VHR LAND COVER MAP OF ROME OBTAINED USING A CITIZEN SCIENCE APPROACH

Fabio Del Frate, Zina Mitraka, Giovanni Schiavon
(University of Tor Vergata, Rome, Italy)

Francesco Carbone
(GEO-K SRL, Rome, Italy)

INTRODUCTION

- It has been shown that advanced automatic algorithms may give satisfactory results but they have to be driven by specialized personnel (technician or scientists) and a manual interaction is often required to reach the desired accuracies, especially for VHR imagery

What about….

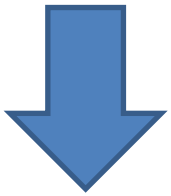
- Perform classifications in a effective way using the powerful and unique human brain ability to recognize patterns, which is superior to any algorithm

- Engage and reach out to the public in a two-ways relationship, providing a new type of education, enabling CS and the public to learn more about science but also contributing to it.

INTRODUCTION

very promising but….

1. We cannot ask the volunteers to manually paint millions of pixels



2. We have to provide them effective, user friendly and free SW packages to perform a significant part of the work automatically (also having some fun..)

3. Moreover we have provide the tools to check the quality of the produced results

EUMAPPER

Effective

- learn complex patterns, taking into account any nonlinear complex relationship between the explicative and the dependent variables
- No a-priori knowledge about probabilistic distributions required. NN learn the statistical relationships directly from the data

"….We must conclude by emphasizing that, similar to the 2007 contest, NNs have provided the best individual performances…"

from the Data Fusion Contest organized by the Data Fusion Technical Committee (DFTC) of the Geoscience and Remote Sensing Society (GRSS) in 2008

