PyData Rome

1st Meeting, 30 November 2022

BinarioF, Rome - Italy

Deep Learning for Land Use and Land Cover Classification

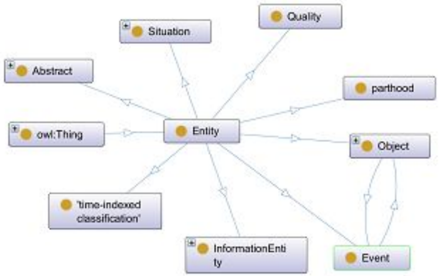
Luigi Selmi

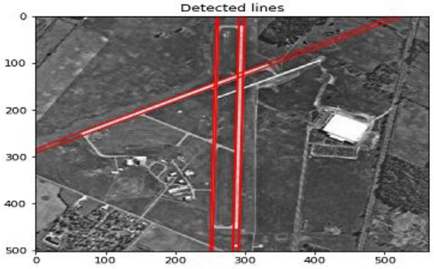
lselmi@datiaperti.it

Luigi Selmi

Working Experience: software engineer in private enterprises and public research institutions in Italy, Switzerland, Germany

Academic Background: Atmospheric Physics and Remote Sensing (Sapienza University of Rome)





Outline

- The Problem
  - Land Use & Land Cover Classification
- The Physics
  - Reflectance spectra
- Copernicus
  - Sentinel-2 and the MSI Instrument
  - Open Access Hub and Copernicus DIAS
- Deep Learning
  - Deep Learning architectures
  - The EuroSAT dataset
  - Data augmentation techniques
  - Fine-tuning
  - Model evaluation
  - One more validation test
- Conclusion

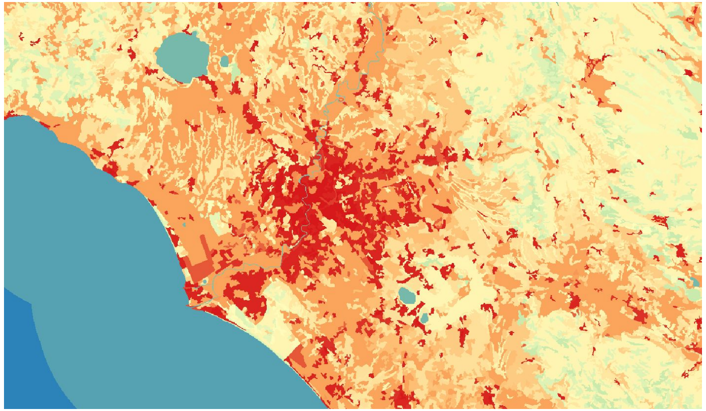
Land Use / Land Cover Classification

Goals:

- Resource planning and monitoring
- Change detection

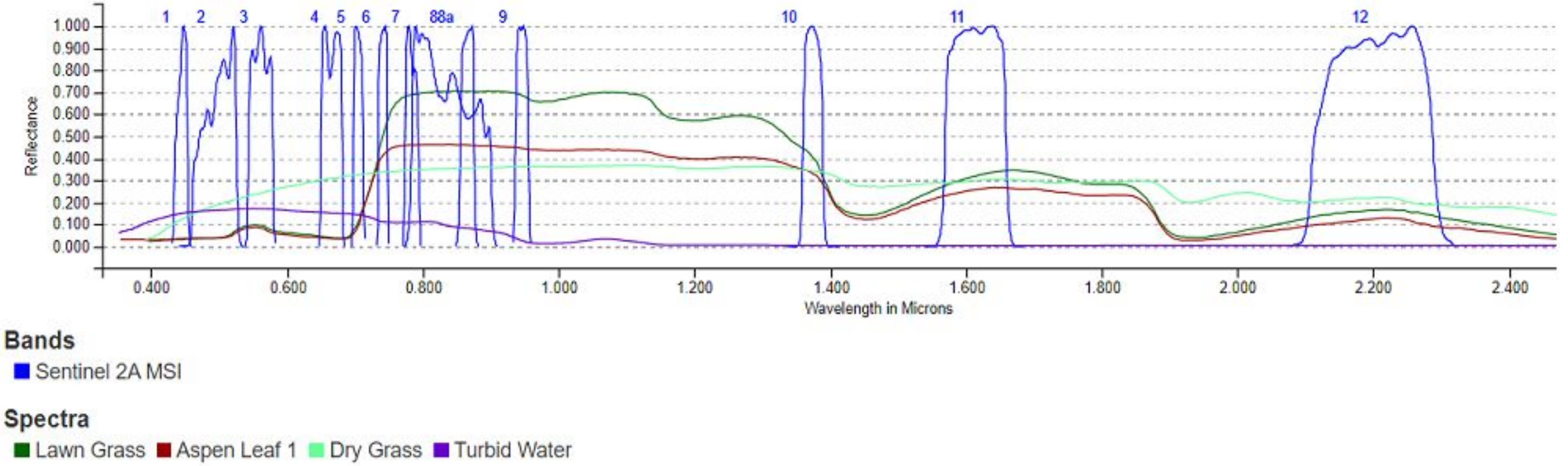
Example:

