Tor Vergata Activities in Satellite Image Classification

F. Pacifici, F. Del Frate, D. Solimini,
G. Licciardi, M. Lazzarini, C. Pratola

f.pacifici@disp.uniroma2.it

Earth Observation Laboratory, DISP, Tor Vergata University, Rome, Italy

Rome, Italy - May 12, 2008

In the last decades Tor Vergata University Earth Observation Laboratory has gained significant experience in remote sensing.

- HIGH RESOLUTION IMAGES AND LAND COVER CLASSIFICATION
- HYPERSPECTRAL IMAGES AND LAND COVER CLASSIFICATION
- URBAN AREAS: CLASSIFICATION AND CHANGE DETECTION
- CLASSIFICATION AND MONITORING OF CROPS
- RADAR POLARIMETRY AND LAND COVER CLASSIFICATION
- RETRIEVAL OF FOREST BIOMASS
- POLINSAR FOR FOREST BIOMASS RETRIEVAL
- MICROWAVE SCATTERING FROM SOIL
- MICROWAVE SCATTERING FROM VEGETATION
- MICROWAVE EMISSION AND SMOS
- RETRIEVAL OF SOIL MOISTURE
- MONITORING WETLAND
- NEW TECHNOLOGIES: BISTATIC RADAR
- SAR INTERFEROMETRY FOR TECTONICS
- SAR MTI
- SAR DETECTION AND MONITORING OF OIL SPILLS
- REMOTE SENSING OF THE ATMOSPHERE: OZONE
- REMOTE SENSING OF THE ATMOSPHERE: TEMPERATURE AND WATER VAPOR
- NEURAL NETWORKS FOR EARTH OBSERVATION
- EUROPEAN RADAR-OPTICAL RESEARCH ASSEMBLAGE (ERA-ORA)

In particular, relevant results have been obtained in the use of neural networks algorithms for classification and retrieval problems applied to remotely sensed data (both SAR and optical imagery).

Goals

1. Automatic classification of urban areas with high and very high resolution optical imagery (Landsat/QuickBird)
2. Automatic classification of urban areas with high resolution SAR imagery (ERS/ENVISAT)
3. Multi-scale textural analysis of panchromatic imagery for urban land-use (QuickBird/WorldView-1)
4. Change detection of urban areas with very high resolution optical imagery (QuickBird)
5. Change detection of urban areas with aerial optical imagery
6. WorldView-1 colorization

Automatic classification of urban areas with high and very high resolution optical imagery (Landsat/QB)

GeoInformazione

TOR VERGATA

Image Information Mining (IIM)

In recent studies, besides assessing the performance of MLP for the classification of high and very high resolution images, we investigated on the **generalization capabilities** of Neural Networks (NNs) with the purpose of using them as a tool for **fully automatic classification** of large archives of satellite images:

→ NNs potentialities from an Image Information Mining (IIM) point of view. In particular, applications to urban area monitoring have been addressed.

An artificial neural network can be viewed as a mathematical model composed of many nonlinear computational elements, named neurons, operating in parallel and massively connected by links characterized by different weights.

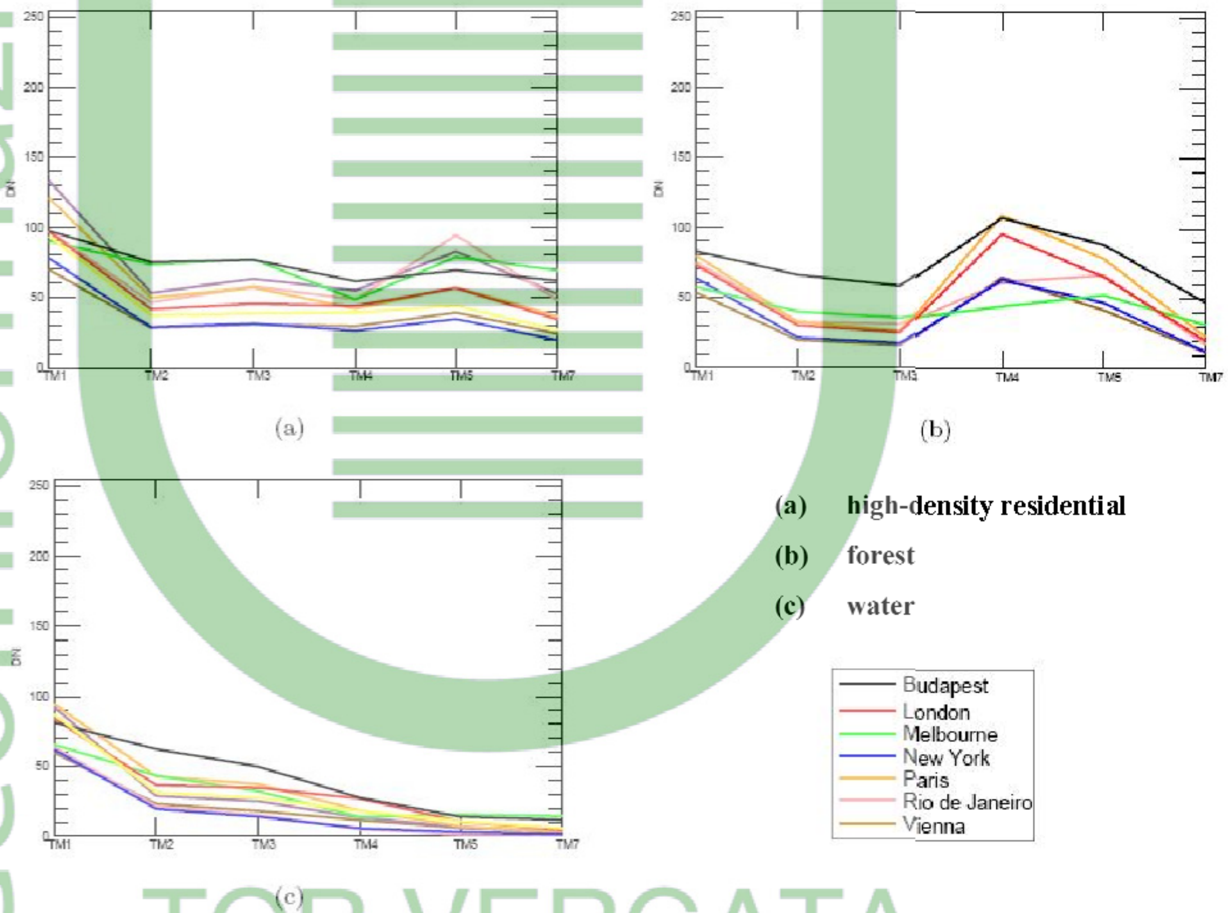
The Landsat dataset consists of a collection of images containing urban areas and located throughout four continents

More than 56000 patterns have been selected to train the final NN dedicated to the Landsat imagery with examples extracted from an overall set of 14 images including urban areas of 12 world big cities, 12 countries.

City	Sensor	Acquisition date
Amsterdam	TM	May 22, 1992
Barcelona	ETM+	Aug 10, 2000
Berlin	ETM+	Aug 14, 2000
Budapest	ETM+	Aug 09, 1999
London	TM	May 20, 1992
Melbourne	ETM+	Oct 05, 2000
New York	TM	Sep 28, 1989
Paris	TM	May 09, 1987
Rio de Janeiro	TM	Jan 18, 1988
Rome	ETM+	Aug 03, 2001
Rome 2	ETM+	Jan 16, 2001
Tokyo	TM	May 21, 1987
Udine	ETM+	Aug 16, 2000
Vienna	TM	Sep 10, 1991

We first analyzed the spectral signatures of the main classes of urban land cover.

Despite the considerable distances among the geographic locations, a good stability of the spectral information has been noted.



(a) high-density residential

(b) forest

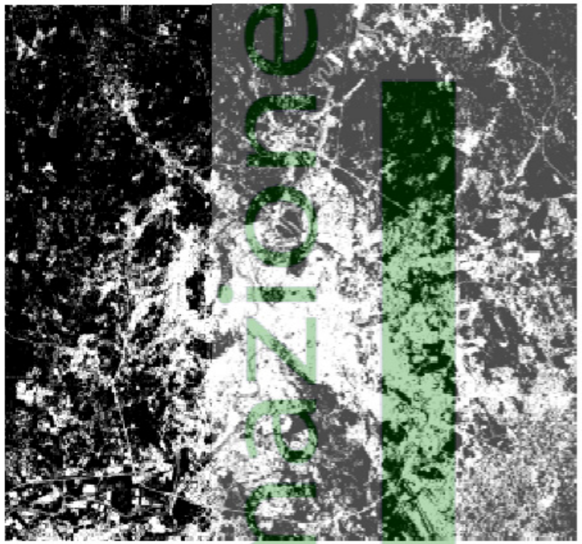
(c) water

Considered classes

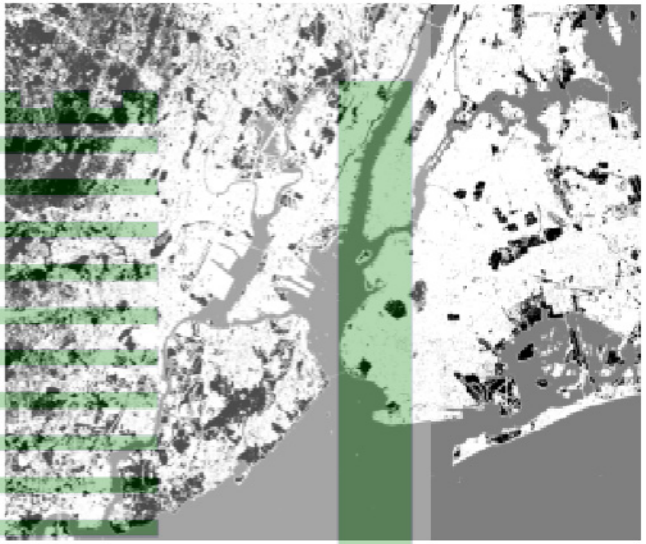
For the classification of each image 3 classes have been considered:

1. **sealed**: includes surfaces with dominant human influence such as continuous and discontinuous urban fabric, industrial, commercial and transport areas including all asphalted surfaces
2. **unsealed**: includes arable lands, permanent crops, pasture, forest and semi-natural areas
3. **water**: includes all water surfaces

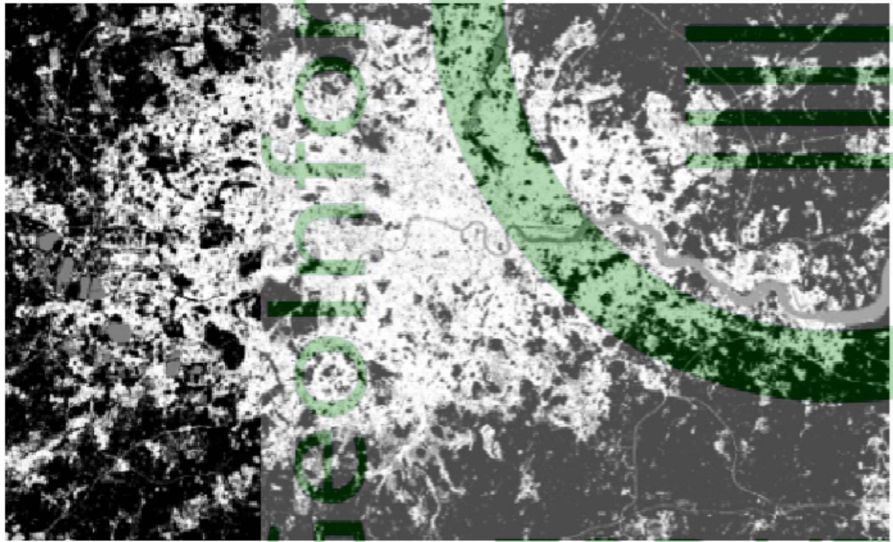
Training images



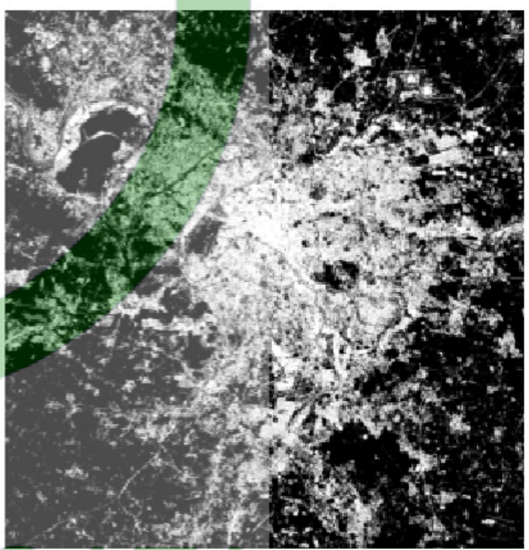
Rome (Italy)



New York (USA)

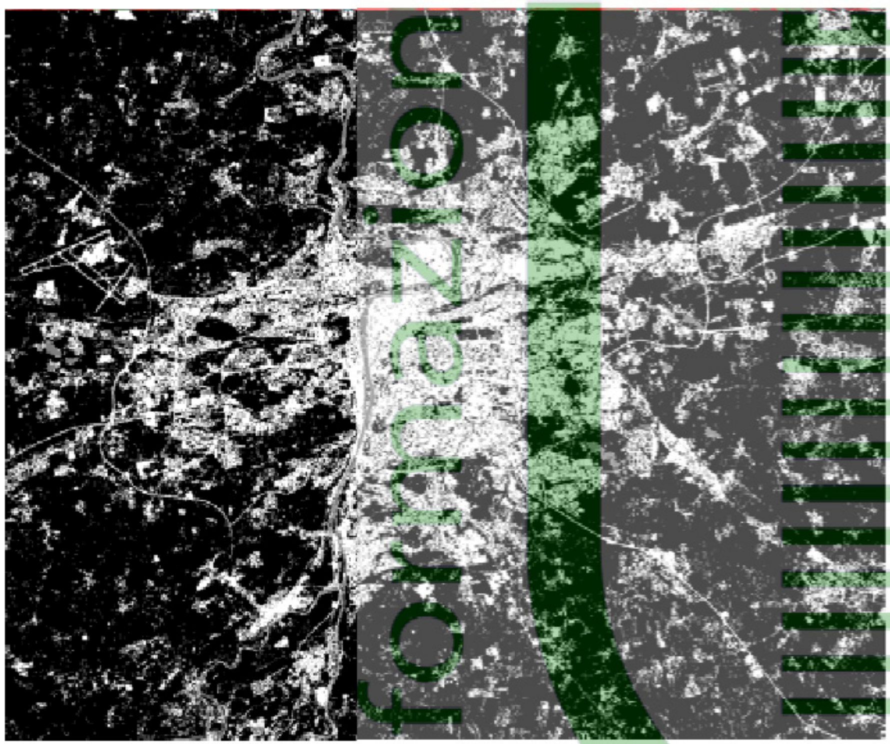


London (UK)

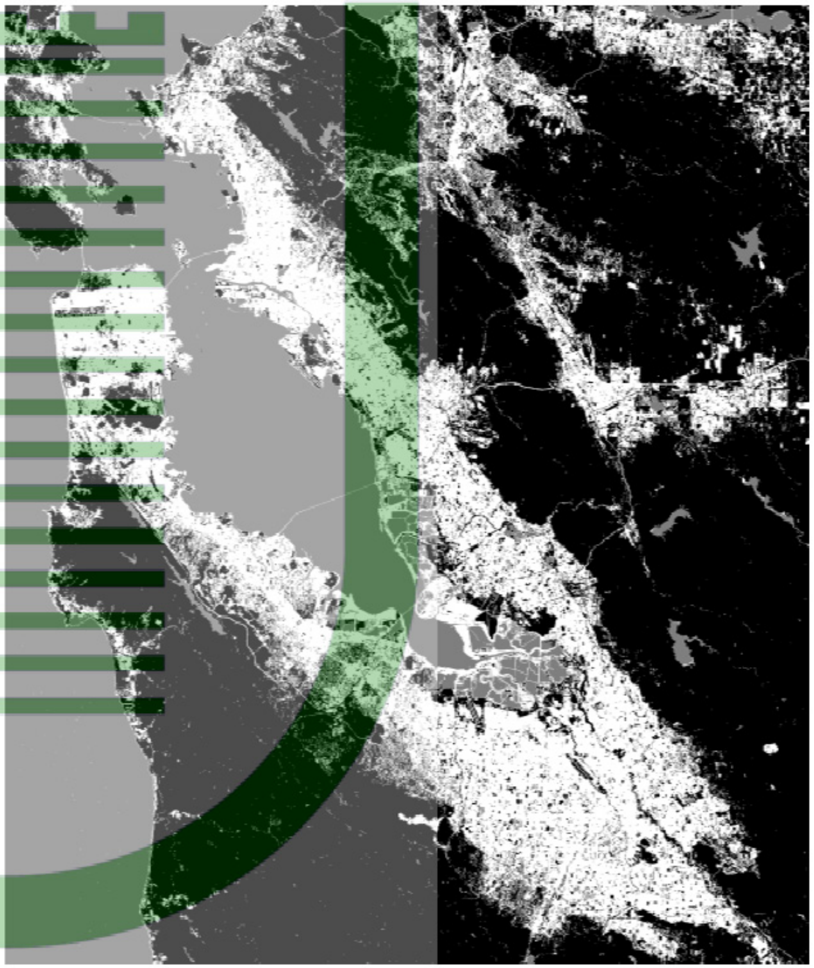


Paris (France)

Validation images



Prague (Czech Republic)



San Francisco (USA)

	Overall Acc. (%)
Training Set	90.9
Validation Set	87.7

Image Information Mining (IIM) – VHR images

A detailed analysis of the pixel-based classification yielded by Neural Networks on very high resolution images such as those provided by the Quickbird or Ikonos platforms was still lacking in the literature.