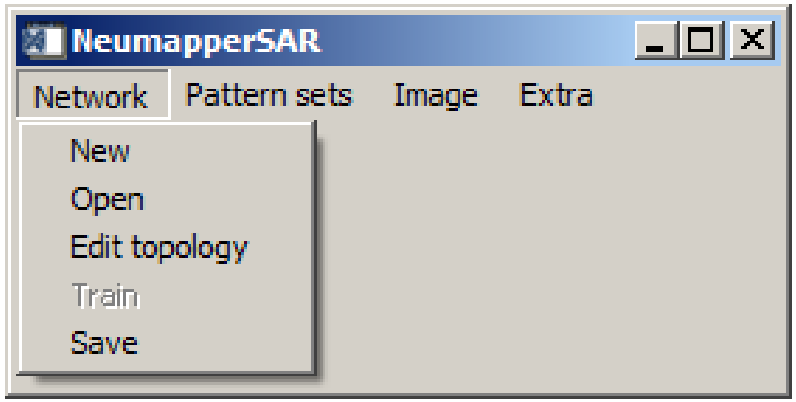
EUMAPPER

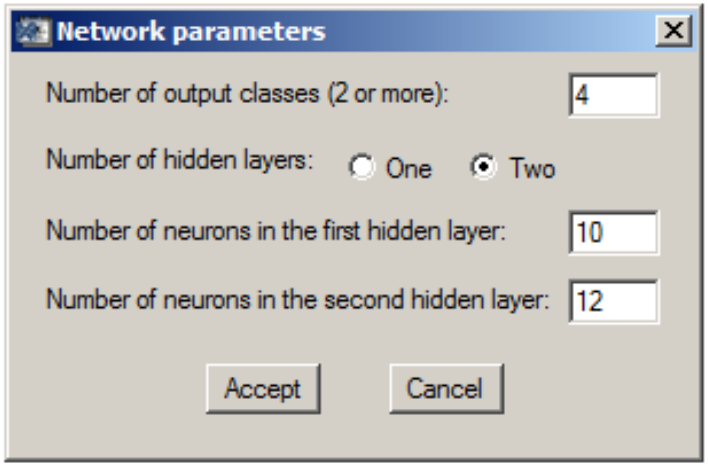
User Friendly

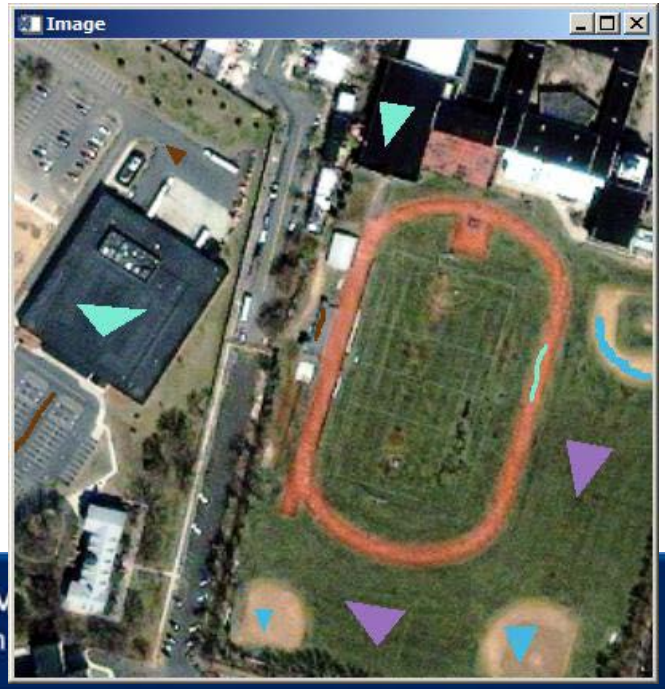
The tool implements in the same environment the various stages in the generation of an artificial neural network (ANN) computational model for automatic pixel-based SAR-image classification. Most of the complicated issue about NN design are masked to the user who has only to do with

a) generation of training data
b) definition of a network topology
c) training phase
d) application of the trained net to the entire image

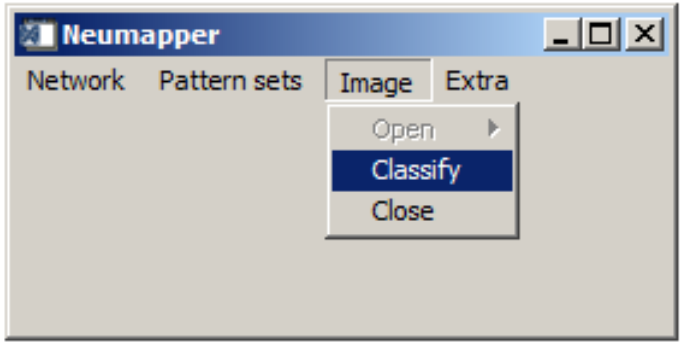
EUMAPPER


NeumapperSAR
Network
Pattern sets
Image
Extra
New
Open
Edit topology
Train
Save


Network parameters
Number of output classes (2 or more):
4
Number of hidden layers:
One
Two
Number of neurons in the first hidden layer:
10
Number of neurons in the second hidden layer:
12
Accept
Cancel



Image


Neumapper
Network
Pattern sets
Image
Extra
Open
Classify
Close

EUMAPPER

Free Download from www.geo-k.co

QUALITY CHECK

The CS is provided a complete manual of instructions

CLASSIFY WITH NEUMAPPER

Input: sample.tif

Output: sample_classified.jp2

Use Neumapper to produce a classification image. Mind to use the predefined color combinations for your classes, shown below.

Bare Soil	→ BS	→ Yellow	(R:255	G:255	B:0)
Water Surface	→ WS	→ Blue	(R:0	G:8	B:255)
Green Vegetation	→ GV	→ Green	(R:4	G:255	B:0)
Buildings	→ BU	→ Red	(R:255	G:0	B:0)
Asphalted Surface	→ AS	→ Purple	(R:87	G:0	B:141)

To ensure that you have these colors in your classification maps you have to choose the following colors in Neumapper:


Bare Soil


Water Surface


Green Vegetation


Buildings


Asphalted Surface

QUALITY CHECK

CLASSIFICATION POST-PROCESSING

Input: sample _classified.bmp

Output: sample _classified_mod.bmp

You have to post-process your classification result by 'correcting' some pixels that were misclassified from Neumapper. To do so, you have to cross-check the resulting image and if you see some misclassified pixels, you have to open you image with Paint and change the color to the correct one. Keep the original Neumapper output as well.

