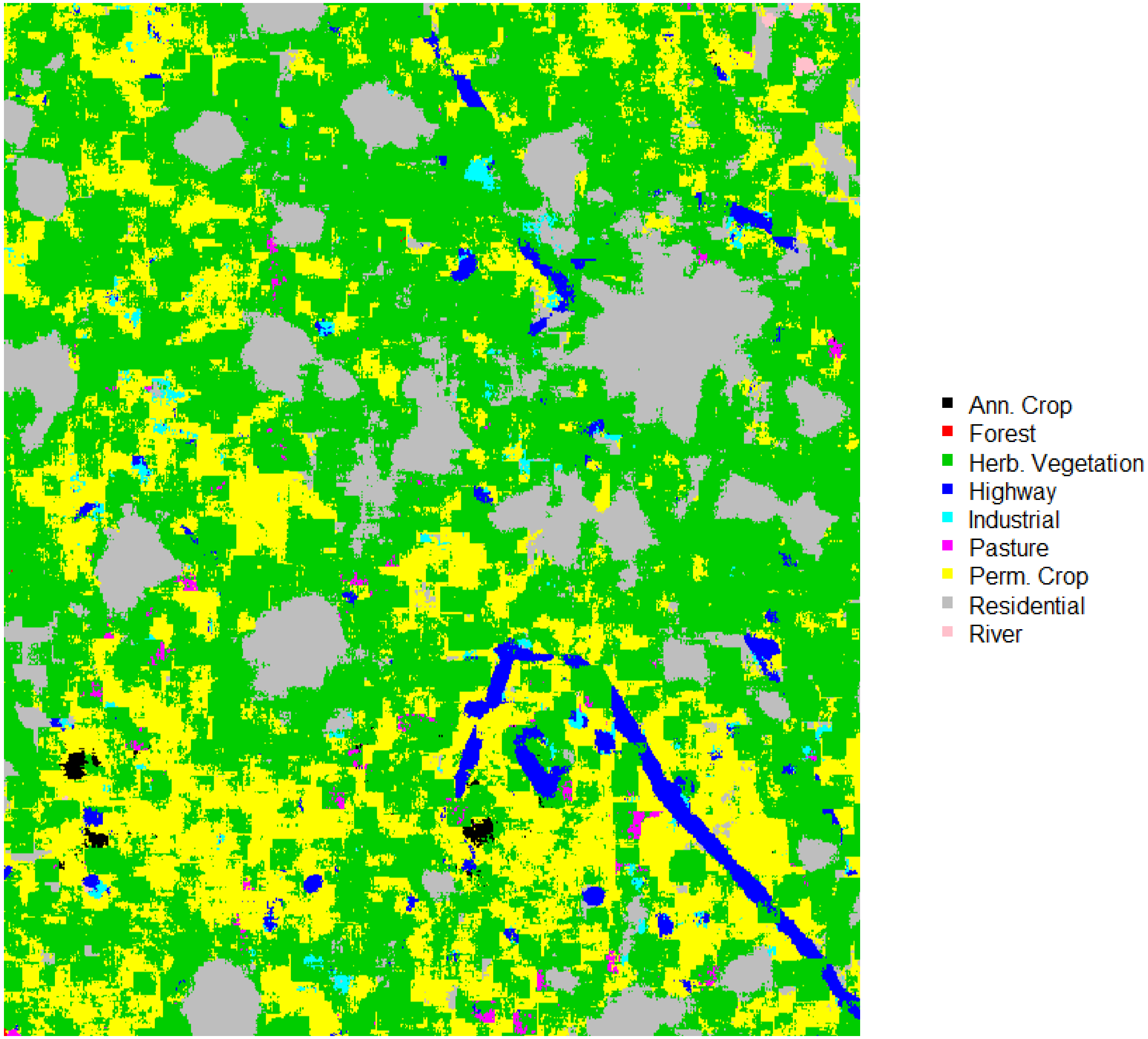
STRIDE

08



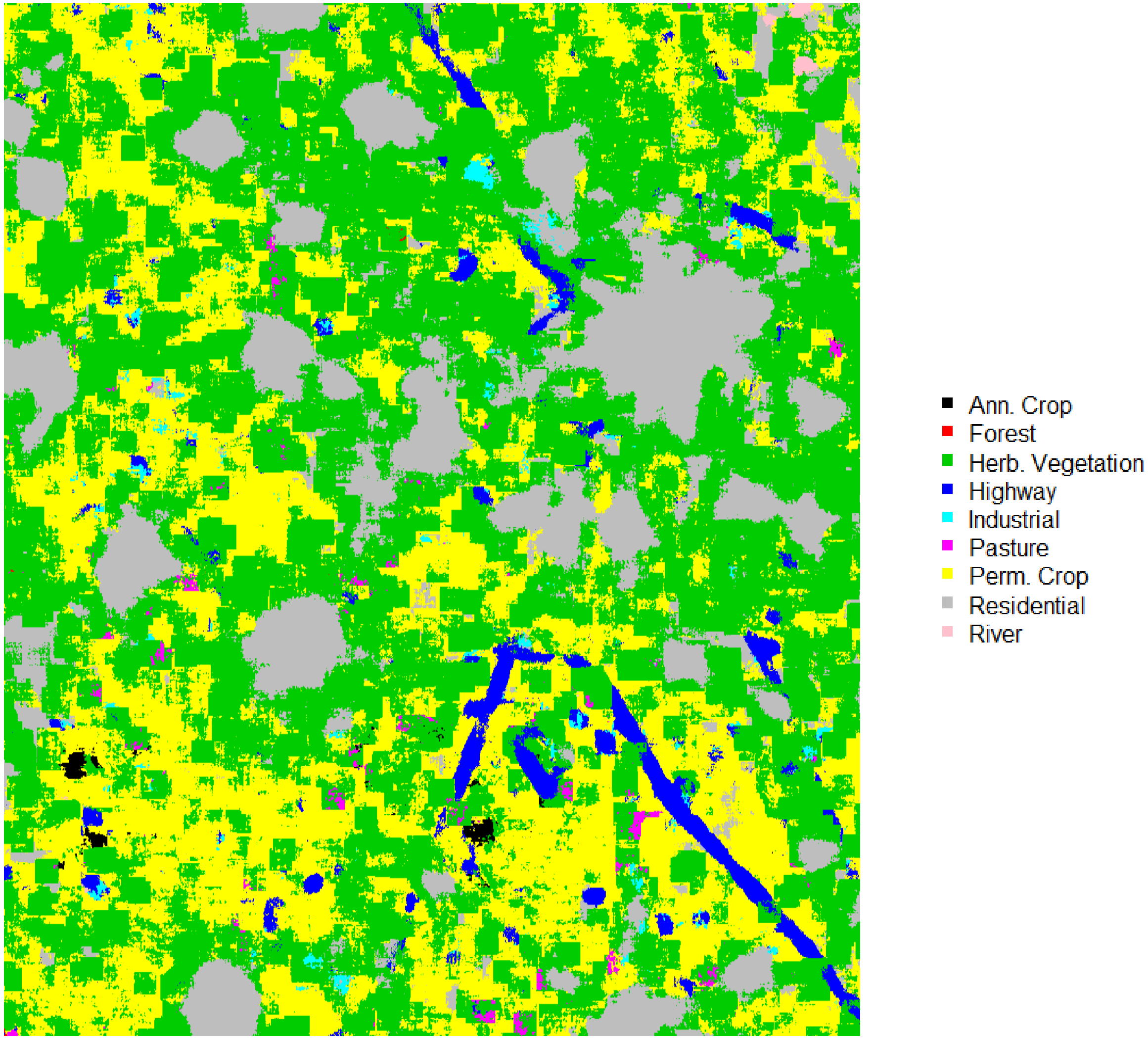
Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