- Corine Land Cover (44 classes, 3 levels)



Corine 2018 100m

Reflectance spectra



Copernicus Sentinel-2 and the MSI Instrument

Constellation of 2 satellites in polar orbit at 786 km altitude.

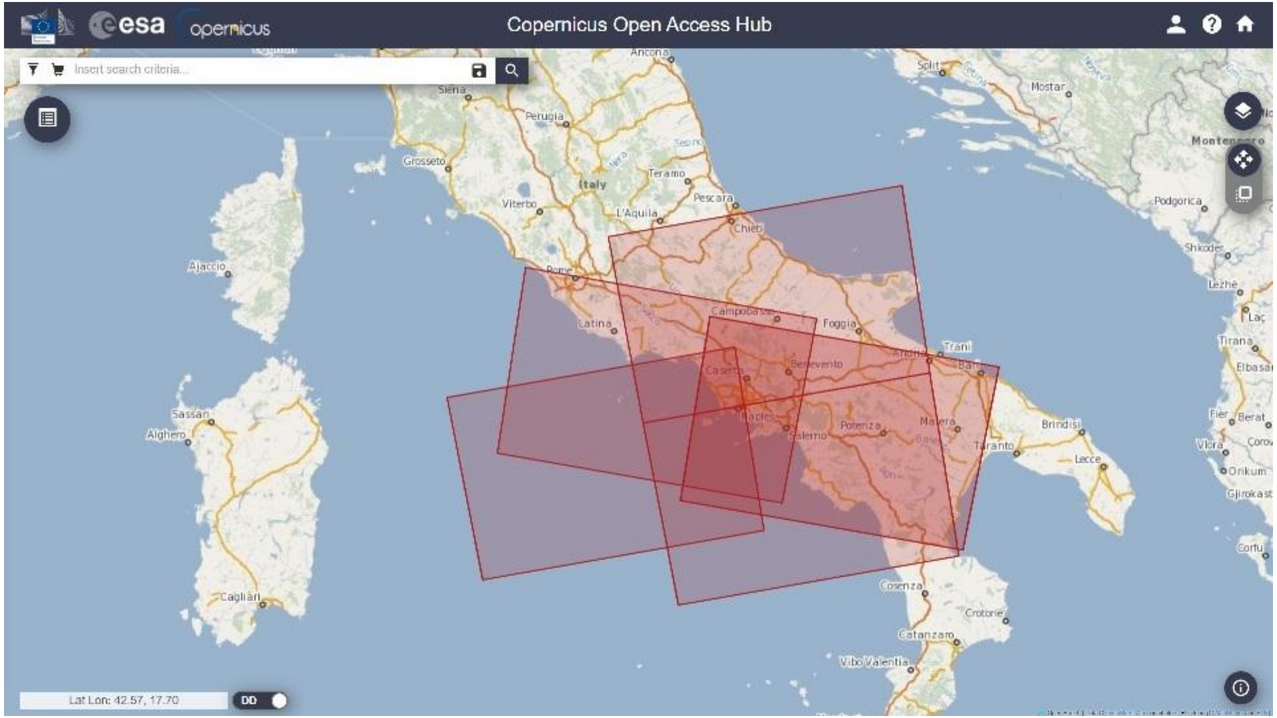
Revisit time: 5 days

Swath: 290 km

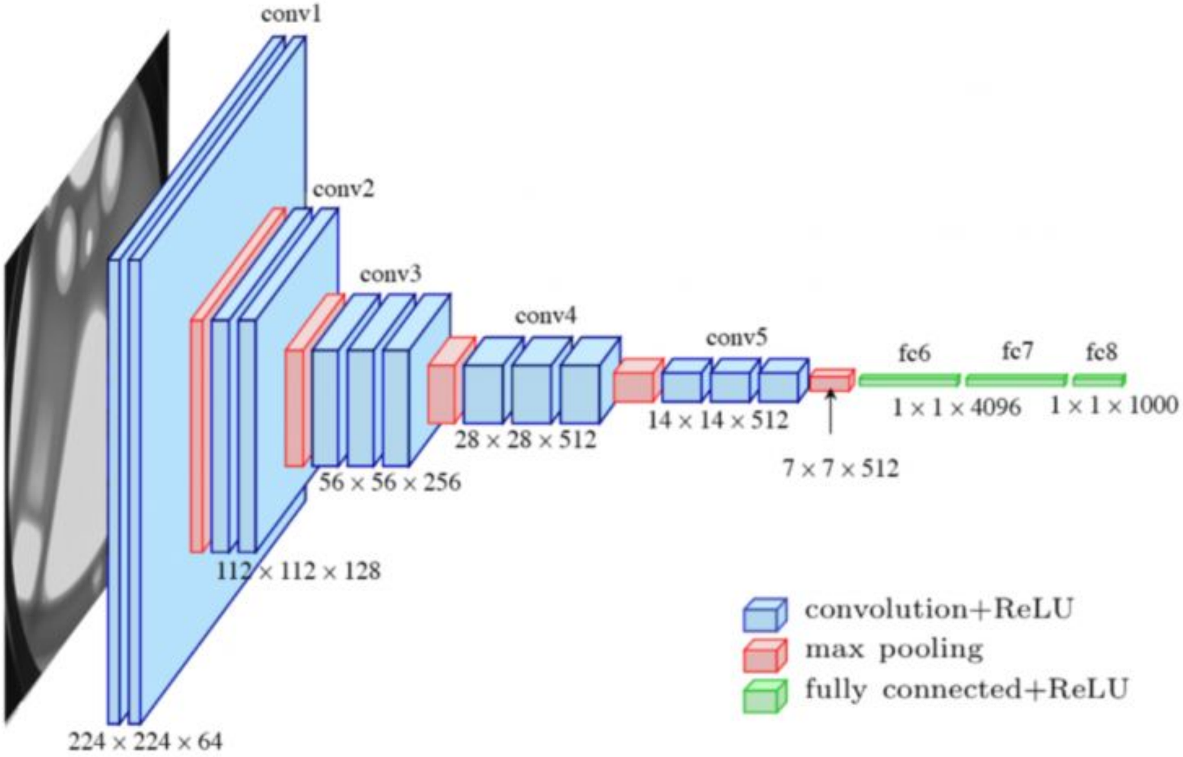
Multi-Spectral Imager with 13 bands from VIS to SWIR

Res.: 10m RGB (60m for other bands)