Note:
Make sure you use the same colors as above!

CLASSIFICATION ACCURACY

Input: sample _classified.bmp, sample.txt, sample_0.npy etc.

Output: (the confusion matrix)

To estimate the accuracy of your classification you have to estimate the confusion matrix. To do so:

- Run the script by typing

```
validatePoints.exe <imagename> <classifimage> <groundtruth> -a <attempt>
```

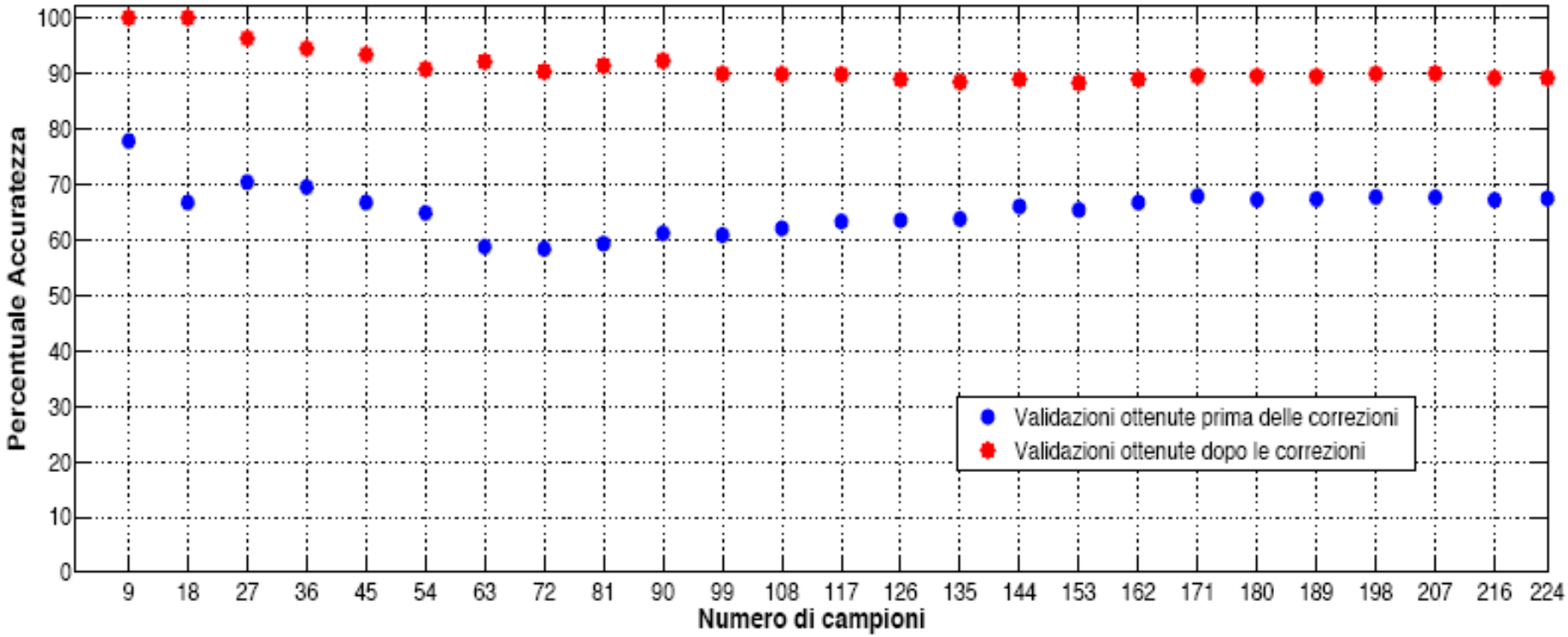
- for example for the sample image:

```
validatePoints.exe sample.tif sample_class.bmp sample.txt -a 1
```

This will estimate the confusion matrix for you which is the number of point validated correctly according to your ground truth.