STRIDE
04



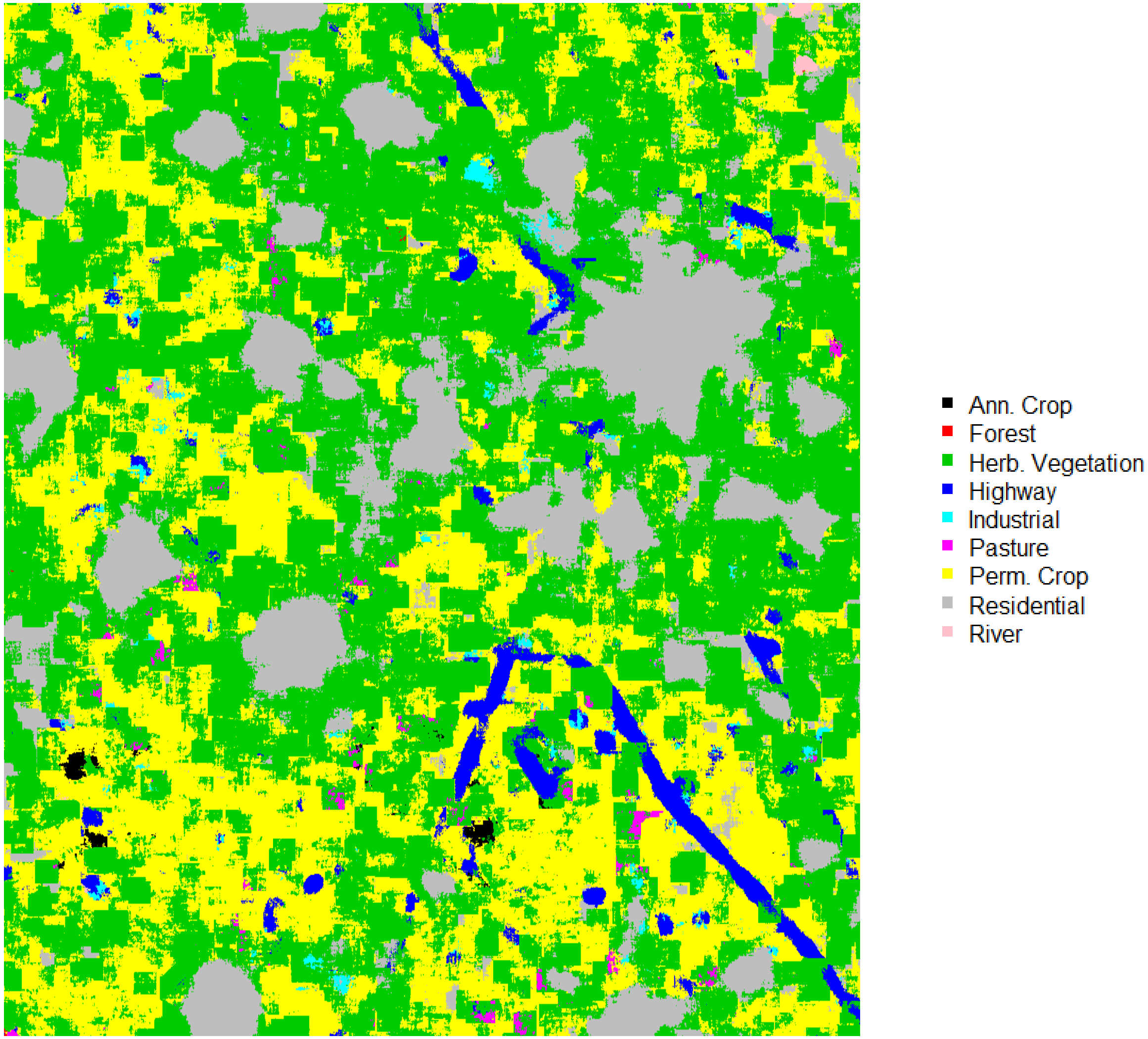
Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