→ Here, we mean that the NNs might be used to retrieve from the archive all the images that contain or do not contain a specific class of land cover, or where the ratio between areas corresponding to different classes is within/out predefined ranges.

	Aquisition Date	Dimension (pixels)	Off Nadir Angle	Location
QB1	03/13/2003	2415x1650	8 Degrees	Rome, SE outskirts
QB2	05/29/2002	2352x1491	11 Degrees	Rome, SE outskirts
QB3	07/19/2004	2415x1650	23 Degrees	Rome, NE outskirts
QB4	07/19/2004	2415x1450	23 Degrees	Rome Downtown
QB5	07/22/2005	2223x1450	12 Degrees	Nettuno

Spectral signatures

The considered land cover classes were

1. buildings
2. roads
3. vegetated areas
4. bare soil (including not asphalted road and artificial excavations)

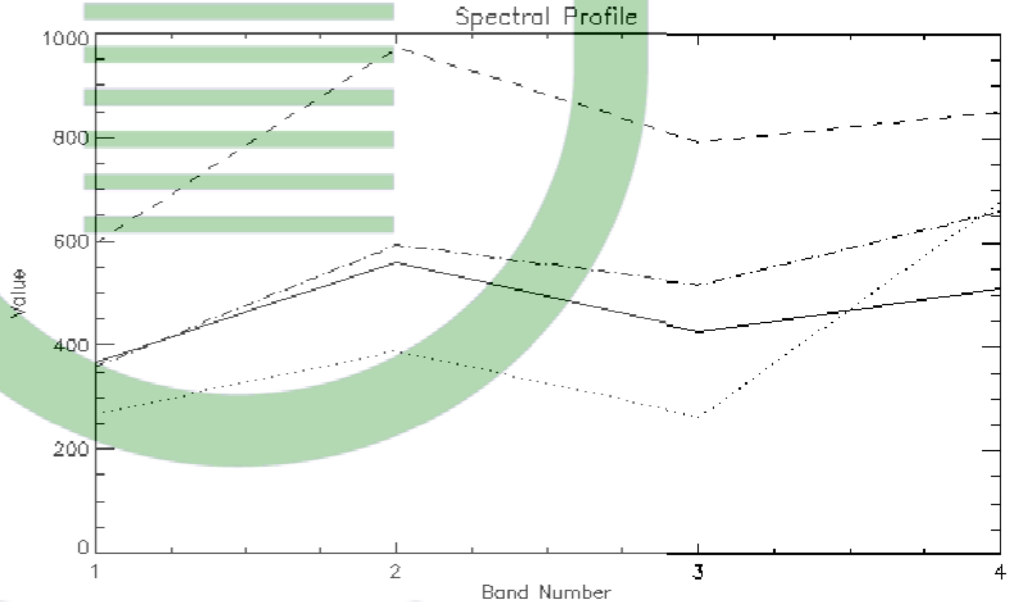
Spectral analysis from image QB1 for the classes:

•buildings (dashed line),

•asphalted surface (solid line),

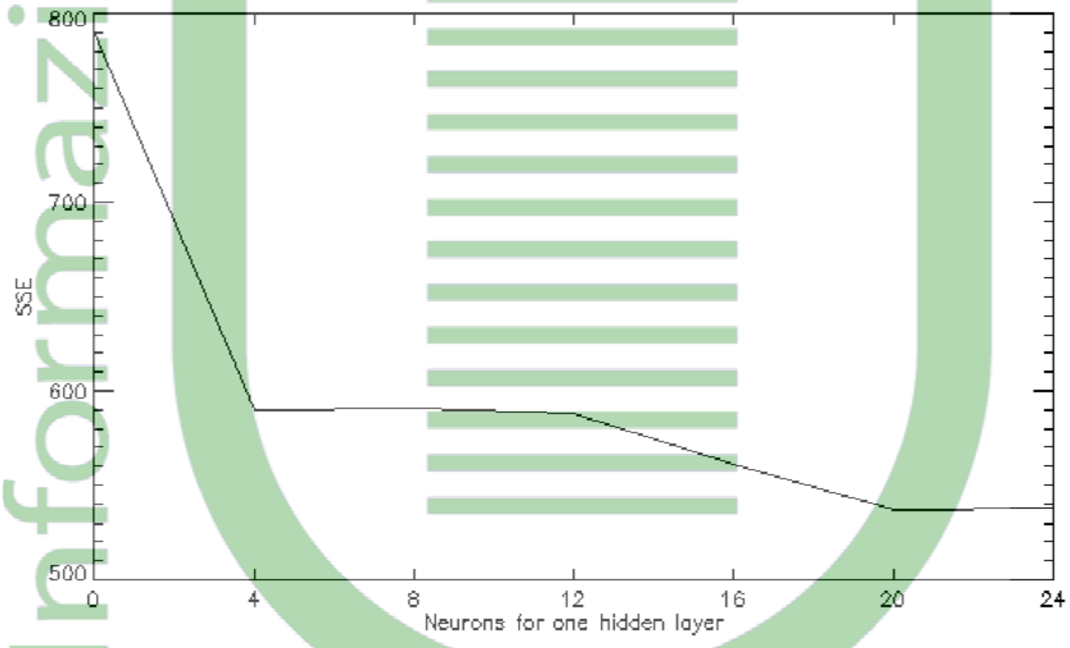
•bare soil (dash-dotted line),

•vegetation (dotted line).



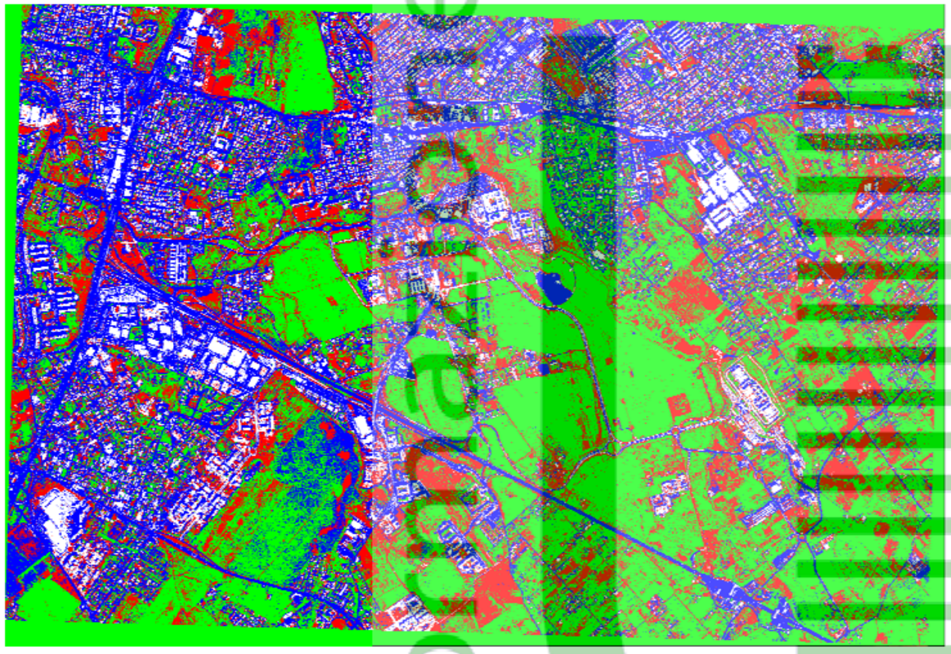
NN optimization

As a first step, we want to assess and optimize the neural network approach for the pixel-based classification of a single very high resolution image.



It can be seen that, considering both the SSE error and the network complexity, the best results were obtained with a 4-20-20-4 topology.

Training image



The QuickBird image taken over the Tor Vergata University campus (South-East of Rome, Italy), on March 13, 2003.

Besides the buildings in the campus, different industrial, commercial and residential areas can be distinguished in the image.

red: bare soil

blue: asphalted surface

white: buildings

green: vegetation

Classified as	True			
	Vegetation	*Asphalt*	*Building*	*Bare soil*
Vegetation	14864	33	750	2207
Asphalt	132	44785	68	27
Building	1225	29	12634	783
Bare soil	230	2	512	3229
Accuracy = 93%				

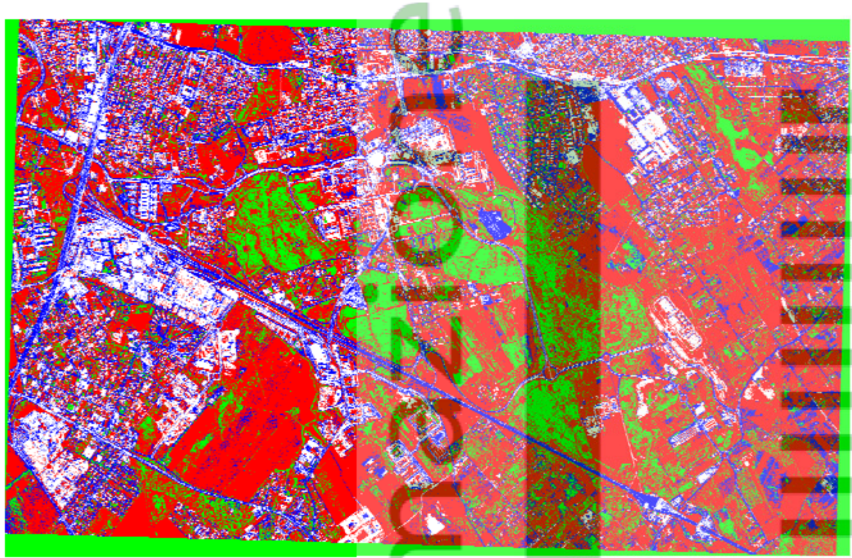
Unique network (1/2)

Once the network topology for this kind of problem has been optimized and the performance assessed, we move to investigate on the capability of a unique network to provide classification on different images rather than on a single one.

To underline the complexity of this new problem we tested the already designed network, positively processing the QB1 image, on another QB image.

The new QuickBird image (QB2) is taken on the same area of the first one, but in a different season and at a slightly different incident angle.

Unique network (2/2)

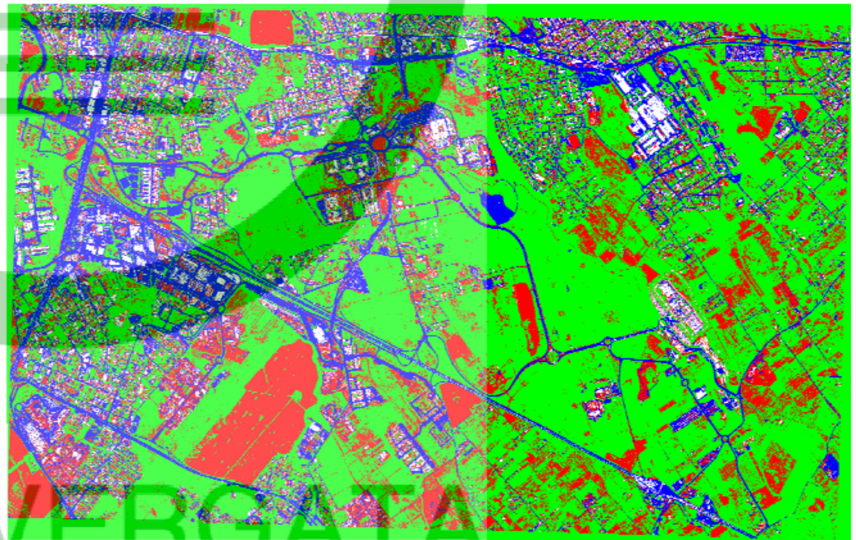


Automatic classification map from the image QB2 with a net trained with examples taken from image QB1 (accuracy=56%)

QB2 using a network trained with examples taken from image QB2 (accuracy=95%)

QB2 with a net trained with examples taken from QB1, QB3, QB4, QB5 (accuracy=87%)

Red: bare soil, blue: asphalted surface, white: buildings, green: vegetation.



Conclusions

This work can be considered as a first step in demonstrating how NNs can contribute at the development of IIM in Earth Observation.

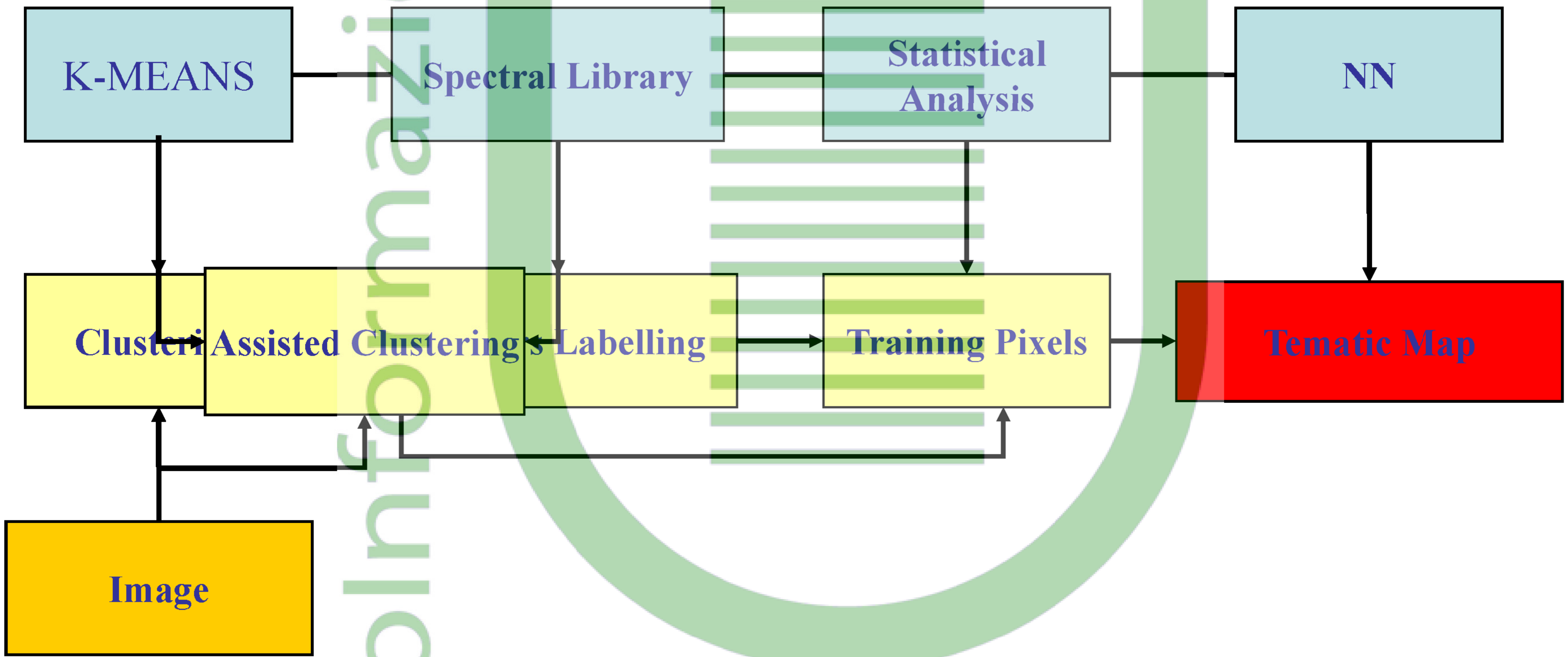
In both high and very high resolution cases, the purpose was to train a NN in order to generalize the image data set considered in the training phase so that the new images in large archives could be processed in real time.

The network performance seems to be satisfactory, especially if we take into account that the procedures are completely automatic. In fact, the maps automatically provided on new images, that is not considered in the training phase, show good agreement with those that would be obtained with careful visual inspection or with the available ground-truth.