QUALITY CHECK

QUALITY CHECK



THE PROJECT

- Citizen Scientists: mostly "Remote Sensing" students of the Tor Vergata University
- Task: classify a 2000 x 1000 pixels image (no time constraint)
- Classes: water, bare soil, vegetation, asphalt, buildings
- Total number of images: 405 (most of the Rome urban area) → 70 km^2
- Image characteristics: spatial resolution 0.3 m, orthofoto 2011

THE PROJECT

Essential Deliverables for the CS:

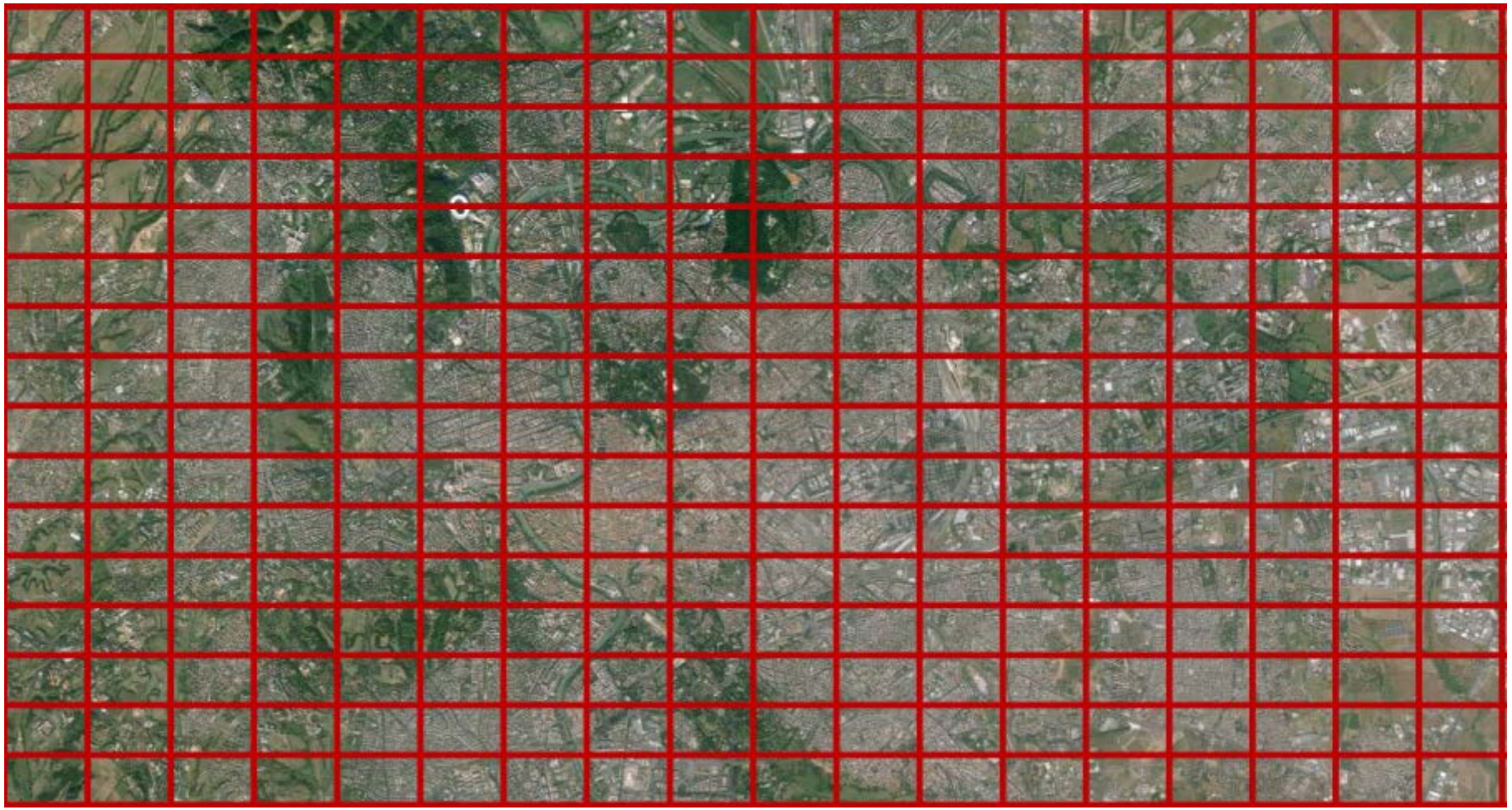
1. *Automatically classified image*
2. *Corrected image*
3. *Accuracies (for both automatic and manual results)*