Copernicus Open Access Hub and DIAS



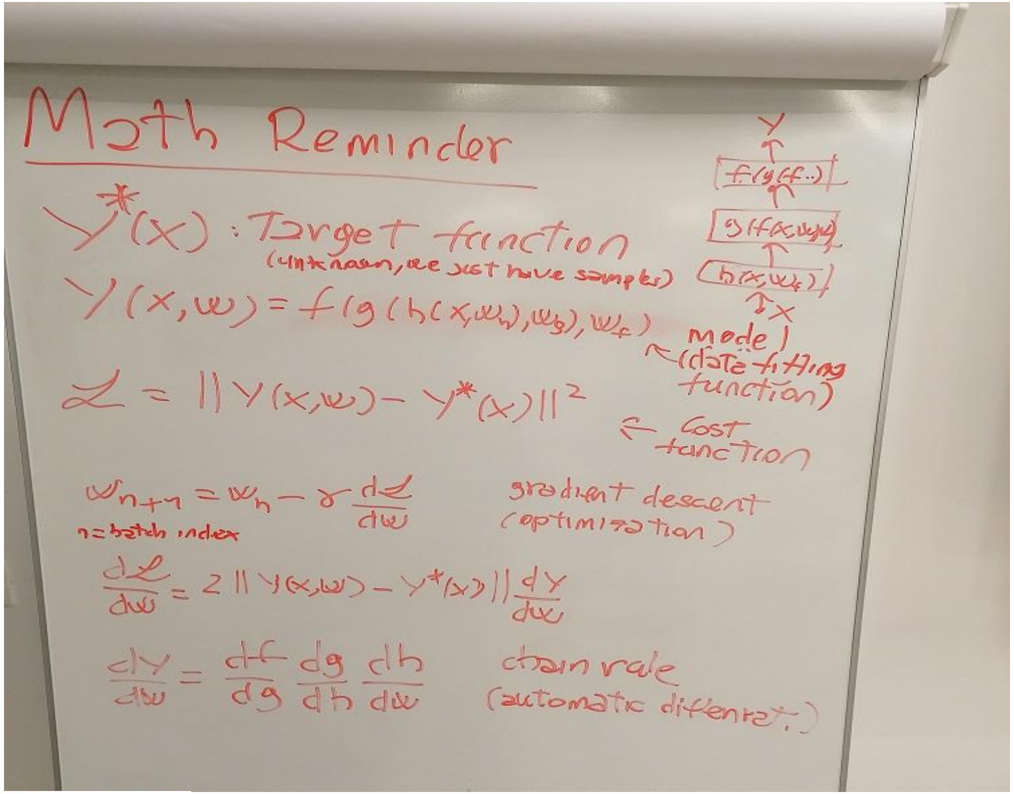
Deep Learning Architectures



Deep Learning Math

DL Math is not that difficult. What matters is the implementation.

PyTorch uses automatic differentiation to compute the gradient of the cost function



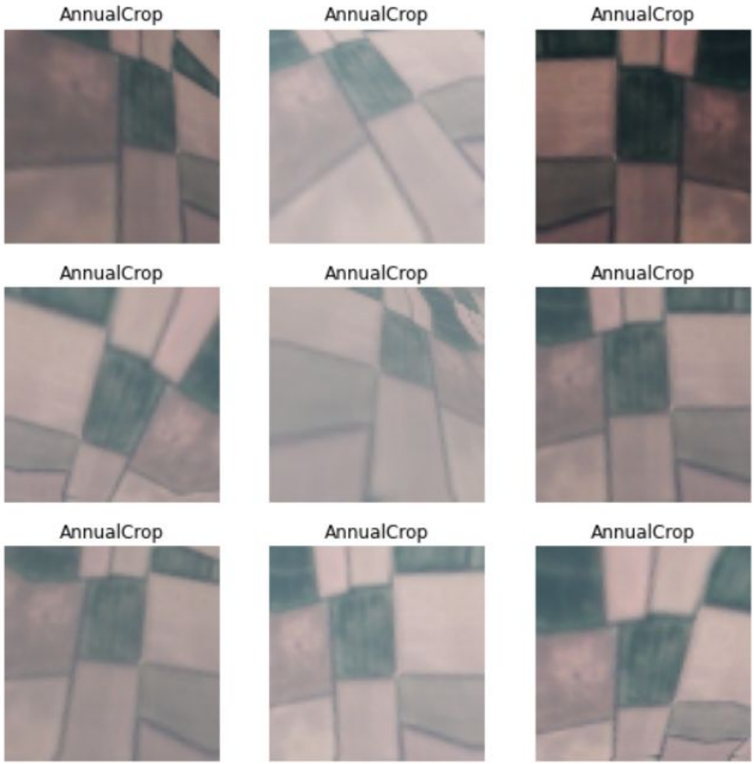
The EuroSAT dataset

10 Classes: Forest , Pasture, Herbaceous Vegetation, Annual Crop, industrial, Residential, Permanent Crop, Highway, Sea Lake, River

Patches from Sentinel-2 imagery: 27000 64x64 RGB images, 2000-3000 per class

Data Augmentation

Affine transformations + Noise



Fine-tuning the ImageNet dataset

The model is trained over the ImageNet dataset, a set of 1.200.000 images from 1000 classes.

The model is fine-tuned using the much smaller EuroSAT images by “freezing” all the model parameters but the last layers.