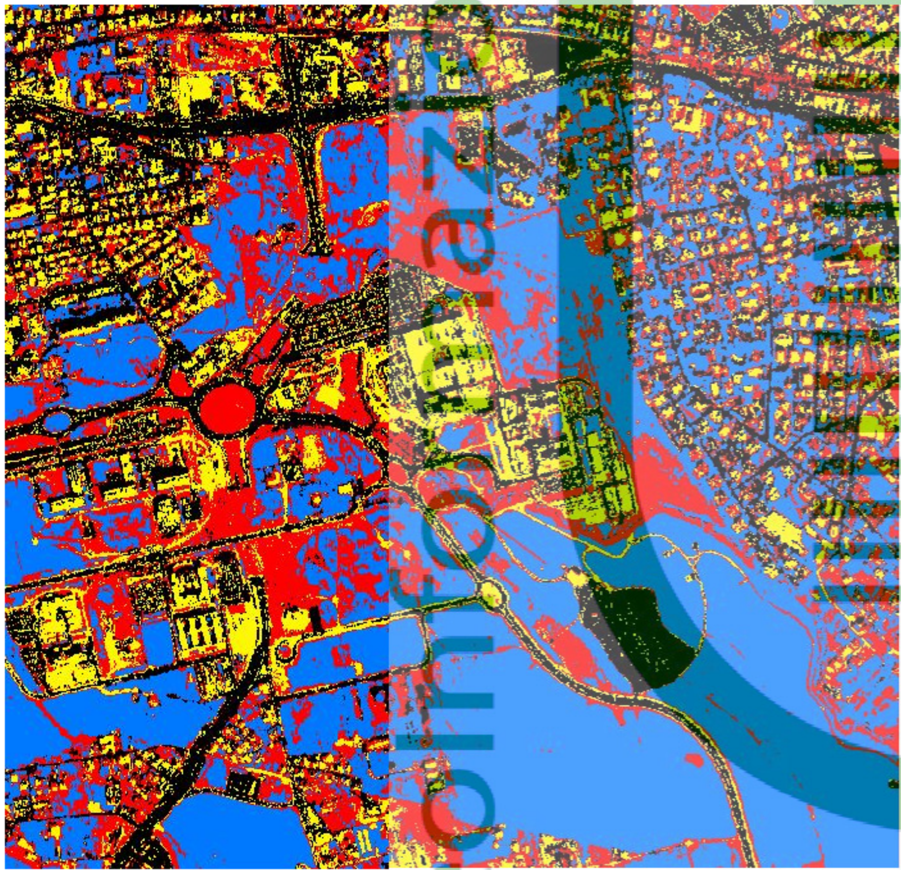
Refereed publications

Journals:

1. F. Del Frate, F. Pacifici, G. Schiavon, C. Solimini, "Use of neural networks for automatic classification from high-resolution images", IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 4, pp. 800-809, April 2007


K-MEANS
Spectral Library
Statistical Analysis
NN
Assisted Clustering
Labelling
Training Pixels
Tematic Map
Image

Supervised NN



yellow: built-up

blue: vegetation

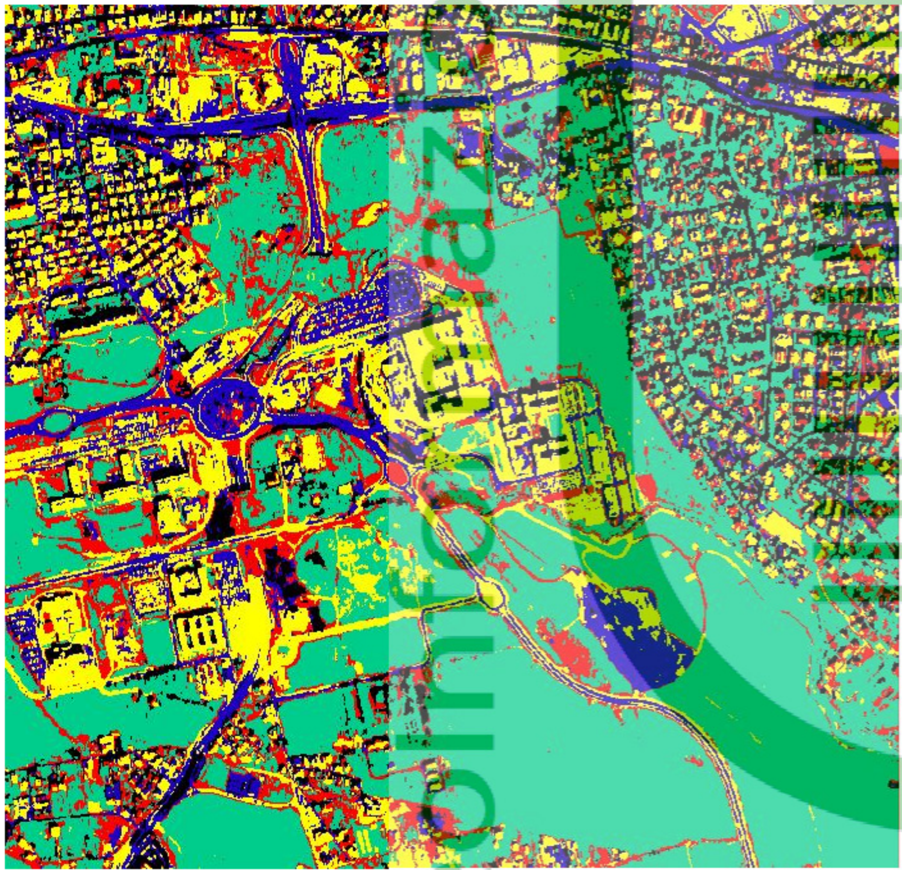
red: bare soil

black: asphalt

	Asphalt	Vegetation	Built-up	Bare Soil
Asphalt	91.60	0.13	15.73	1.60
Vegetation	0.00	98.93	0.00	0.00
Built-up	6.00	0.00	78.40	2.80
Bare Soil	2.40	0.93	5.87	95.60

Overall Accuracy = 89.90%

Full Automatic Scheme



yellow: built-up

blue: asphalt

red: bare soil

green: vegetation

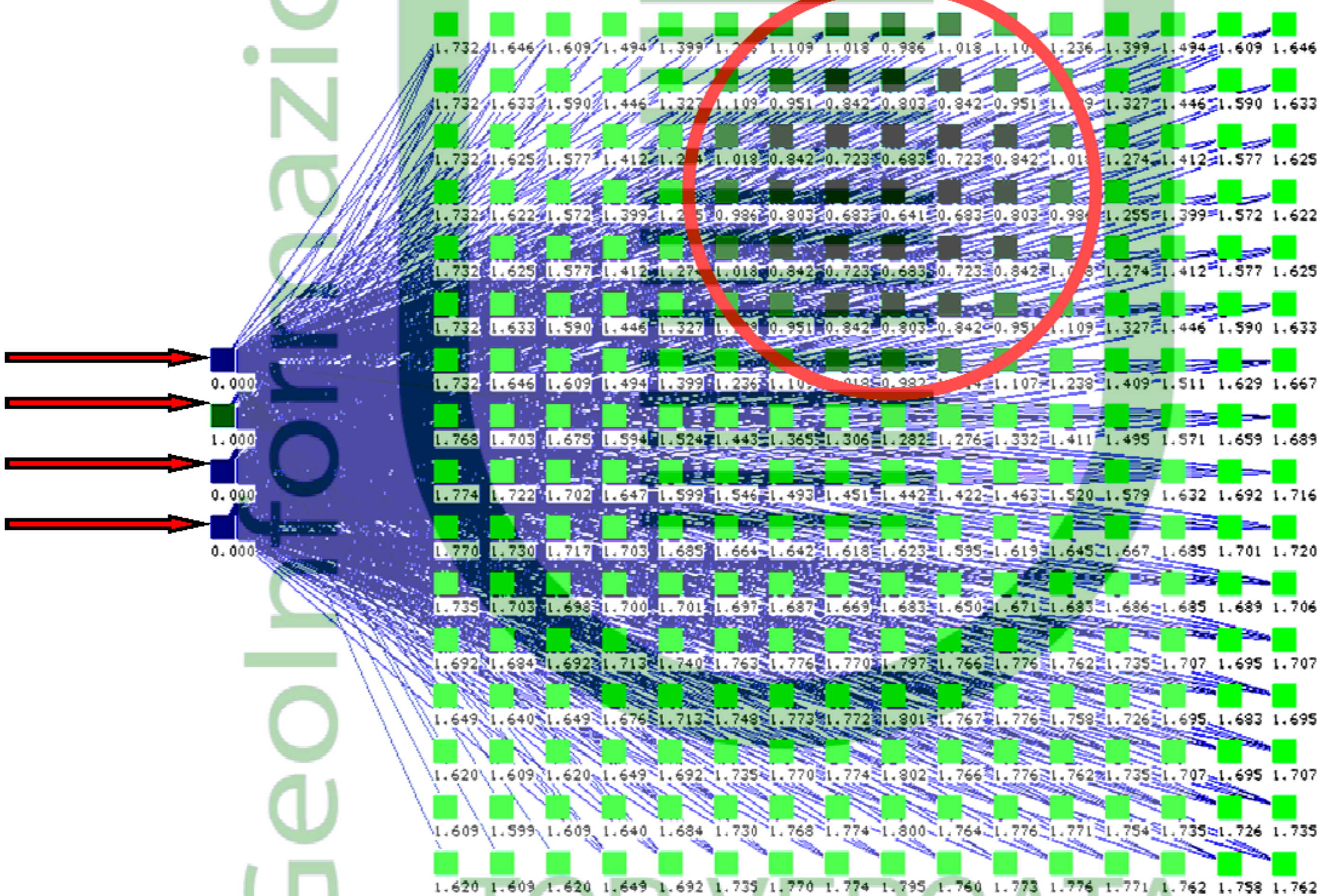
	Asphalt	Vegetation	Built-up	Bare Soil
Asphalt	78.66	11.11	0.00	31.08
Vegetation	18.41	81.76	0.40	14.41
Built-up	0.00	0.00	99.60	2.25
Bare Soil	2.93	7.13	0.00	52.25

Overall Accuracy = 84.89%

vs

Overall Accuracy = 89.90%

The weights are organized such that topologically close nodes are sensitive to inputs that are physically similar.



1 Initialize weights:
Initialize weights from *N* inputs to the *M* outputs nodes to small random values. Set the initial radius of the neighborhood

2 Present New Input

3 Compute Distance to All Nodes:
Compute distances d_j between the input and each output node *j* using

$$d_j = \sum_{i=0}^{N-1} \left(x_i(t) - w_{ij}(t) \right)^2$$

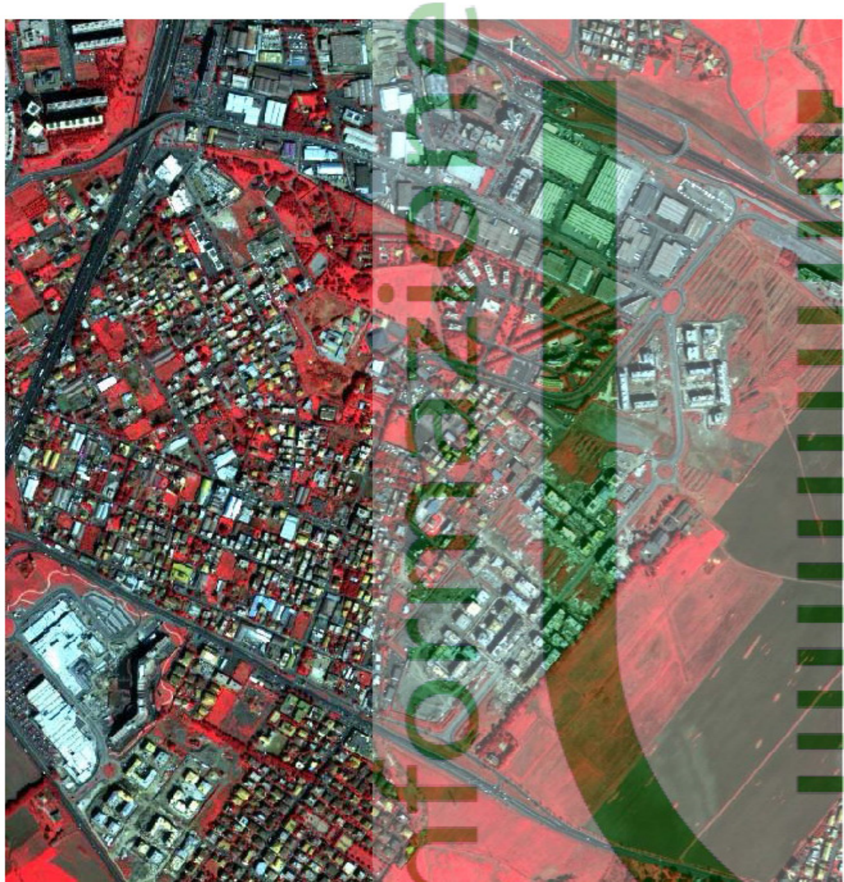
4 Select Output Node with Minimum Distance:
Select node *j** as that output node with minimum d_j

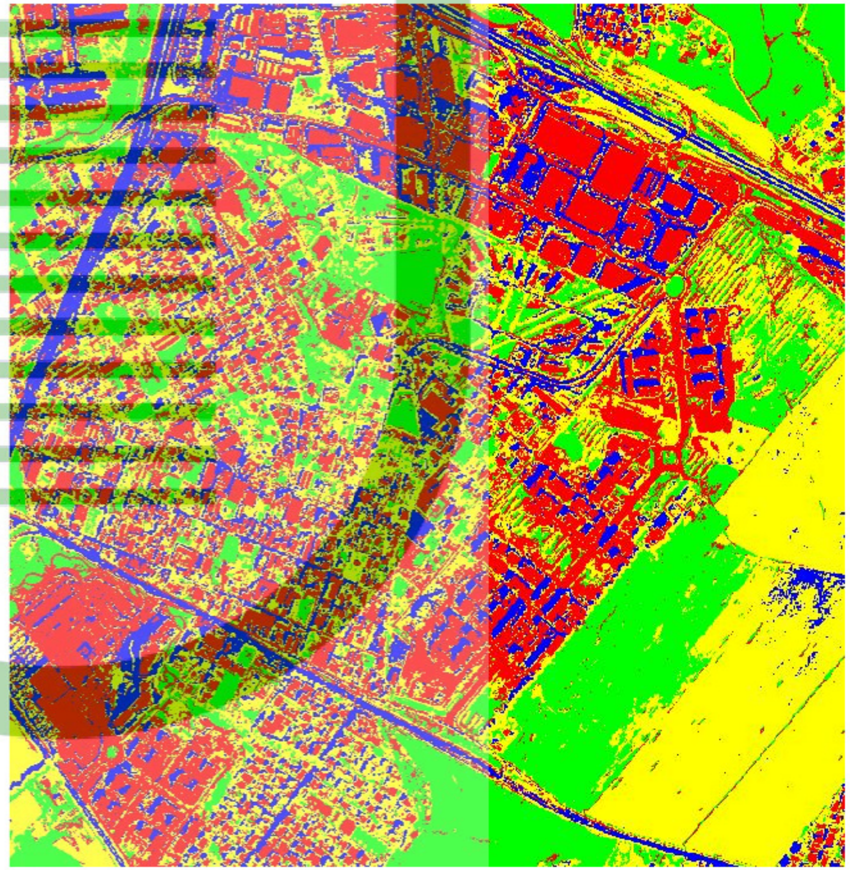
5 Update Weights to Node *j and Neighbours:**
Weights are updated for node *j** and all nodes in the neighborhood defined by $NE_{j*}(t)$. New weights are:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \left(x_i(t) - w_{ij}(t) \right)$$

6 Repeat by Going to Step 2

…on going research: Kohonen Self Organizing Maps (3/3)





yellow: bare soil

blue: asphalted surface

white: buildings

green: vegetation

Automatic classification of urban areas with high resolution SAR imagery (ERS/ENVISAT)

GeoInformazione

TOR VERGATA

Goals

The global view of urban areas makes satellite missions a valid instrument for updating urban maps and carrying out the analysis of settlement dynamics.

Remote sensing in the optical band is a well established tool for producing maps of urban land cover and monitoring changes, but it can suffer from atmospheric limitations.

The management of emergencies over LARGE AREAS relies on near-real time information, irrespective of the time of day and of the cloud cover: to this purpose the availability of SAR acquisitions is essential.

Goals

The new generation satellite missions, such as the Canadian Radarsat 2, the German TerraSAR-X and the Italian COSMO-SkyMed will make available **LARGE ARCHIVES** of images.

In the following, we discuss the identification and use of single-polarization SAR image features for land cover applications.

↓

Simulation of emergency acquisitions in a minimal configuration/processing.