Optional Deliverables for the CS:

1. *Training Patterns*
2. *Network model*
3. *Short technical report*

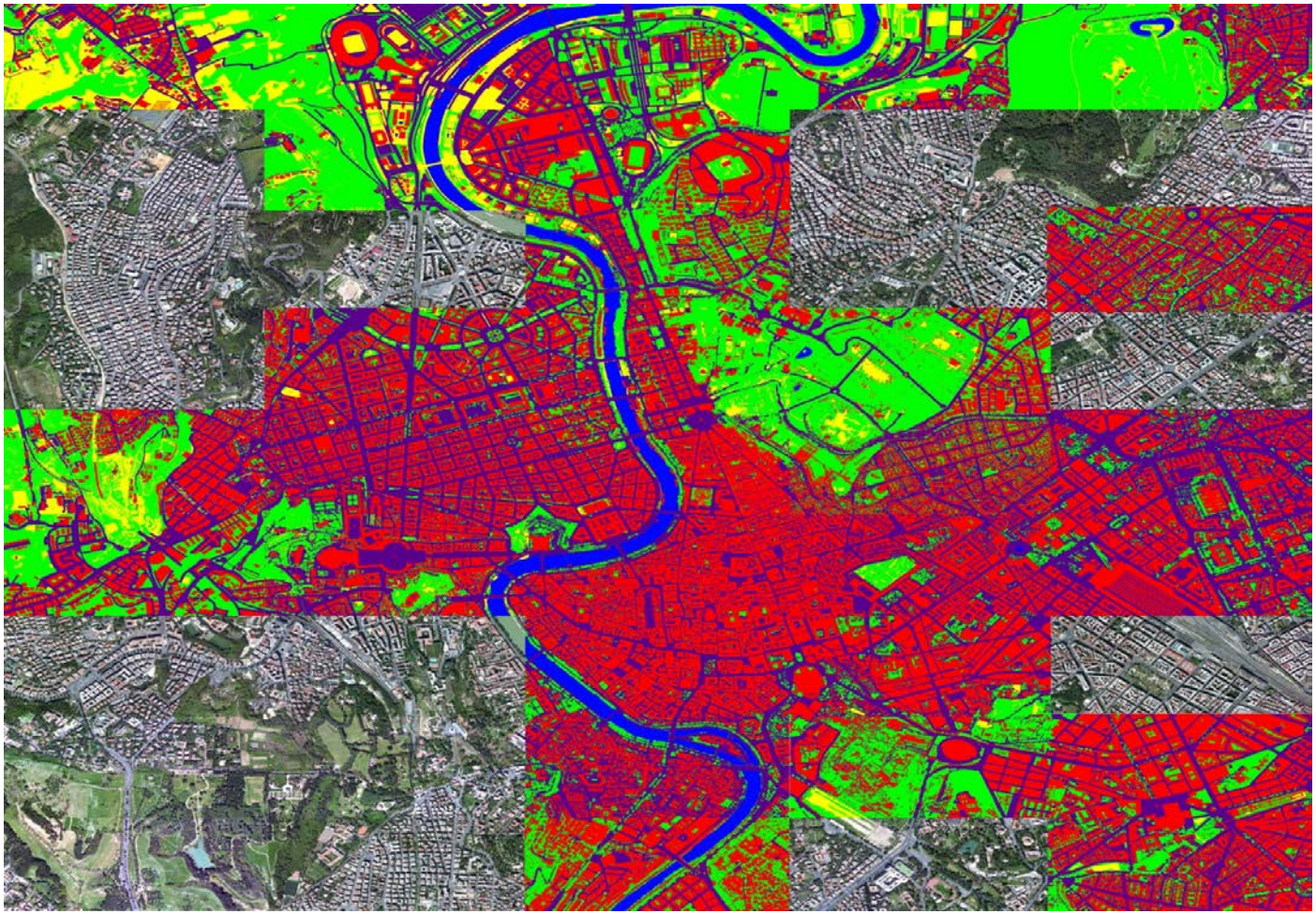
- On-line help assistance has been set-up