STRIDE
02

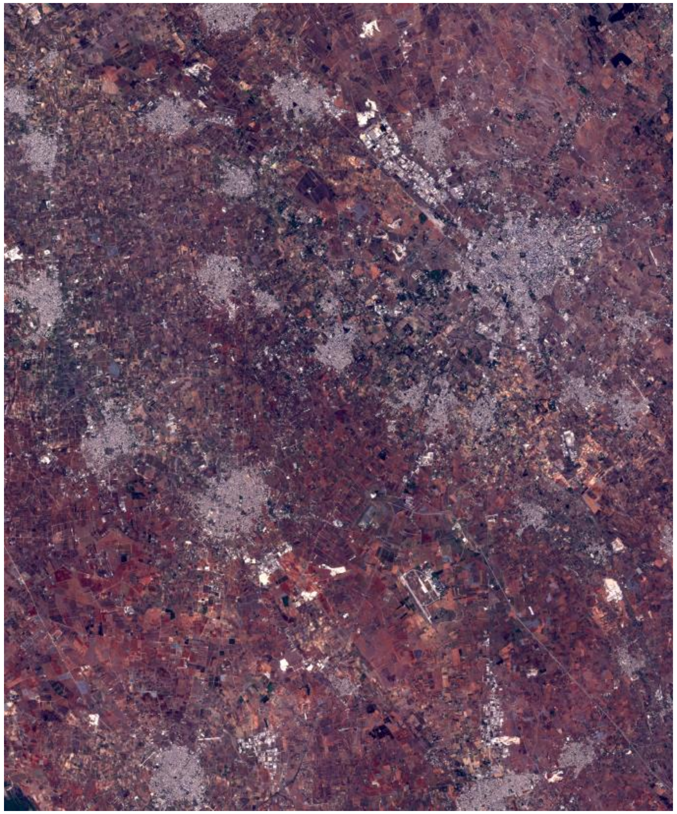

Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

STRIDE

01


Ann. Crop
Forest
Herb. Vegetation
Highway
Industrial
Pasture
Perm. Crop
Residential
River

RESULTS: LAND COVER STATISTICS (Lecce)



LAND COVER (ELC)	AREA SHARE (%)
Annual Crop	0.73
Forest	0.003
Herb. Vegetation	8.36
Highway	4.13
Industrial	4.52
Pasture	0.51
Permanent Crop	72.45
Residential	9.22
River	0.07