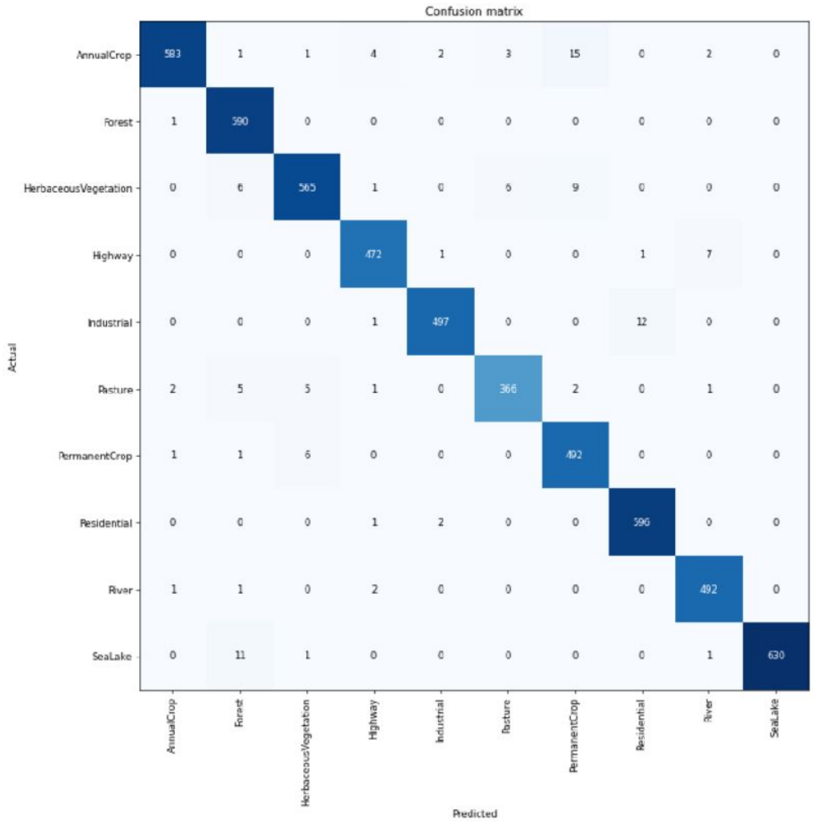
Only 2 millions parameters to be learnt for each epoch instead of more than 25 millions.

Model evaluation

Training set: 80%
Validation set: 20%

Misclassifications:

Annual Crop, Permanent Crop: 15
Industrial, Residential: 12
Sea Lake, Forest: 11
Herbaceous Vegetation, Permanent Crop: 9
Highway, River: 7
Herbaceous Vegetation, Forest: 6
Herbaceous Vegetation, Pasture: 6
Permanent Crop, Herbaceous Vegetation: 6
Pasture, Forest: 5
Pasture, Herbaceous Vegetation: 5



Additional tests

July 2019

Prediction: Highway; Probability: 0.9980

[4.8410e-04, 1.3392e-06, 1.5634e-04, 9.9796e-01, 1.9808e-04,
1.3962e-04, 8.8622e-04, 4.6300e-05, 1.1766e-04, 1.2730e-05]

Aug. 2021

Prediction: Industrial; Probability: 0.7301

[2.8591e-04, 3.4131e-06, 1.4427e-04, 2.6680e-01, 7.3014e-01,
3.1067e-05, 1.4392e-04, 3.6543e-04, 2.0304e-03, 5.0807e-05]

Conclusion

- Satellite imagery is a relevant resource for LULC classification tasks
- A CNN ResNet with “only” 50 layers provides a good accuracy
- A LULC map can be produced quickly in a cost-effective way
- The result may be improved by using all the 13 MSI bands, Sentinel-1 SAR images and a larger dataset for fine-tuning.

Blog post with link to the GitHub repository:

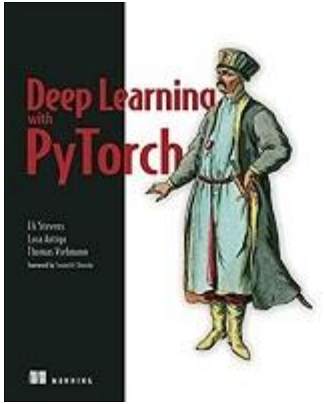
https://www.luigiselmi.eu/eo/lulc-classification-deeplearning.html

References

Links to many free online textbooks about Remote Sensing and Deep Learning are available on my website

https://www.luigiselmi.eu/bookshelf

Other books on the market are:



Thanks for your attention!

Email: lselmi@datiaperti.it

Personal website:

https://www.luigiselmi.eu/

Company website:

https://www.datiaperti.it/