THE PROJECT



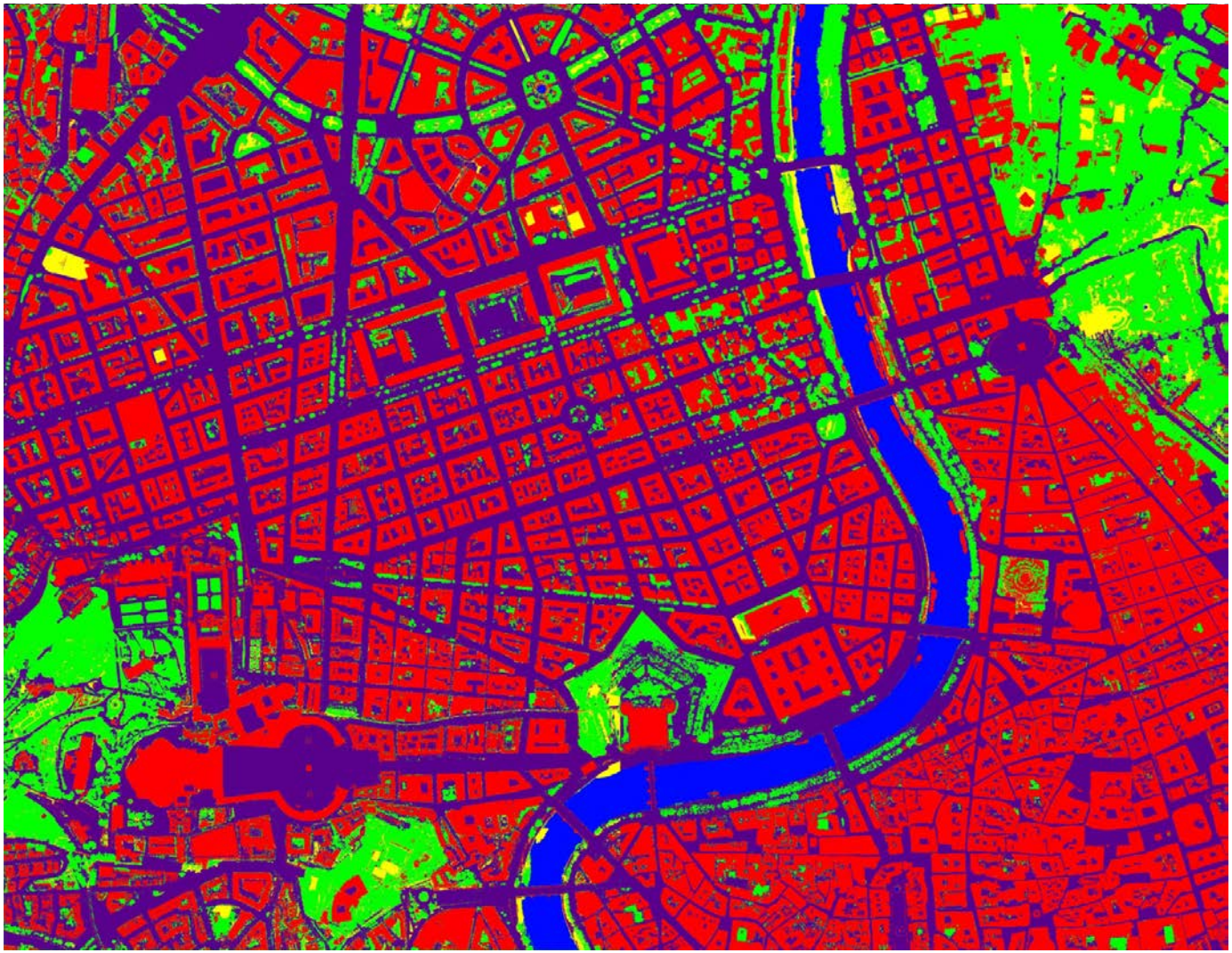
The CS is invited to choose the area where she/he leaves