Istat

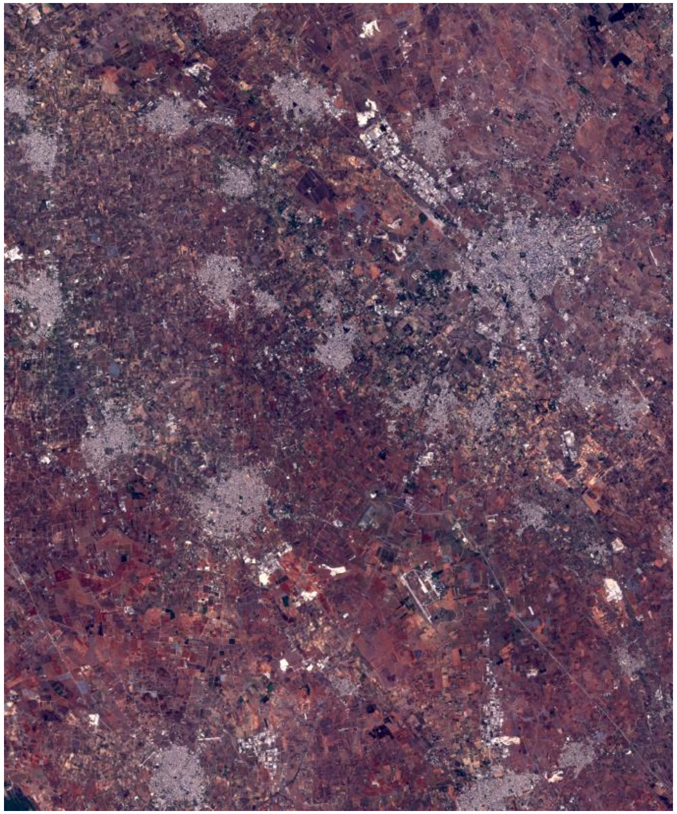
RESULTS: LAND COVER STATISTICS (Pisa)

LAND COVER (ELC)	AREA SHARE (%)
Annual Crop	26.51
Forest	3.16
Herb. Vegetation	9.24
Highway	14.29
Industrial	8.26
Pasture	1.10
Permanent Crop	23.86
Residential	7.03
River	6.54

MAPPING LAND COVER CLASSIFICATIONS

LUCAS	EUROSAT
ARTIFICIAL LAND	Industrial
ARTIFICIAL LAND	Residential
ARTIFICIAL LAND	Highway
BARE LAND & LICHENS/MOSS	*Non contemplata*
WETLANDS	*Non contemplata*
CROPLAND	Annual crop
CROPLAND	Pasture
CROPLAND	Permanent crop
WOODLAND	Forest
SHRUBLAND	*Non contemplata*
GRASSLAND	Herbaceous vegetation
WATER AREAS	River
WATER AREAS	Lake

RESULTS: LAND COVER STATISTICS (Lecce)



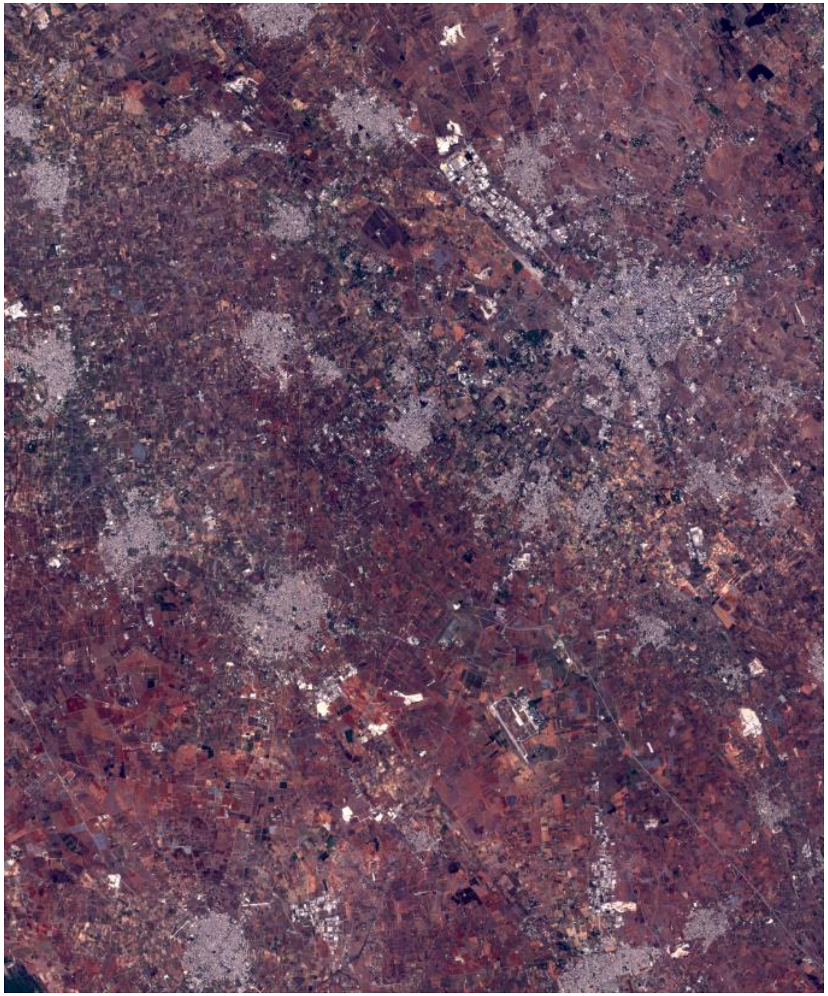
	AREA SHARE (%)	
LAND COVER (LLC)	ESTIMATE	PSEUDO GROUND-TRUTH
Artificial Land	17.88	12.11
Cropland	73.69	87.10
Woodland	0.003	0.22
Shrubland	-	0.002
Grassland	8.36	0.00
Bare Land & Lichens/Moss	-	0.56
Water Areas	0.07	0.001
Wetlands	-	0.008