Three configurations are considered:

- *short-term* classification scheme intended for providing information in near-real time
- *long-term* scheme aimed at observing the urban changes at year time scales
- *fully automatic* classification scheme

Data Set

Site Information		Images Information			
Location	Dimension (km^2)	Acquisition Date	Satellite	B_p (m)	Dimension (pixels)
Rome, Italy	836	January 25, 1994	ERS 1	89	1245 x 1300
		January 31, 1994	ERS 1		
		March 26, 1994	ERS 1	157	
		March 29, 1994	ERS 1		
		July 13, 1994	ERS 1	-	
		February 13, 1999	ERS 1	211	
		February 14, 1999	ERS 2		
		March 20, 1999	ERS 1	65	
		March 21, 1999	ERS 2		
		July 4, 1999	ERS 2	-	

Input Features

Careful selection and suitable processing are required to exploit the various pieces of information embedded in the amplitude, its time-space behavior and phase of the radar return:

backscattering → mean intensity / intensity st. deviation

interferometric coherence → late winter / early summer coherence

textural features → Gray-Level Co-occurence Matrix (GLCM)

↓

Contrast – Energy (literature review)

↓

Window size : 7x7, 11x11 and 15x15

Input Features

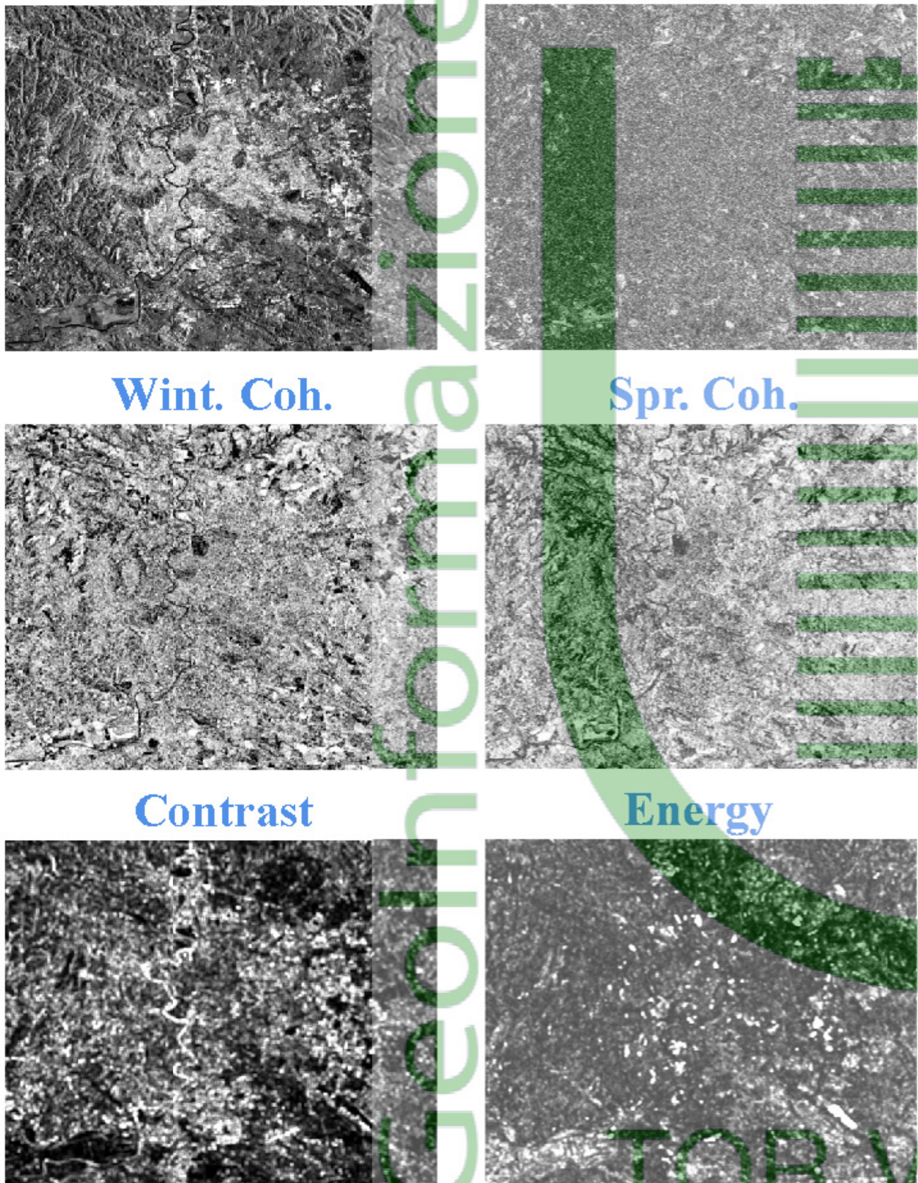
For *long-term* monitoring, the time-average amplitude of the backscattering coefficient, the degree of interferometric coherence corresponding to the late winter and early summer, and the seasonal variations are exploited together with two textural (*Contrast* and *Energy*) parameters of the radar amplitude image.

6 inputs

Short-term monitoring is based on the same data set, abridged by excluding the seasonal variations of the backscattering and one of the coherence values.

4 inputs

Input Features



Long-term	
1	Mean Int.
2	Int. St. Dev.
3	Wint. Coh.
4	Spr. Coh.
5	Contrast
6	Energy

5 images

Short-term	
1	Mean Int.
2	Spr. Coh.
3	Contrast
4	Energy

2 images

Class Definition

These sets of 6 and 4 parameters respectively are exploited to discriminate among seven urban/sub-urban classes, including water surfaces (WS), vegetation (VE), forest (FO), asphalted surfaces (AS), isolated large buildings (IB) and continuous high/low density residential areas (HD/LD).

Classes	TR	VS
Asphalt (AS)	511	219
Forest (FO)	2592	1326
High Density (HD)	648	278
Isolated Buildings (IB)	122	433
Low Density (LD)	4900	6130
Vegetation (VE)	3535	6008
Water (WS)	892	382
Total	**13200**	**14776**

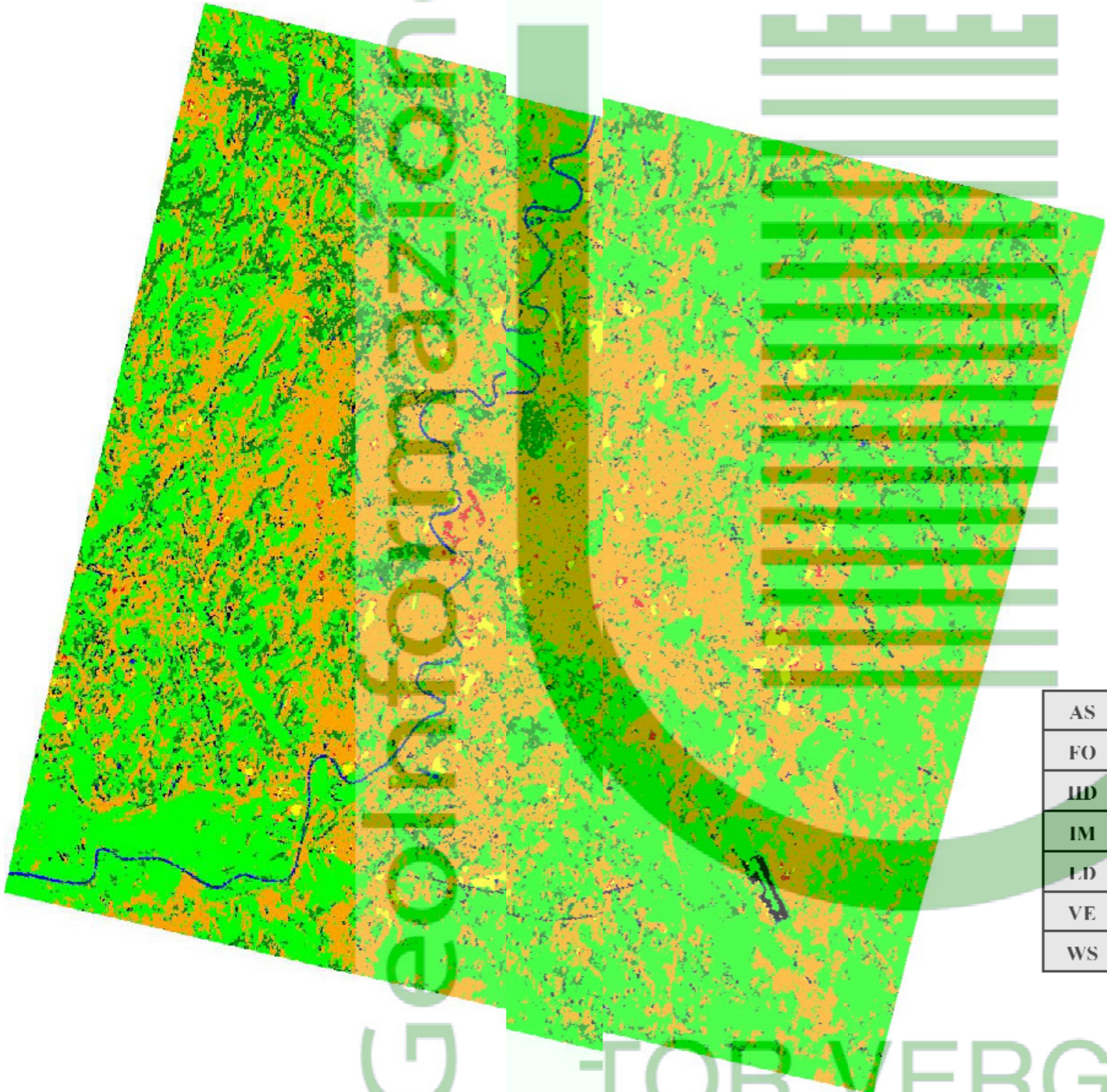
Neural Network Approach

The pieces of information extracted from the SAR images are fused and processed by a supervised Multi-Layer Perceptron (MLP) neural network which is known to show a considerable ease in using multi-domain data sources.

We recorded the classification accuracies yielded by a varying number of hidden neurons, starting from a small topology (6-12-12-7) to end with a large one (6-100-100-7). The variance of the accuracy for different initializations of the weights was computed to monitor the stability of the algorithm.

The Magnitude Based Pruning procedure has then been applied to thin the net.

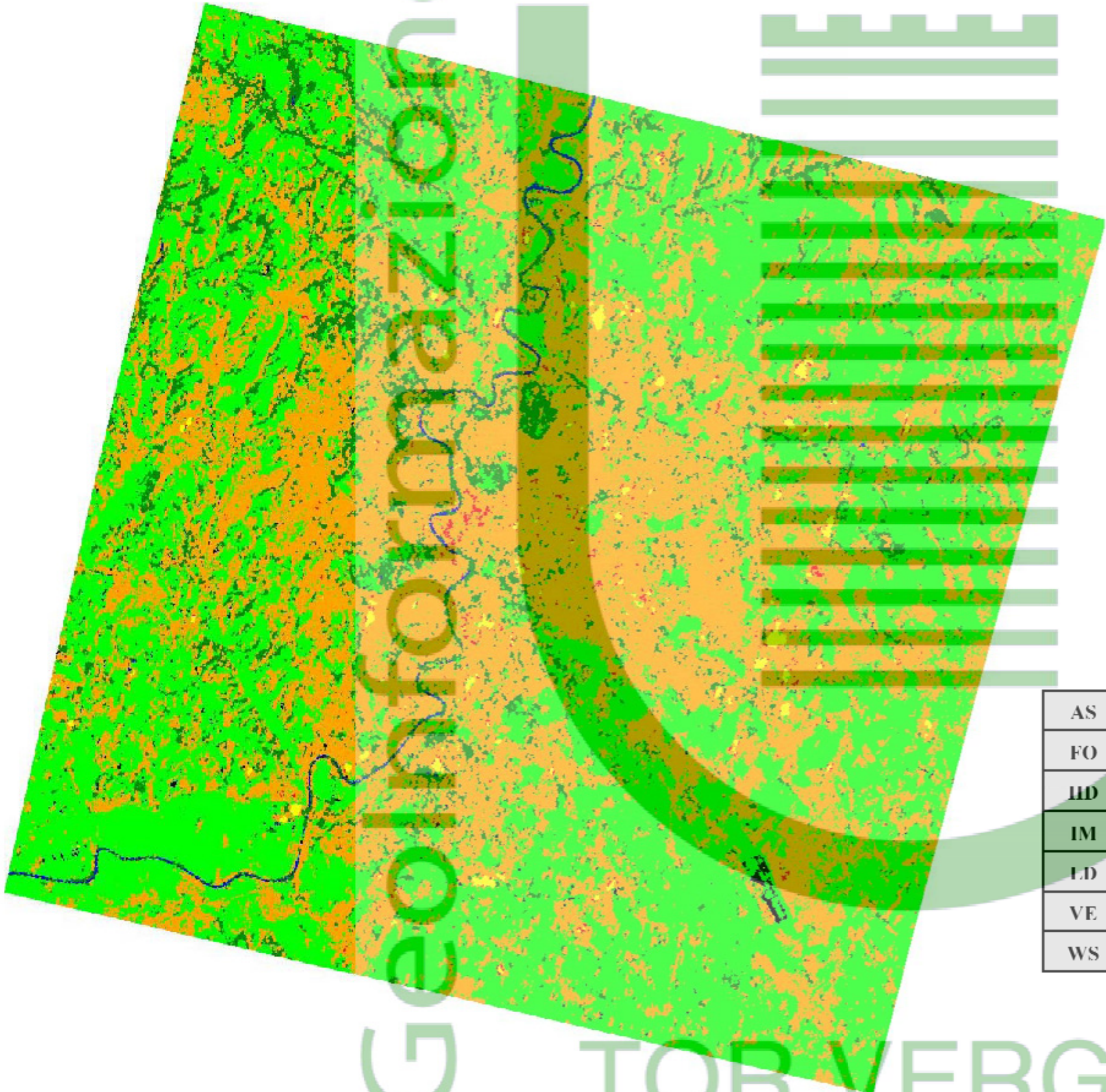
Short-term classification - 4 inputs 1994



Classes	
Asphalt (AS)	
Forest (FO)	
High Density (HD)	
Isolated Buildings (IB)	
Low Density (LD)	
Vegetation (VE)	
Water (WS)	

	AS	FO	HD	IM	LD	VE	WS
AS	74.18	2.75	0.00	0.00	0.00	19.23	3.85
FO	0.00	81.39	0.00	0.00	16.71	10.05	1.49
HD	0.00	0.00	36.32	1.42	61.32	0.94	0.00
IM	0.00	0.00	1.12	96.21	2.68	0.00	0.00
LD	0.07	0.95	0.23	0.11	96.99	1.65	0.00
VE	0.39	1.08	0.00	0.00	7.66	93.50	0.01
WS	1.22	1.52	0.00	0.00	0.30	0.30	93.92
	Overall Er. (%)		8.12		k-Coefficient		0.863

Short-term classification - 4 inputs 1999



Classes
Asphalt (AS)
Forest (FO)
High Density (HD)
Isolated Buildings (IB)
Low Density (LD)
Vegetation (VE)
Water (WS)

	AS	FO	HD	IM	LD	VE	WS
AS	52.97	1.83	0.00	0.00	1.37	42.01	1.83
FO	0.15	79.34	0.00	0.00	9.05	8.82	2.64
HD	0.00	0.00	23.02	1.80	73.74	1.44	0.00
IM	0.00	0.00	0.23	94.00	5.54	0.23	0.00
LD	0.03	0.78	0.52	0.10	96.15	2.40	0.02
VE	0.80	2.18	0.08	0.02	7.86	88.96	0.10
WS	4.71	9.69	0.00	0.00	1.05	1.31	83.25
	Overall Er. (%)		10.69		k-Coefficient		0.834

Short-term classification

The classification accuracy is slightly less than 0.84.

OSS: due to the decametric size of the resolution cells at ground, mixed pixels are likely to occur, especially in a sub-urban landscape, where heterogeneous land covers coexist within short distances.

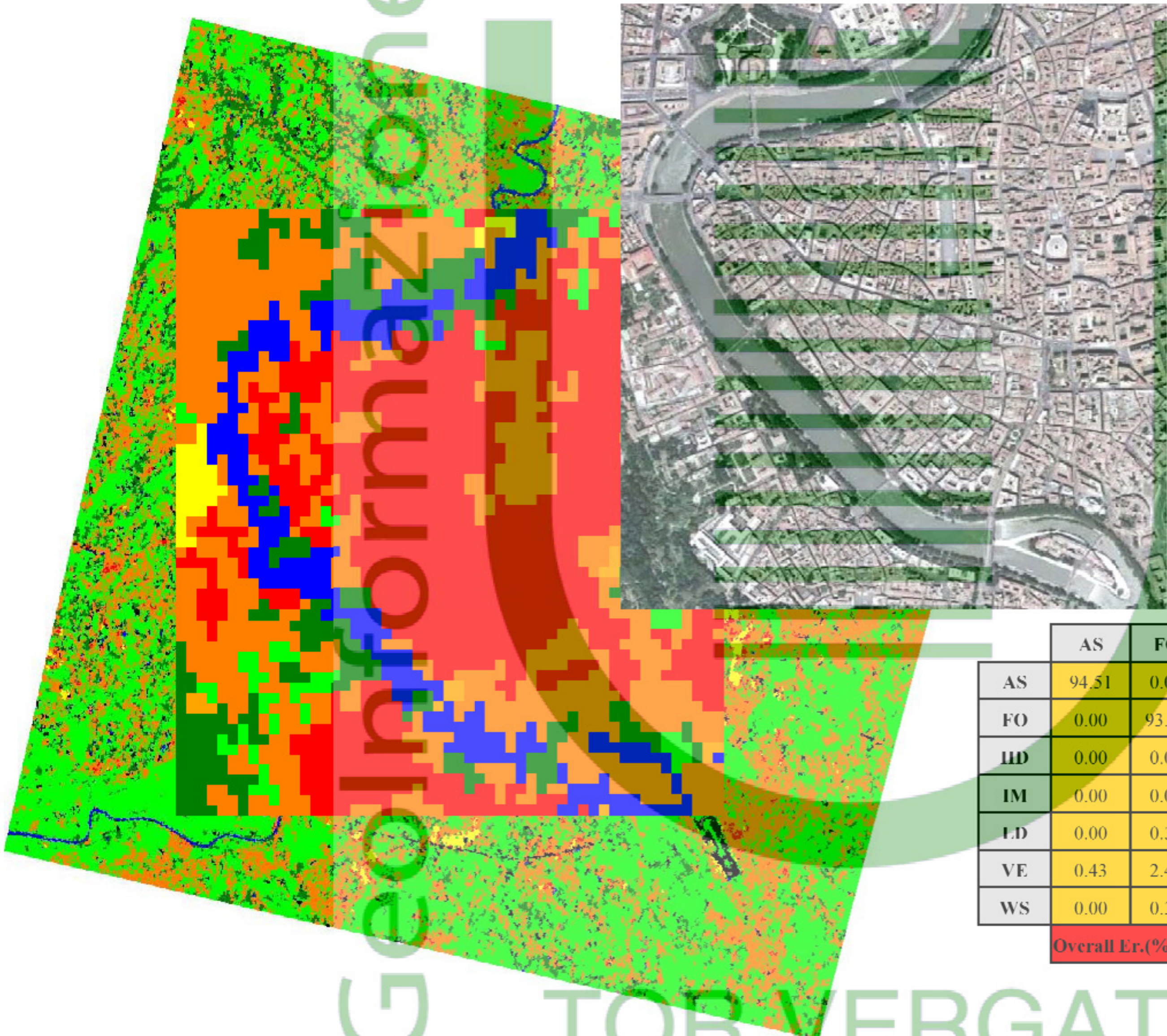
As expected from physical considerations, the classification errors mainly consist in misclassification between:

- high- and low-density residential areas
- asphalt and low vegetation
- low vegetation and low-density residential areas

↓

Adding the seasonal information is expected to improve the results.