THE PROJECT

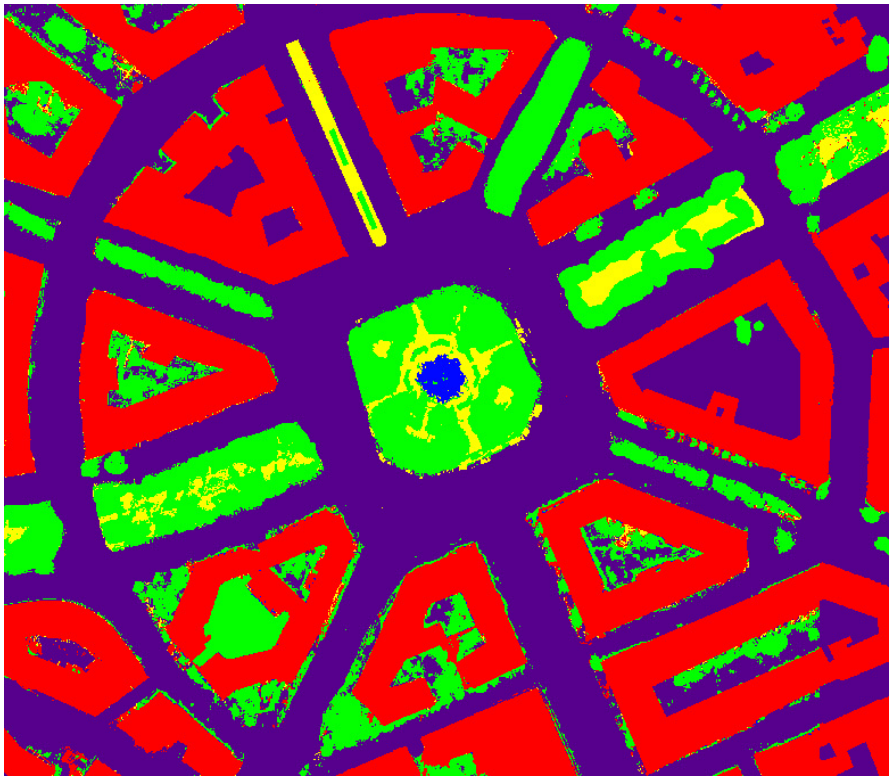


Current Status: 261/405

THE PROJECT



THE PROJECT



RESULTS

- At the moment about 45 km^2 classified with more than 95% of accuracy
- Some corrections needed when mosaicing the tiles
- Main ambiguities stem from the courtyards which are inside the buildings, sometime classified as asphalted areas instead of built areas
- Other (scientific) outputs:

Automatic vs Manual
Stability on NN parameters

CONCLUSIONS

- A very detailed land cover map of Rome urban area is being produced (hopefully completed by the end of this year) using a CS approach
- Partial results have been already exploited for validation activities in scientific studies
- Good feedbacks from the CS
- Plan to extend the community but we need to find proper motivations