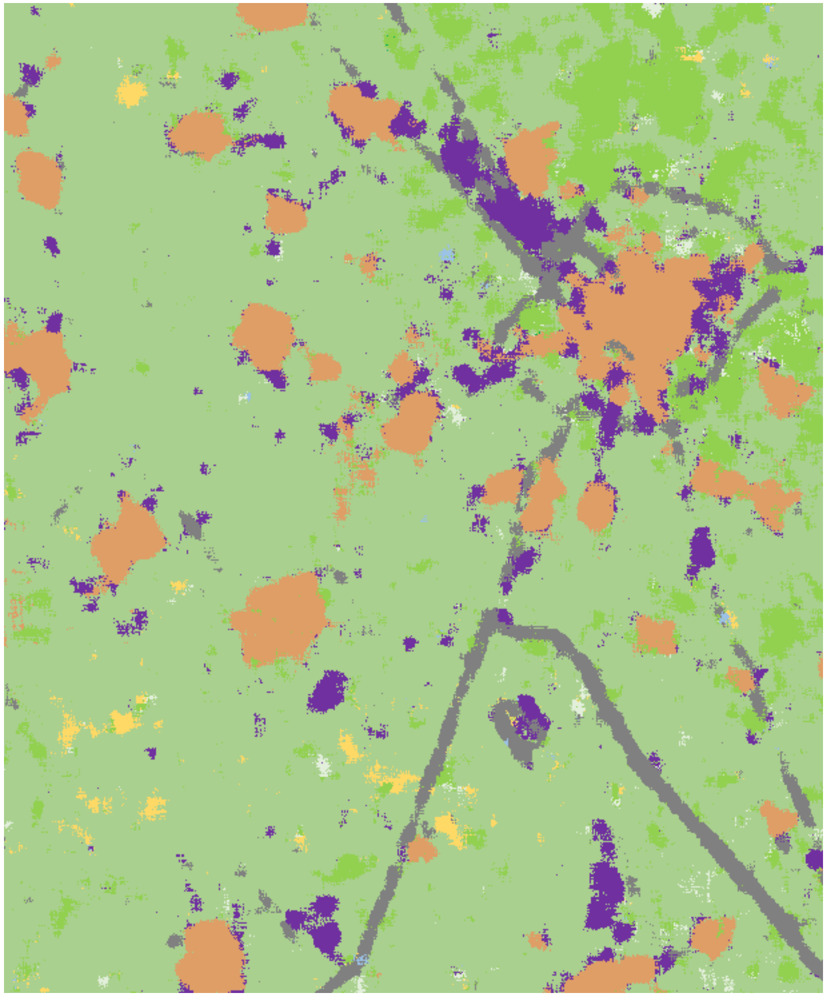
Istat

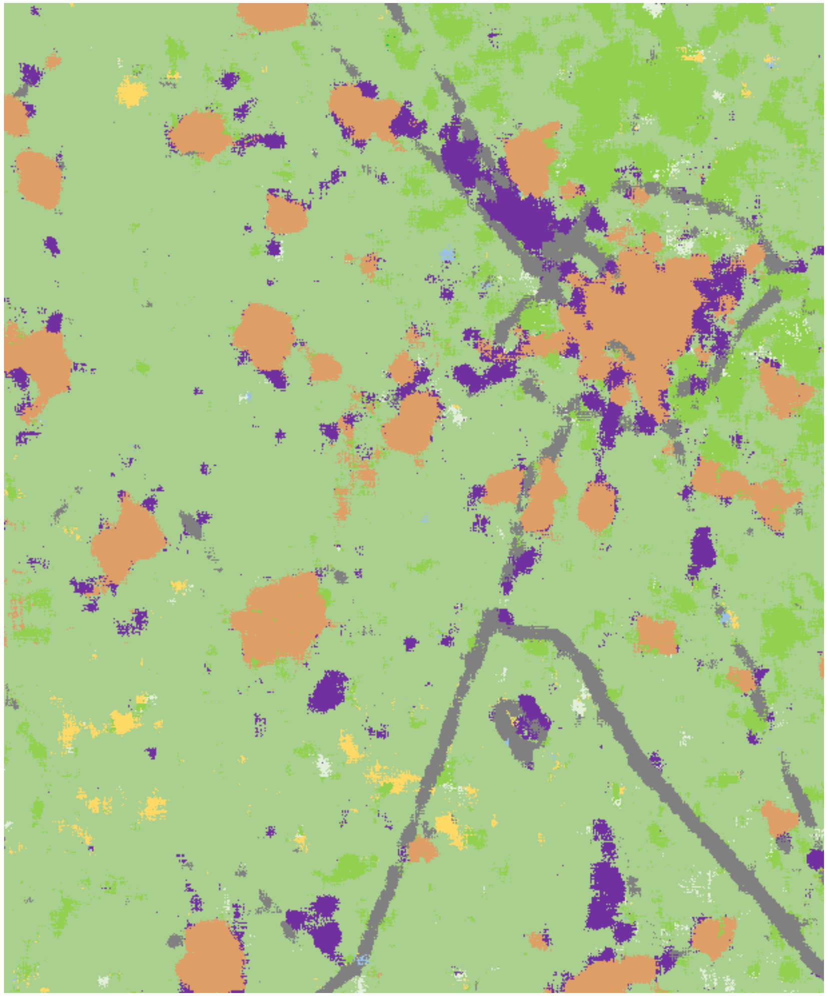
RESULTS: LAND COVER STATISTICS (Pisa)