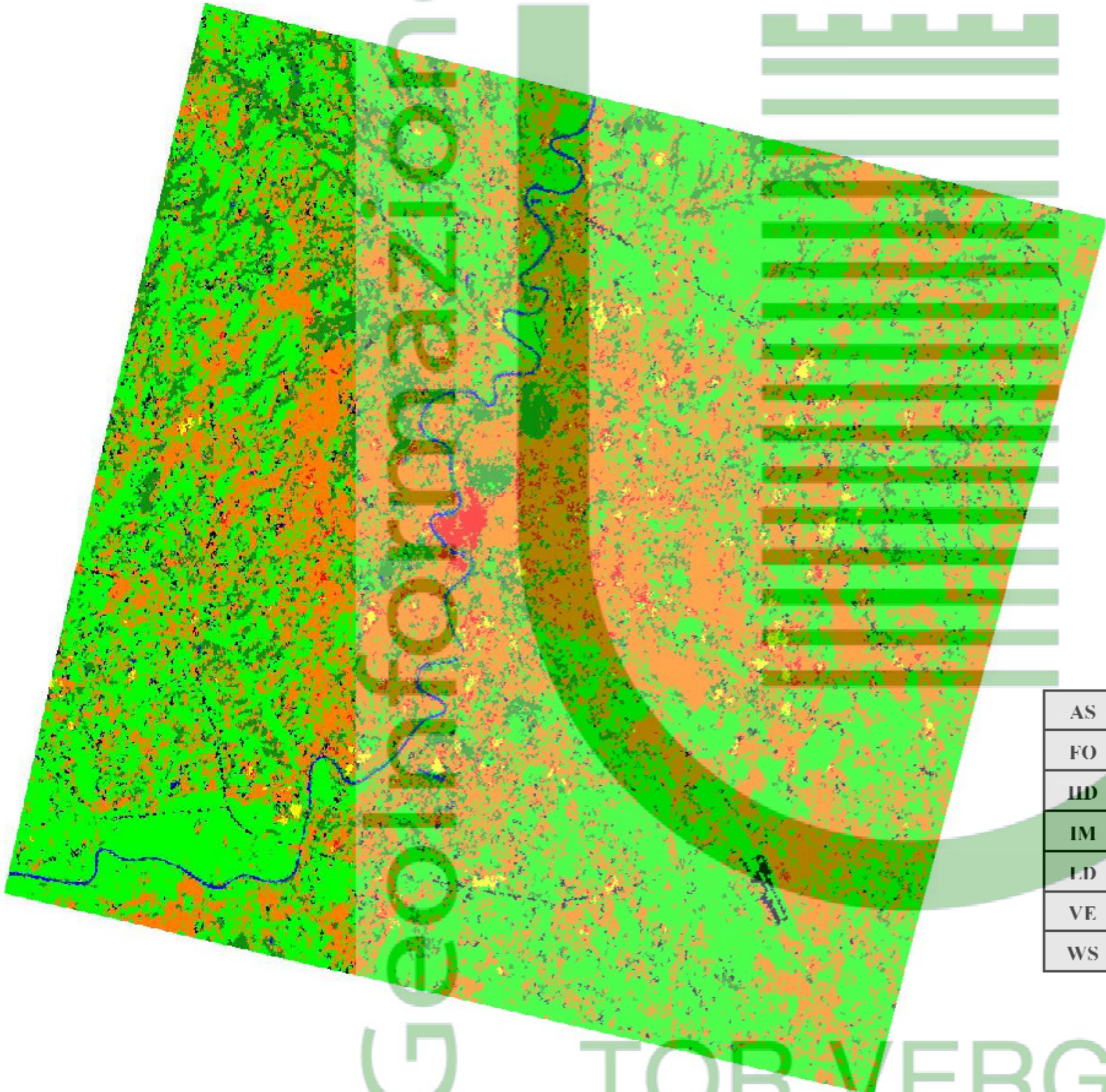
Long-term classification - 6 inputs 1994



Classes	
Asphalt (AS)	
Forest (FO)	
High Density (HD)	
Isolated Buildings (IB)	
Low Density (LD)	
Vegetation (VE)	
Water (WS)	

	AS	FO	HD	IM	LD	VE	WS
AS	94.51	0.00	0.00	0.55	0.55	4.40	0.00
FO	0.00	93.07	0.00	0.00	2.85	3.80	0.27
HD	0.00	0.00	87.74	0.00	11.32	0.94	0.00
IM	0.00	0.00	3.35	93.75	2.68	0.22	0.00
LD	0.00	0.34	1.67	0.08	95.39	2.52	0.00
VE	0.43	2.47	0.00	0.15	1.80	95.16	0.00
WS	0.00	0.30	0.00	0.00	0.00	0.00	96.96
	Overall Er.(%)		4.91		k-Coefficient		0.923

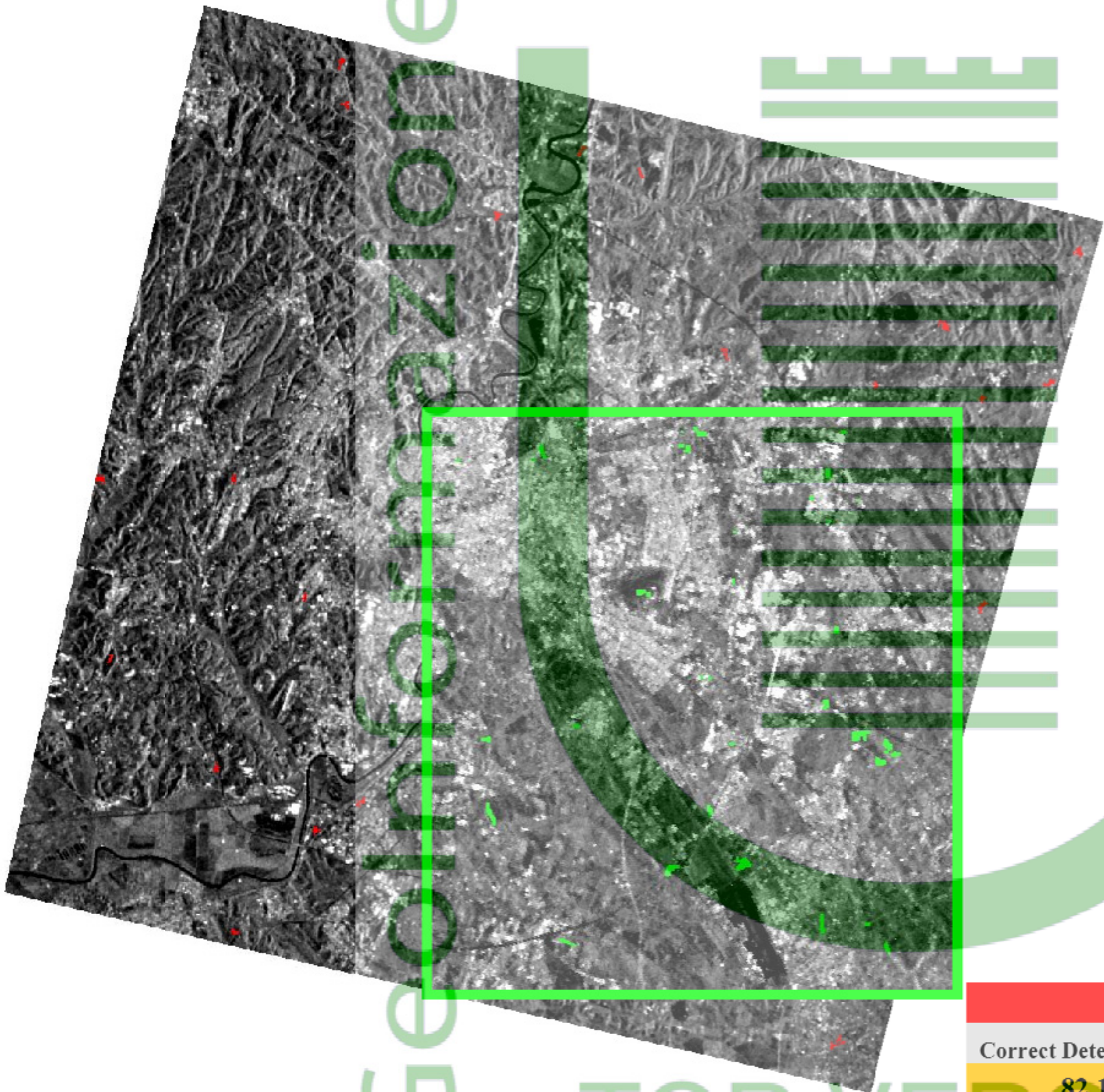
Long-term classification - 6 inputs 1999

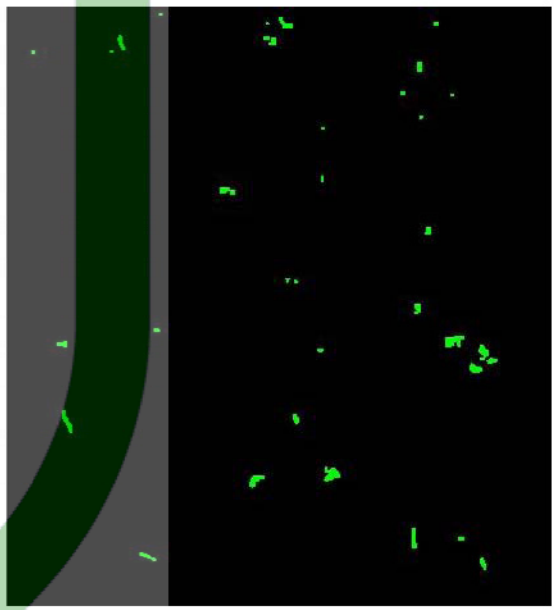


Classes	
Asphalt (AS)	
Forest (FO)	
High Density (HD)	
Isolated Buildings (IB)	
Low Density (LD)	
Vegetation (VE)	
Water (WS)	

	AS	FO	HD	IM	LD	VE	WS
AS	87.21	0.91	0.00	0.00	1.37	10.05	0.46
FO	0.60	93.21	0.00	0.00	3.17	2.71	0.30
HD	0.00	0.72	90.29	0.00	8.99	0.00	0.00
IM	0.00	0.00	1.85	85.68	3.46	9.01	0.00
LD	0.03	0.57	2.07	0.03	95.73	1.52	0.05
VE	0.15	3.01	0.02	0.08	3.30	93.43	0.02
WS	0.79	3.66	0.00	0.00	0.26	1.05	94.24
	Overall Er.(%)		6.00		k-Coefficient		0.909

Change Detection 1994-1999





Rome SE: 160km^2 - 423x465pixels		
Correct Detection (%)	False Alarms (%)	Missed Alarms (%)
82.17	17.83	0.26

The fully automatic mode

The results reported in previous sections have shown how SAR imagery and neural networks may be effective in producing classification and change detection maps.

However each image is processed by its own network (trained off-line) which takes into account different conditions (e.g. baseline, soil moisture, …).

This might not meet the need of a fully automatic scheme for a fast processing chain.

GOAL: design a neural network capable of classifying images whose pixels have not considered at all during the training phase.

The fully automatic mode

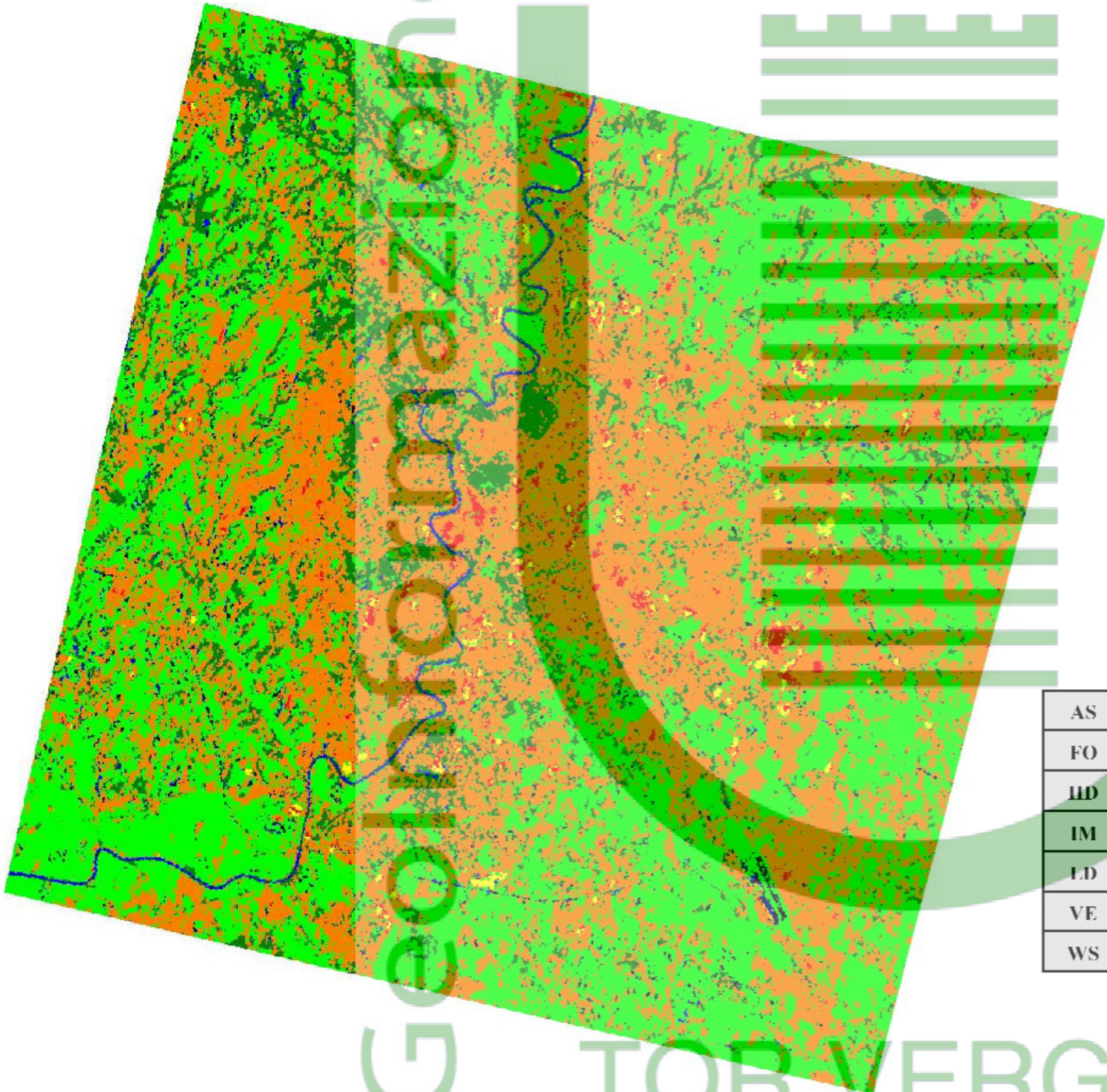
For this purpose we used another set of images over the same test site, but corresponding to another year.

Site Information		Images Information			
Location	**Dimension (km^2)**	**Acquisition Date**	**Satellite**	**B_p (m)**	**Dimension (pixels)**
Rome, Italy	836	February 24, 1996	ERS 1	12	1245 x 1300
		February 25, 1996	ERS 2		
		March 30, 1996	ERS 1	106	
		March 31, 1996	ERS 2		
		July 14, 1996	ERS 2	-	

The classification procedure follows the same scheme illustrated before in terms of the physical quantities to be considered as input and classes to be discriminated.

A set consisting of 26,400 pixels has been created stemming **ONLY** from 1994 and 1999 images.