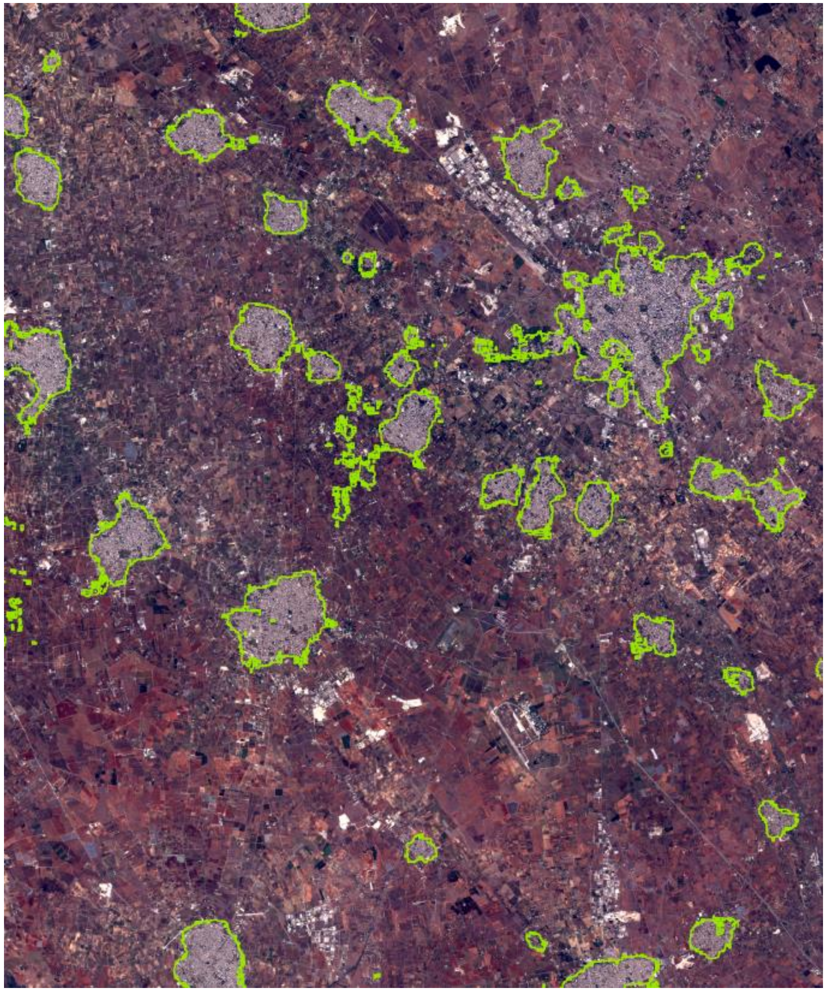
	AREA SHARE (%)	
LAND COVER (LLC)	ESTIMATE	PSEUDO GROUND-TRUTH
Artificial Land	29.59	15.58
Cropland	51.47	66.96
Woodland	3.16	13.35
Shrubland	-	1.11
Grassland	9.24	0.01
Bare Land & Lichens/Moss	-	0.17
Water Areas	6.54	2.25
Wetlands	-	0.57

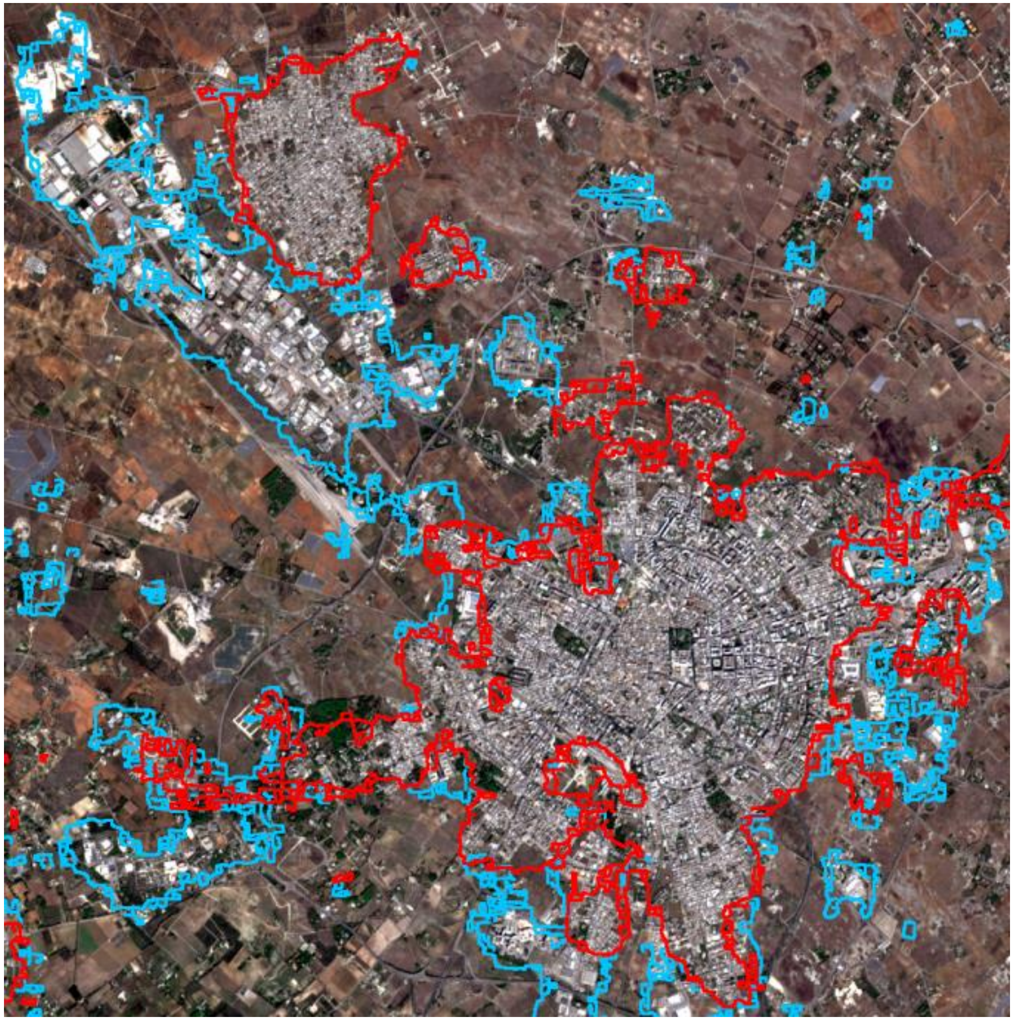
Istat





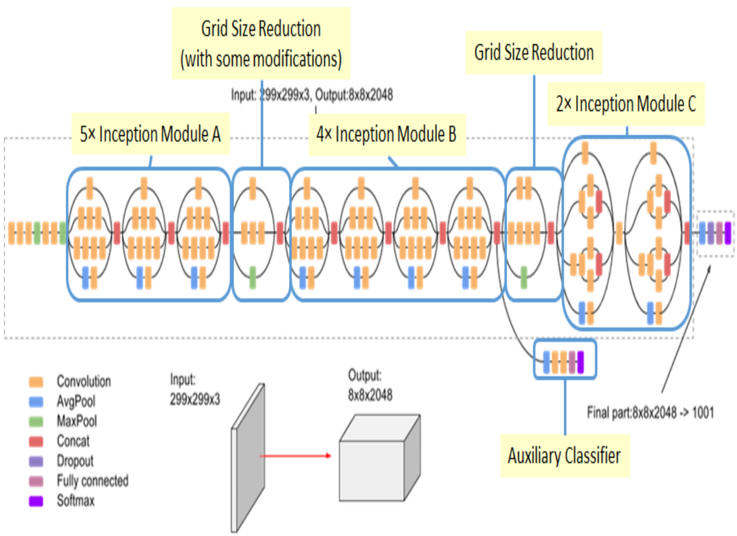


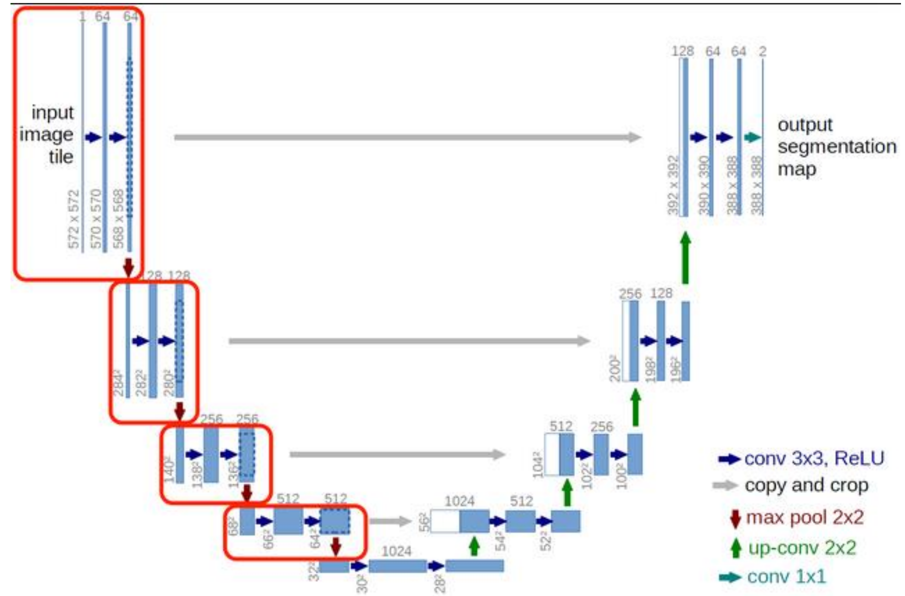


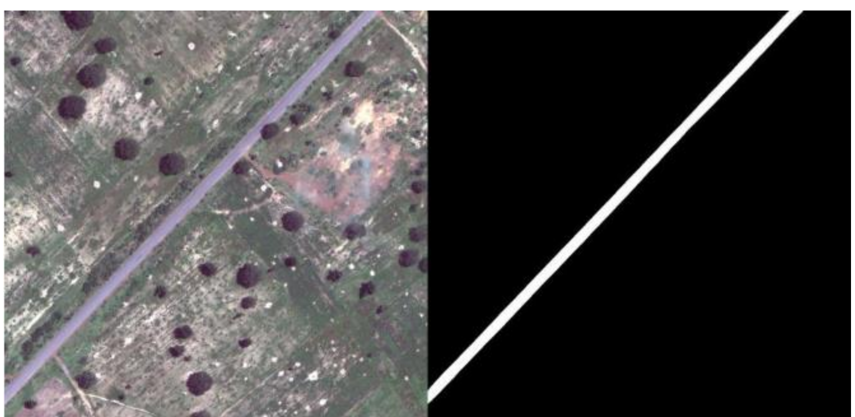


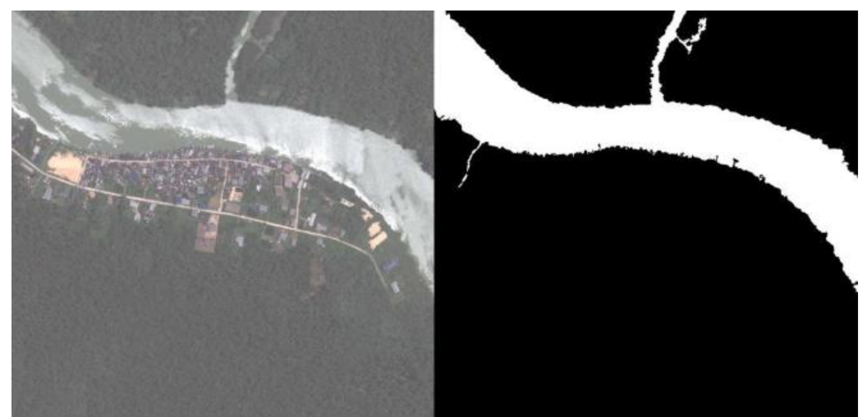
[Google Maps]

Future Research: CNN + UNET









References

B. Bischke, P. Bhardwaj, A. Gautam, P. Helber, D. Borth, and A. Dengel. Detection of Flooding Events in Social Multimedia and Satellite Imagery using Deep Neural Networks. In MediaEval, 2017.

B. Bischke, P. Helber, C. Schulze, V. Srinivasan, and D. Borth. The Multimedia Satellite Task: Emergency Response for Flooding Events. In MediaEval, 2017.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. science, 313(5786), 504-507.

Bernasconi, E., Pugliese., F., Zardetto, D., Scannapieco, M., Satellite-Net: Automatic Extraction of Land Cover Indicators from Satellite Imagery by Deep Learning, In NTTS 2019 proceedings.

Sagiroglu, S., & Sinanc, D. (2013, May). Big data: A review. In 2013 International Conference on Collaboration Technologies and Systems (CTS) (pp. 42-47). IEEE.

Bengio Y, LeCun Y (2007) Scaling learning algorithms towards, AI. In: Bottou L, Chapelle O, DeCoste D, Weston J (eds). Large Scale Kernel Machines. MIT Press, Cambridge, MA Vol. 34. pp 321–360.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-First AAAI Conference on Artificial Intelligence.

Büttner, G. (2014). CORINE land cover and land cover change products. In Land Use and Land Cover Mapping in Europe (pp. 55-74). Springer, Dordrecht.

Istat

Aknowledgements

Thank you for attention.

Fabrizio De Fausti
Francesco Pugliese
Diego Zardetto
Monica Scannapieco