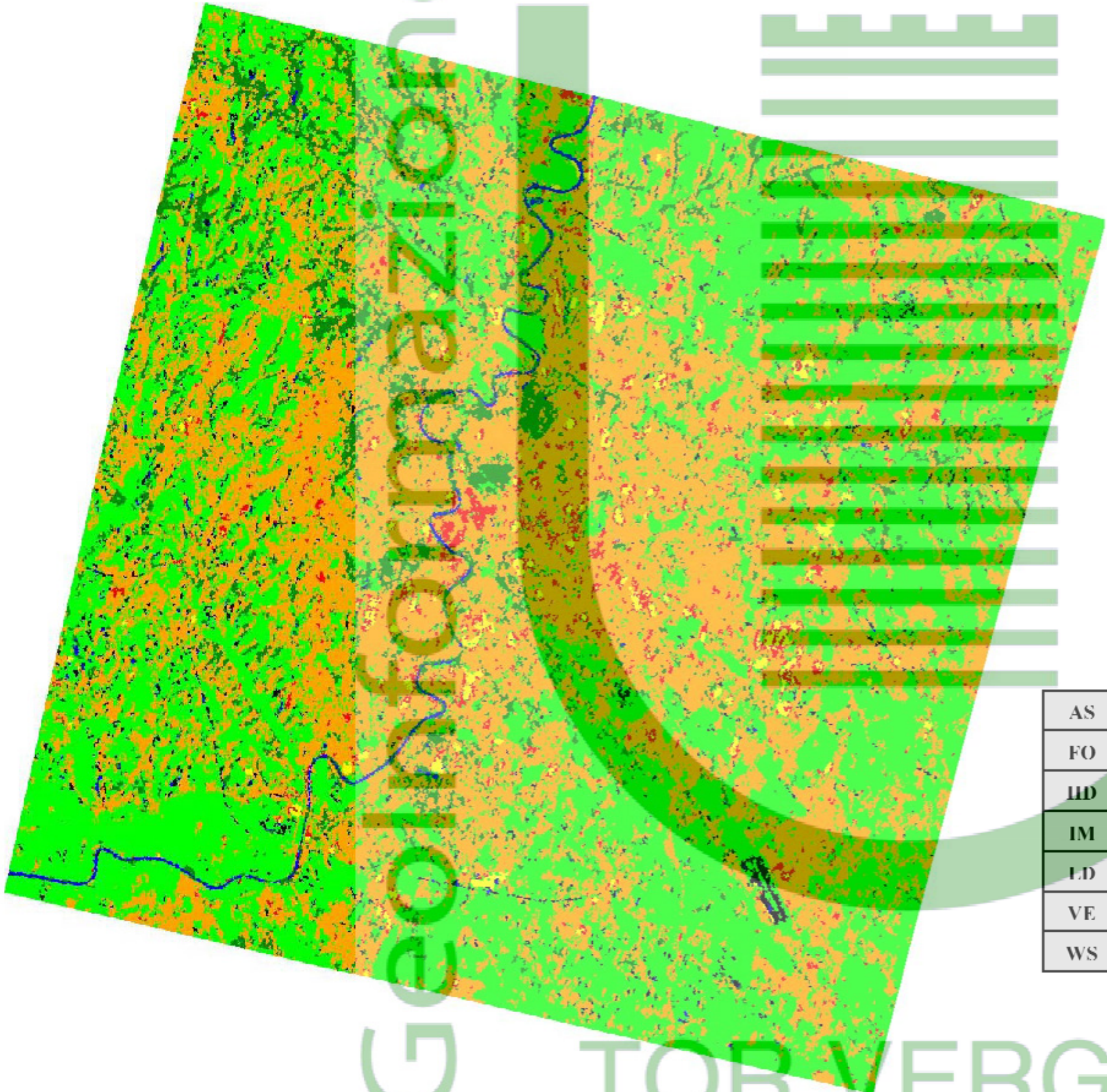
Short-term automatic mode - 4 inputs 1996



Classes	
Asphalt (AS)	black
Forest (FO)	dark green
High Density (HD)	red
Isolated Buildings (IB)	yellow
Low Density (LD)	orange
Vegetation (VE)	green
Water (WS)	blue

	AS	FO	HD	IM	LD	VE	WS
AS	46.15	1.65	0.00	0.00	1.10	43.41	7.69
FO	1.90	69.16	0.00	0.00	15.22	10.60	3.13
HD	0.00	0.47	26.42	0.00	72.64	0.47	0.00
IM	0.00	0.00	4.91	78.13	13.84	3.13	0.00
LD	0.05	2.79	0.39	0.00	91.34	5.39	0.03
VE	1.29	2.55	0.00	0.00	8.08	87.94	0.13
WS	1.22	3.95	0.00	0.00	0.00	3.95	88.15
	Overall Er.(%)		13.52		k-Coefficient		0.780

Long-term automatic mode - 6 inputs 1996



Classes	
Asphalt (AS)	
Forest (FO)	
High Density (HD)	
Isolated Buildings (IB)	
Low Density (LD)	
Vegetation (VE)	
Water (WS)	

	AS	FO	HD	IM	LD	VE	WS
AS	47.80	0.55	0.00	0.55	5.49	42.86	2.75
FO	0.41	71.74	0.00	0.00	15.22	11.14	1.49
HD	0.00	0.00	43.40	0.00	56.13	0.47	0.00
IM	0.00	0.00	16.52	79.46	3.13	0.89	0.00
LD	0.02	0.64	5.92	0.67	89.28	3.48	0.00
VE	1.35	1.44	0.21	0.01	4.71	92.23	0.04
WS	0.91	6.69	0.00	0.00	0.91	2.43	86.32
	Overall Er.(%)		11.73		k-Coefficient		0.811

Long-term automatic mode - 6 inputs 1996

The origin of most of the errors relies in the misclassification of HD as LW, which, given the contiguity of the two classes, can be recognized as a minor drawback.

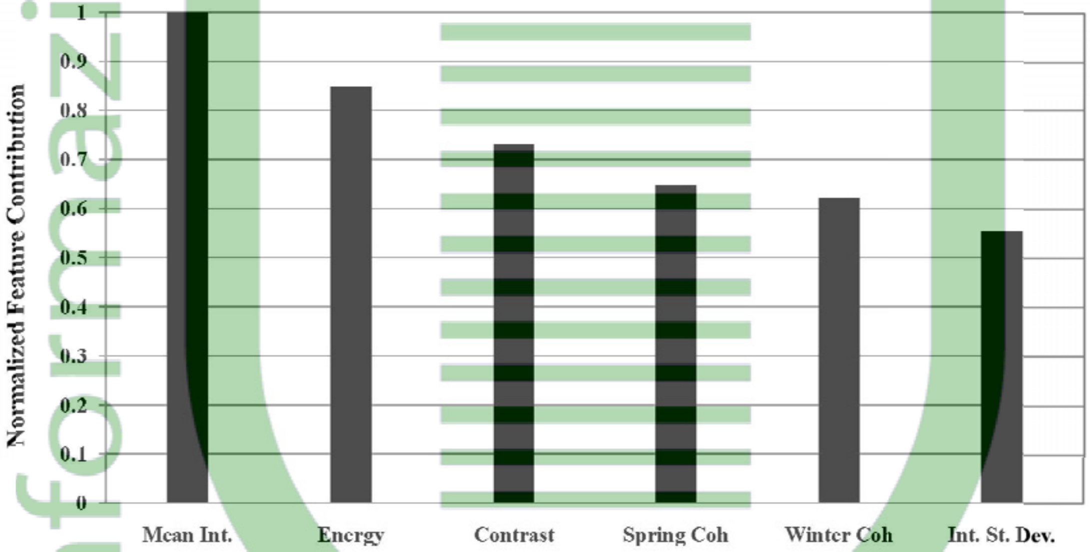
Classes	
Asphalt (AS)	
Forest (FO)	
High Density (HD)	
Isolated Buildings (IB)	
Low Density (LD)	
Vegetation (VE)	
Water (WS)	

If we merge these two classes, the overall accuracy reaches 91.5% (*K*-Coeff.=0.860) which can represent a satisfactory target for this type of application.

	AS	FO	HD	IM	LD	VE	WS
AS	47.80	0.55	0.0[illegible]	0.55	5.49	42.86	2.75
FO	0.41	71.7[illegible]	0.00	0.00	15.22	11.14	1.49
HD	0.00	0.00	43.40	0.00	56.13	0.47	0.00
IM	0.00	0.00	16.52	79.46	3.13	0.89	0.00
LD	0.02	0.64	5.92	0.67	89.28	3.48	0.00
VE	1.35	1.44	0.21	0.01	4.71	92.23	0.04
WS	0.91	6.69	0.00	0.00	0.91	2.43	86.32
	Overall Er.(%)		11.73		k-Coefficient		0.811

Feature Contribution

The relative information content of the six input quantities has been evaluated by computing the contribution of each input.



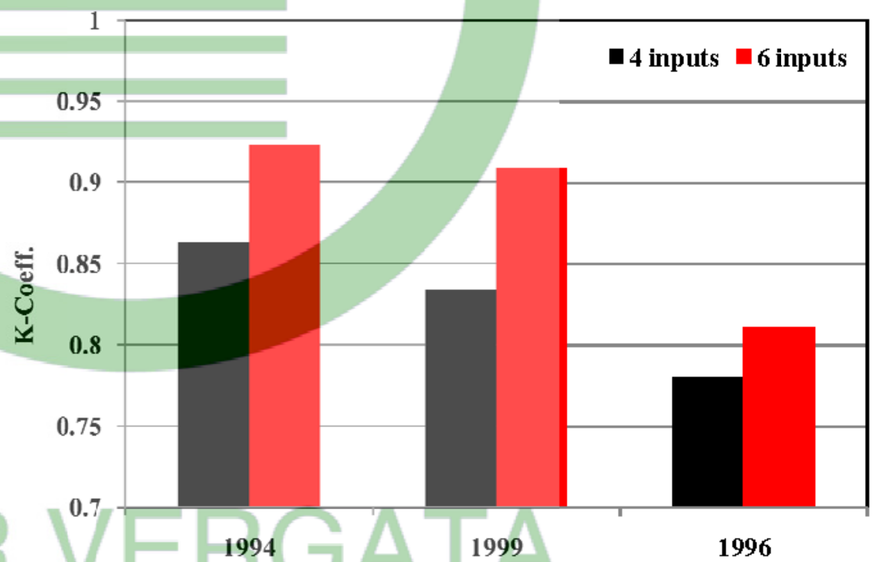
The backscattering intensity carries the maximum information, followed by energy, contrast and spring coherence, while the standard deviation of intensity contributes the least.

Conclusions

We explored the potential of single-polarization decametric SAR data by discussing the extraction of suitable features and by using them in producing land cover maps through a Neural Network algorithm.

Backscattering intensity, GLCM energy and contrast turned out to be the most effective parameters in classifying the landscape.

The accuracy in classify AUTOMATICALLY 7 types of surface from a single interferometric acquisition exceeded 86%, with a k-Coeff. larger than 0.78.



Refereed publications

Journals:

1. F. Pacifici, F. Del Frate, D. Solimini, "Monitoring Urban Land Cover in Rome, Italy, and its Changes by Single-polarization Multi-temporal SAR Images", accepted for publication on IEEE Transactions on Geoscience and Remote Sensing

Conference Proceedings:

1. F. Del Frate, F. Pacifici, D. Solimini, "Urban Land Cover in Rome, monitored by single- parameter multi-temporal images", Urban Remote Sensing Joint Event 2007, Paris, France, April 11-13, 2007

2. F. Del Frate, F. Pacifici, D. Solimini, A. Burini, "Urban land cover classification: potential of high and very high resolution SAR imagery", International Geoscience and Remote Sensing Symposium 2007, Barcelona, Spain, July 23-27, 2007

3. F. Del Frate, F. Pacifici, D. Solimini, "Use of Neural Networks for Automatic Classification of SAR imagery", ESA-EUSC 2008: Image Information Mining: pursuing automation of geospatial intelligence for environment and security, ESA/ESRIN, Frascati, Rome, Italy, March 4-6, 2008

Multi-scale textural analysis of panchromatic imagery for urban land-use (QuickBird/WorldView-1)

Goals (1/3)

Although urban areas currently cover only 2% of the land surface, they have a **global impact** due to the size of the associated energy, food, water and raw material demands. Therefore, the analysis of the urban environment represents one of the most important areas for the remote sensing community.

During the last two decades, significant progress has been made in planning and launching satellites with instruments, in both the optical/IR and microwave regions of the spectrum. These instruments are well suited for Earth observation with an increasing degree of finer resolution (spatial and spectral domains).

The introduction of data from the **WorldView 1/2** satellites is poised to make a major contribution towards the advancement of the commercial remote sensing industry.

Goals (2/3)

Remote sensing data from these systems have a potential for an increase in accurate mapping of the urban environment with a sub-meter ground resolution. At the same time, the increase in volume of data creates additional problems in terms of information extraction using automatic classification.

Urban areas are composed of numerous materials (concrete, asphalt, metal, plastic, glass, water, grass, shrubs, trees and soil) arranged by humans in complex ways to build housing, transportation systems, utilities, commercial buildings and recreational areas. Considering these conditions, a simple building may appear as a complex structure with many architectural details surrounded by gardens, trees, buildings, roads, social and technical infrastructure and many temporary objects, such as cars, buses or daily markets.

Goals (3/3)

There is no multi-spectral information in panchromatic data and even though they tend to have a higher spatial resolution, they are rarely used for automatic urban classification studies. Therefore, it is essential to use other analytical criteria to classify LAND-USE in these very high resolution panchromatic satellite images.

In this study, a multi-scale textural analysis is carried out to optimize the classification accuracy of urban land-use in very high spatial resolution panchromatic imagery.

Las Vegas – Data Set and Regions Of Interest

QB - May 10, 2002

Urban environment with regular structures

Presence of cars in the parking lots

Small shadows

Medium off-nadir angle (12.8°)

Land-Use Classes	TR	VS
Bare Soil	4255	44675
Commercial	1822	19126
Drainage Channel	1143	12001
Highway	2836	29774
Parking Lots	2257	23695
Residential	7007	73563
Roads	6098	64023
Short Vegetation	1793	18823
Soil	1472	15437
Trees	1043	10945
Water	118	1236
TOTAL ROIs	29844	313298

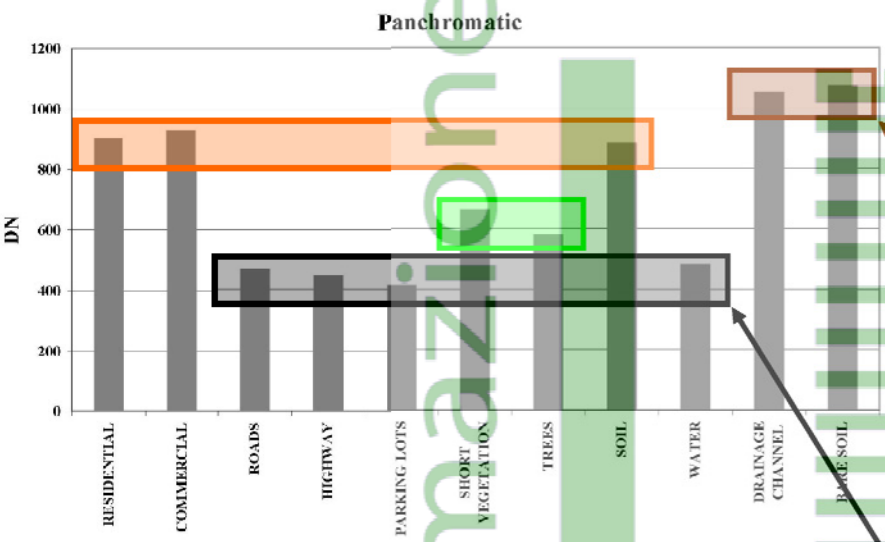
Las Vegas – Classification Map (Panchromatic ONLY)

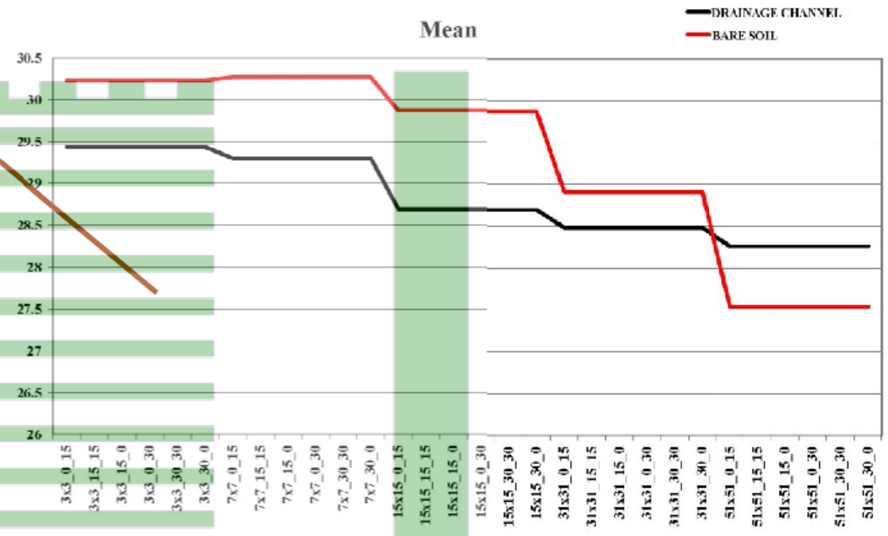
Land-Use Classes
Bare Soil
Commercial
Drainage Channel
Highway
Parking Lots
Residential
Roads
Short Vegetation
Soil
Trees
Water

Overall Error = 50.2%

***k*-Coeff. = 0.378**

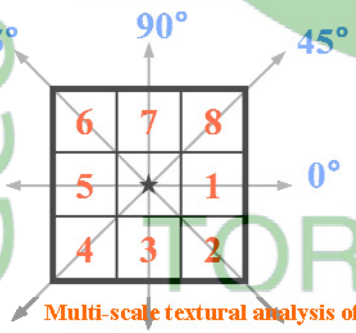
Las Vegas – Textural Analysis

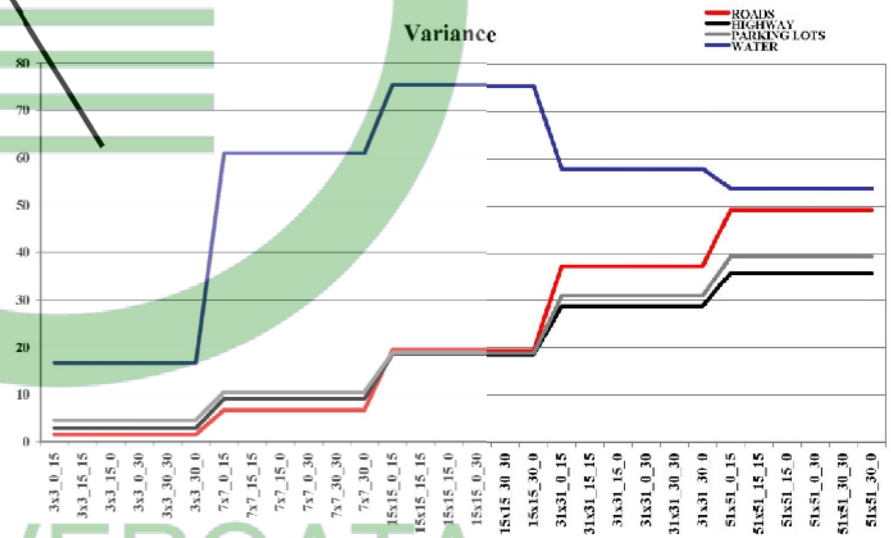




We analyze the contextual information of each pixel by mean of 8 textural features.

Textural Features	Windows Size (pixel)	Steps (pixel)	Directioins (°)	Total Inputs
Mean	3x3	15	0	241
Variance	7x7	30	45	
Homogeneity	15x15		90	
Contrast	31x31			
Dissimilarity	51x51			
Entropy				
Second Moment				
Correlation				





Las Vegas – Homogeneity

15x15_0_15

Vertical structures

15x15_15_0

Horizontal structures

15x15_15_15

Wide structures

31x31_0_15

Vertical structures

Las Vegas – FINAL Classification Map (1/2)

The classification process with a large input space **rarely** yields high classification accuracies due to **information redundancy** of certain inputs.

Feature selection based on Pruning Neural Networks (which eliminates the weakest connections) was used to reduce the number of textural features. The remaining input features totaled 169.

Las Vegas – FINAL Classification Map (2/2)

Land-Use Classes
Bare Soil
Commercial
Drainage Channel
Highway
Parking Lots
Residential
Roads
Short Vegetation
Soil
Trees
Water

Overall Error = 6.8%

***k*-Coeff. = 0.920**

Rome – Data Set and Regions Of Interest

QB - July 19, 2004

2 Different urban environments:
- old style architecture
- new style architecture

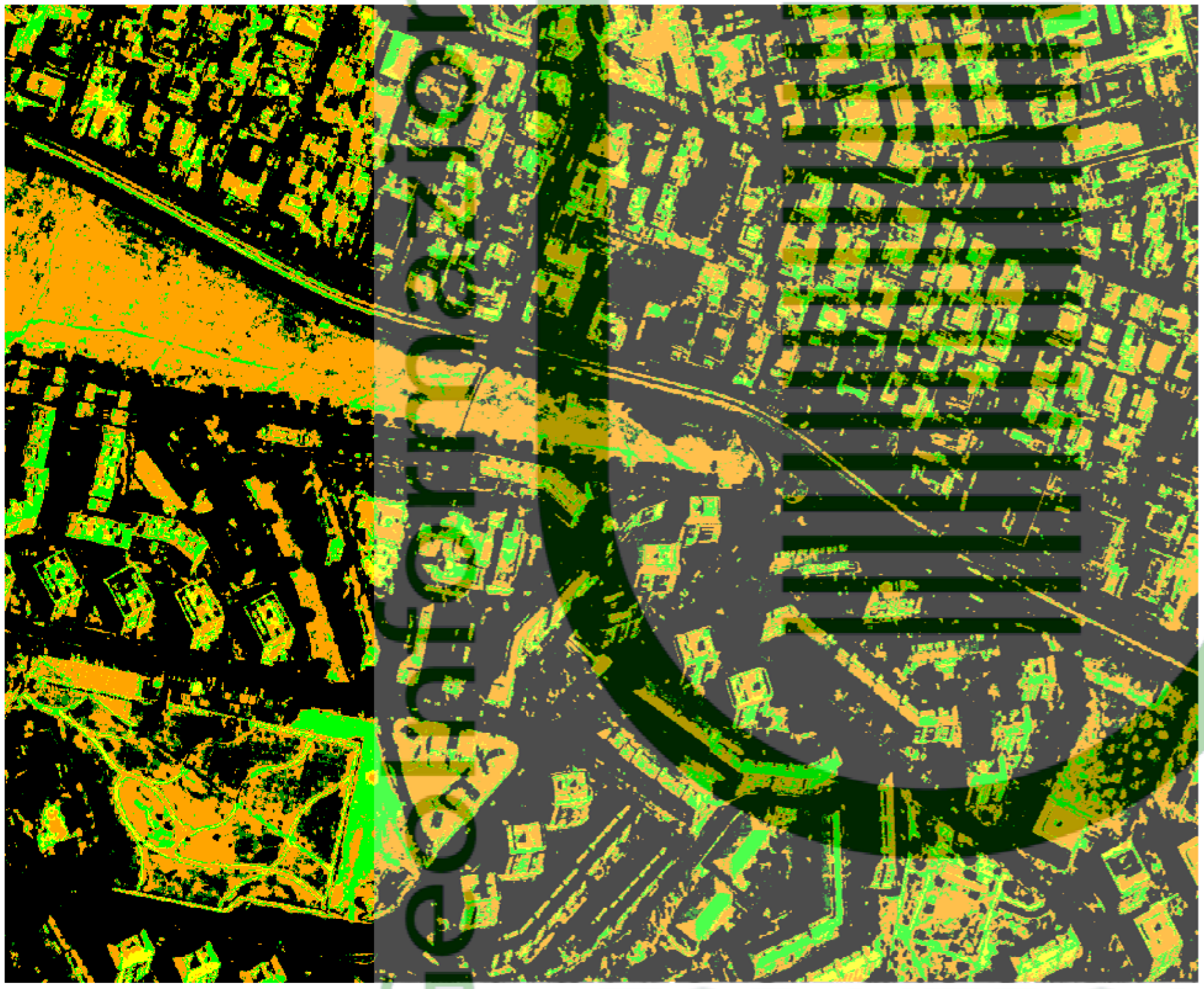
Many temporary objects:
- cars
- buses

Long shadows

High off-nadir angle (23°)

Land-Use Classes	TR	VS
Bare Soil	4127	38572
Blocks	20472	44672
Buildings	27188	77034
Light Train	2606	6727
Roads	35531	69002
Soil	3506	5776
Tower	9187	19365
Trees	13632	38624
Short Vegetation	10443	29587
TOTAL ROIs	126692	329359

Rome – Classification Map (Panchromatic ONLY)

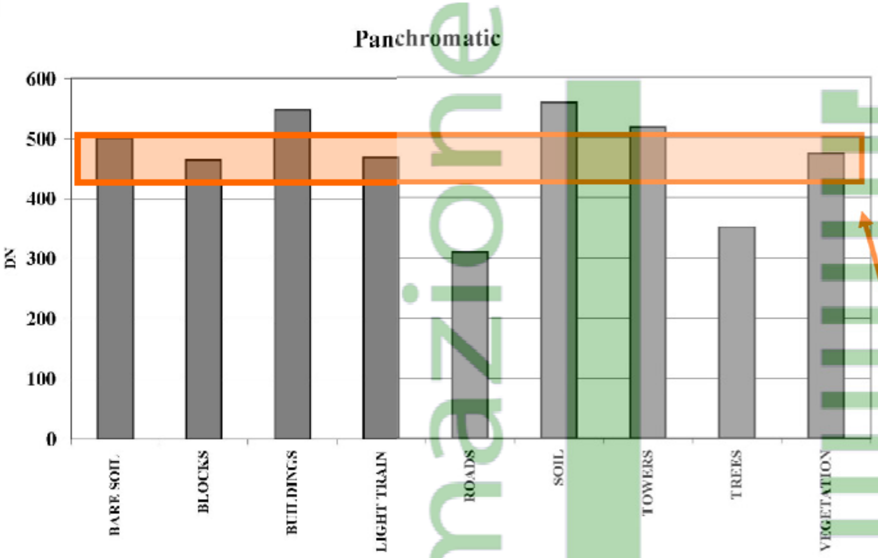


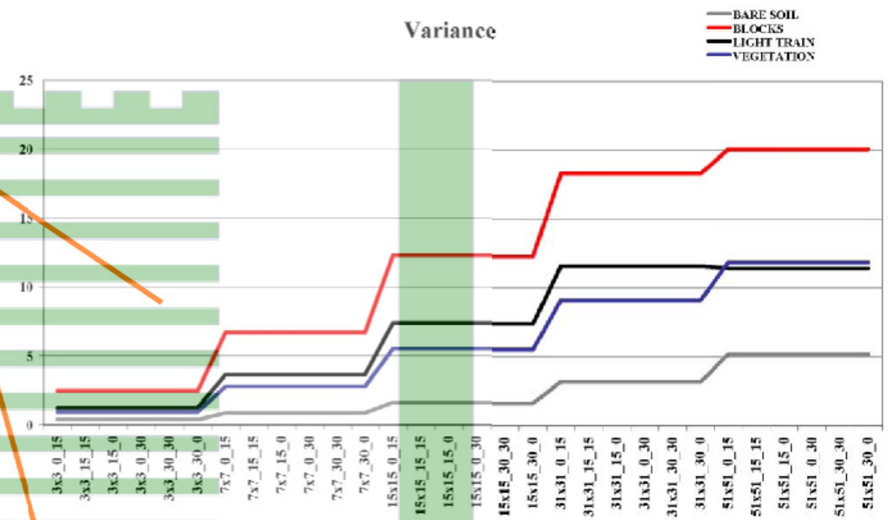
Land-Use Classes	
	Bare Soil
	Blocks
	Buildings
	Light Train
	Roads
	Soil
	Tower
	Trees
	Short Vegetation

Overall Error = 66.0%

***k*-Coeff. = 0.184**

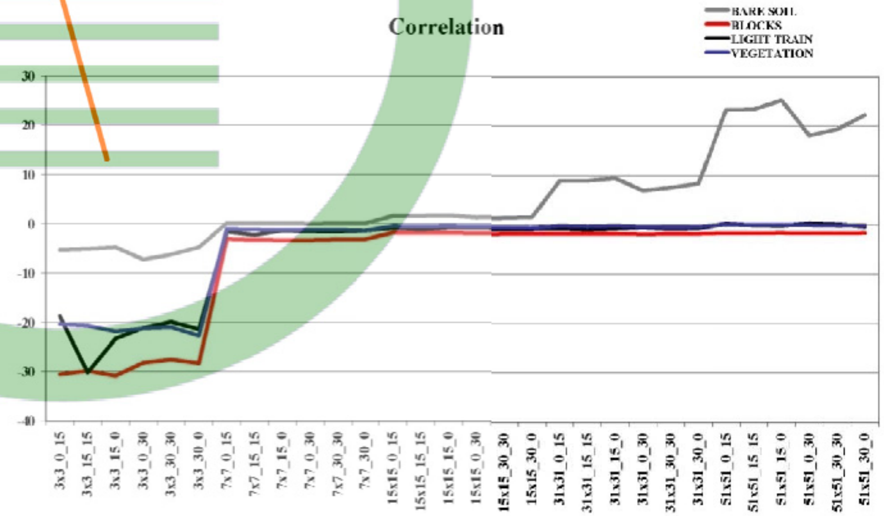
Rome – Textural Analysis





We analyze the contextual information of each pixel by mean of 8 textural features.

Textural Features	Windows Size (pixel)	Steps (pixel)	Directioins (°)	Total Inputs
Mean	3x3	15	0	241
Variance	7x7	30	45	
Homogeneity	15x15		90	
Contrast	31x31			
Dissimilarity	51x51			
Entropy				
Second Moment				
Correlation				



Rome – Dissimilarity

7x7_15_15 | 15x15_15_15

31x31_15_15 | 51x51_15_15

Rome – FINAL Classification Map

Again, feature selection was applied to reduce the number of textural features eliminating the redundant inputs. The selected input features were 140.

Land-Use Classes
Bare Soil
Blocks
Buildings
Light Train
Roads
Soil
Tower
Trees
Short Vegetation

Overall Error = 5.0%

k-Coeff. = 0.941

Extended Pruning (1/3)

The **EXTENDED PRUNING** is the process of eliminating the **least contributing inputs** in order to identify the optimal textural features to be used in successive classifications.

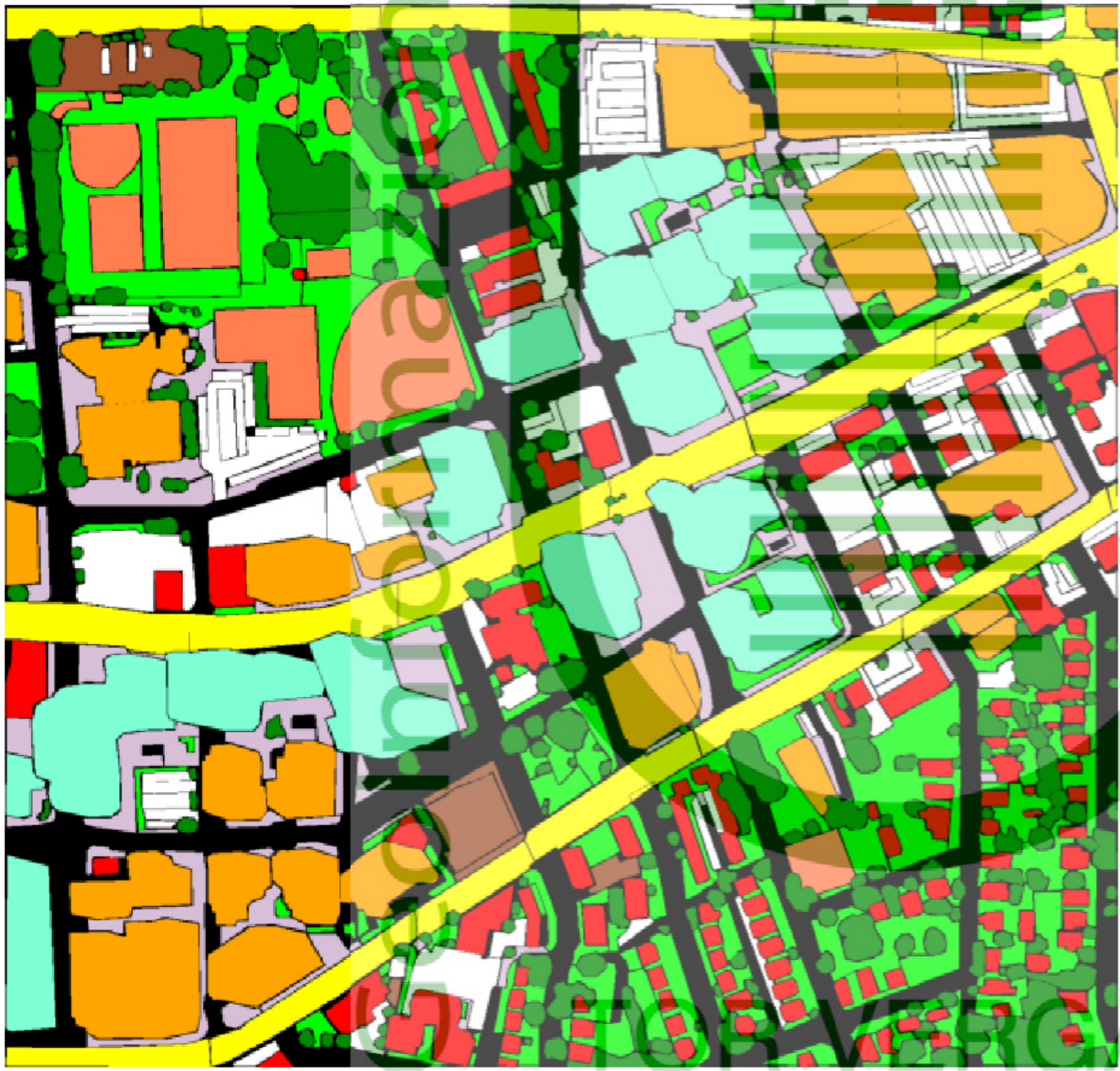
This further input reduction is obtained by **accepting** a decrease in the classification accuracy.

LAS VEGAS	Classification Error (%)	*k*-Coeff.	Inputs	Connections
Panchromatic	50.2	0.378	1	20
Full NN	7.1	0.916	241	4820
Pruned NN	6.8	0.920	169	1227
Ext. Pruning	12.0	0.859	59	183

ROME	Classification Error (%)	*k*-Coeff.	Inputs	Connections
Panchromatic	66.0	0.184	1	40
Full NN	16.9	0.798	241	9640
Pruned NN	5.0	0.941	140	2110
Ext. Pruning	15.1	0.820	61	652

Washington – Data Set and Regions Of Interest

WV1 - December 18, 2007



3 Different urban environments:
- small residential houses
- large buildings
- tall large buildings

Very long shadows
(24.9° sun elevation)

Very high off-nadir angle
(27.8°)

Winter image:
- trees without leaves

Land-Use Classes	TR	VS
Buildings	24178	76159
Highway	17985	56653
Parking Lots	17019	53611
Residential	14195	44714
Roads	20618	64946
Soil	2553	8043
Sport Facilities	8270	26051
Tall Buildings	21047	66297
Trees	18535	58386
Vegetation	23403	73720
Walk side	12203	38439
TOTAL ROIs	180006	567019

Washington – Classification Map (Panchromatic ONLY)

Land-Use Classes	
	Buildings
	Highway
	Parking Lots
	Residential
	Roads
	Soil
	Sport Facilities
	Tall Buildings
	Trees
	Vegetation
	Walk side

Overall Error = 31.4%

k-Coeff. = 0.187

Washington – FINAL Classification Map

Land-Use Classes
Buildings
Highway
Parking Lots
Residential
Roads
Soil
Sport Facilities
Tall Buildings
Trees
Vegetation
Walk side

Overall Error = 87.9%

***k*-Coeff. = 0.864**

Conclusions

The **multi-scale textural analysis** made it possible to classify urban LAND-USE on a **per-pixel** basis overcoming the spectral information deficit of panchromatic imagery.

The obtained classification maps demonstrated that it is possible classify urban LAND-USE with accuracies over 85% in terms of *k*-coefficient using only information from neighborhood pixels at multi-scale levels. Significantly, it is possible to distinguish between different asphalt surfaces, such as **ROADS**, **HIGHWAYS** and **PARKING LOTS**, or different architectures, such as **BUILDINGS**, **APARTMENT BLOCKS** and **TOWERS**.

The analysis of the results indicates the importance of using window dimensions greater than **15x15 pixels** for images with a 50-60 cm resolution. **Dissimilarity** and **Contrast** turned out to be the most important textural features.

Refereed publications

Conference Proceedings:

1. F. Pacifici, M. Chini, W. J. Emery, "Urban land-use textural analysis for very high resolution optical imagery", ESA-EUSC 2008: Image Information Mining: pursuing automation of geospatial intelligence for environment and security, ESA/ESRIN, Frascati, Rome, Italy, March 4-6, 2008

2. F. Pacifici, M. Chini, W. J. Emery, "Urban land-use multi-scale textural analysis", International Geoscience and Remote Sensing Symposium 2008, Boston, USA, July 6-11, 2008

Seminaries:

1. Multi-scale Textural Analysis of Panchromatic Imagery: Urban Land-Use Classification, DigitalGlobe, Longmont, Colorado, USA, November 26, 2007

Change detection of urban areas with very high resolution optical imagery (QuickBird)

Motivation (1/2)

The analysis of the urban environment represents one of the most challenging areas for the remote sensing community. With the availability of commercial *very high resolution* satellite imagery, it is possible to identify small-scale features such as single houses or small structures.

A typical urban surface is composed primarily of roofs and roads characterized by different:

- age
- quality
- composition

Moreover, bare soil and vegetation, which may include trees, plants and parks, make the landscape more complicated.

Motivation (2/2)

Different categories of change can be identified when comparing two or more acquisitions of the same scene:

- newly built houses, roof variations and widened roads which are *important information* for public housing authorities
- streets with or without traffic jams, trees with more or less leaves, or parking areas where cars occupy different places can result in a simple modification of an existing object and not a conversion from one object to another. These transient features are *not important* change detection targets

Moreover, *technical aspects* should be further taken into account:

- different solar conditions (shadowing)
- different off-nadir angles

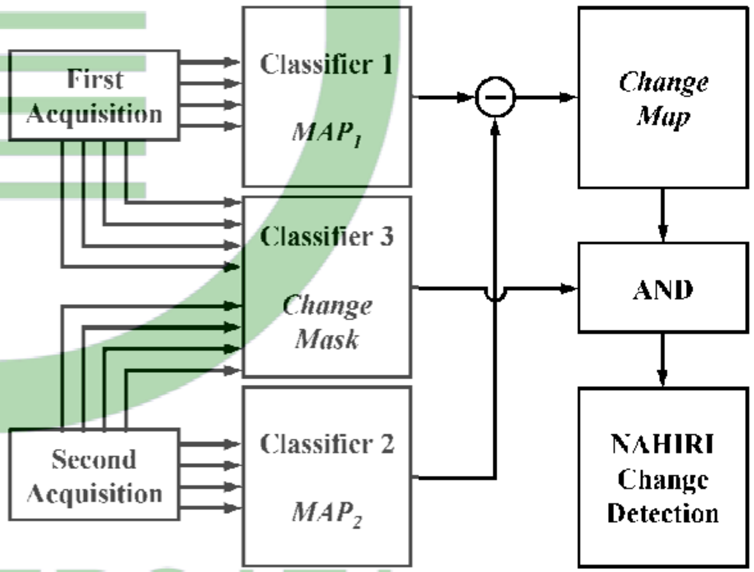
may cause false signals which increase the difficulty of interpreting changes.

Goals

Effectiveness and robustness of a new method for urban change detection that greatly reduces the human effort needed to analyze the imagery:

NAHIRI: Neural Architecture for HIgh Resolution Imagery

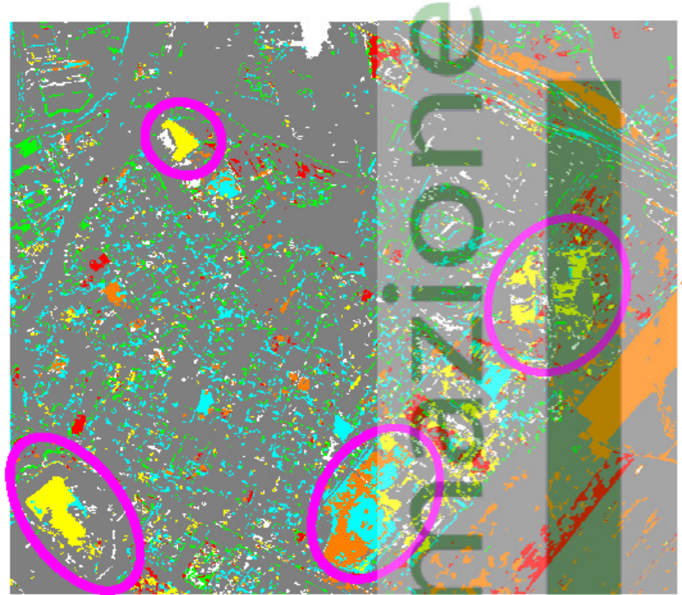
is a change detection algorithm based on Neural Networks able to exploit in parallel both the *multi-spectral* and the *multi-temporal* information to discriminate between real changes and false alarms.



Test Area 1: Campus of the Tor Vergata University

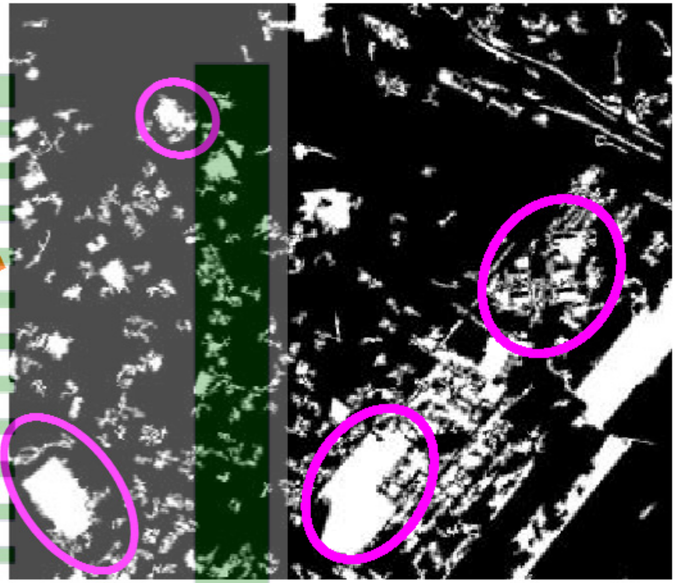
The campus of the Tor Vergata University is located in the upper part of the image (2002)

Test Area 1: Campus of the Tor Vergata University



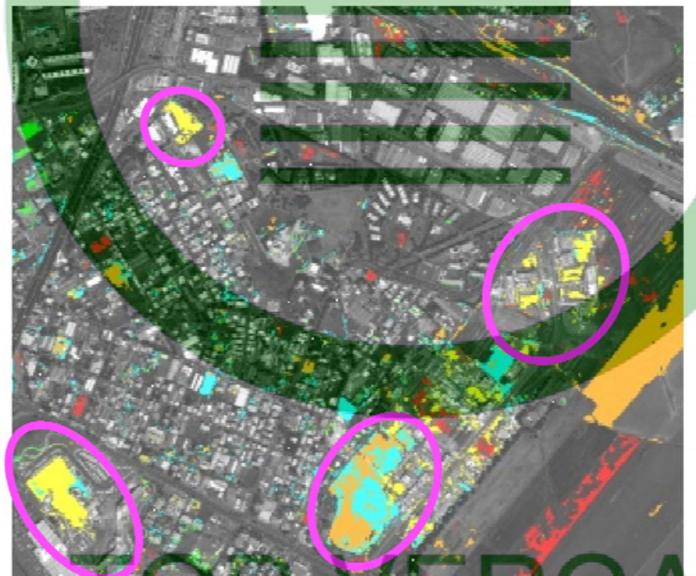
Change Map

AND



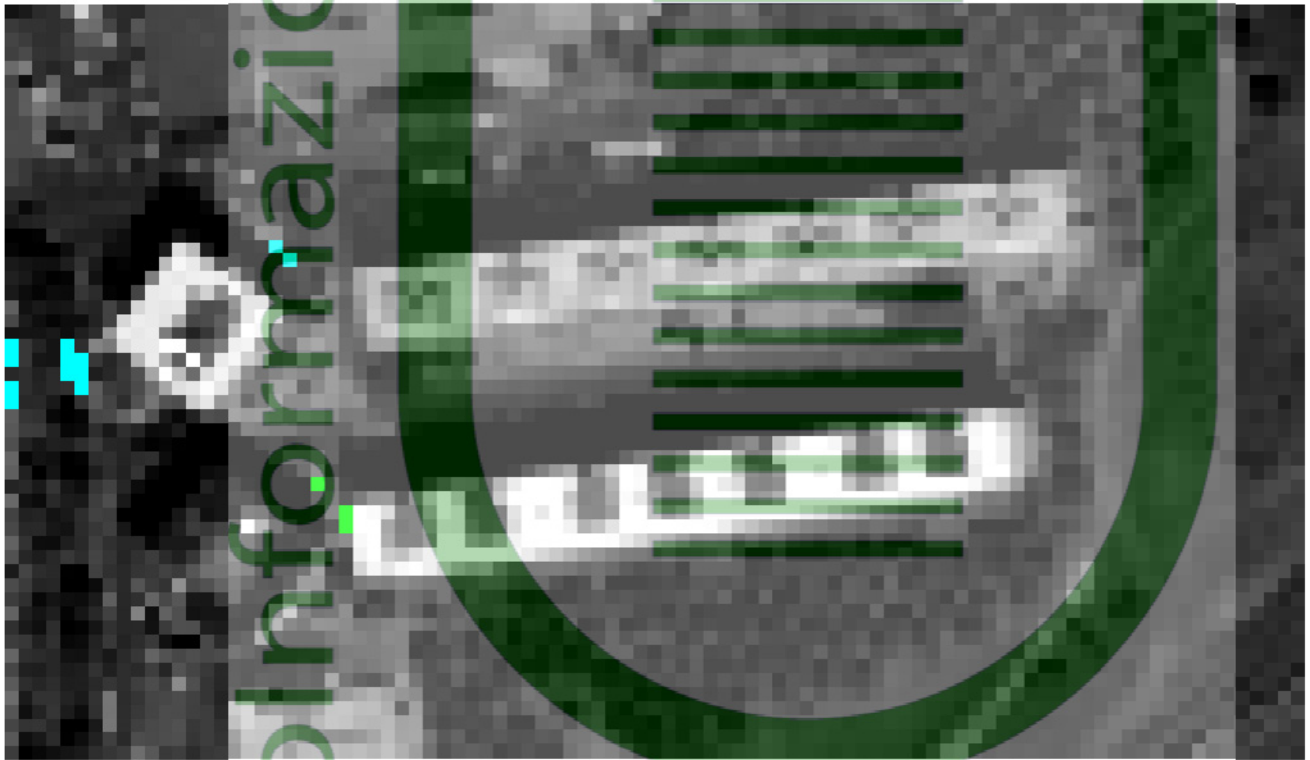
Change Mask

	2003		
2002	Vegetation	Man-made	Soil
Vegetation	Gray	Cyan	Orange
Man-made	Green	Gray	White
Soil	Red	Yellow	Gray



NAHIRI CD

Test Area 1: Campus of the Tor Vergata University



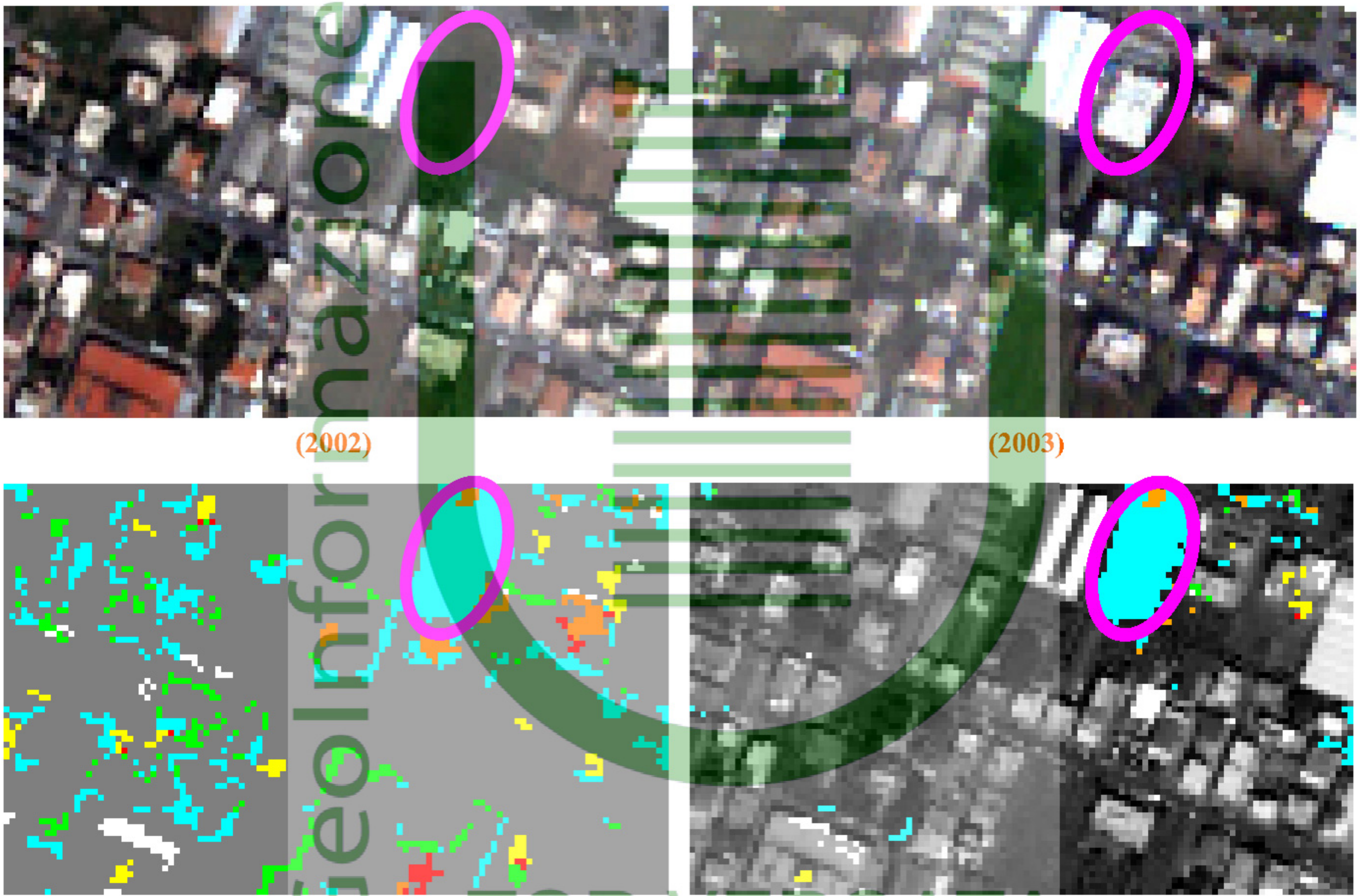
(2002)

(2003)

PCC

NAHIRI

Test Area 1: Campus of the Tor Vergata University



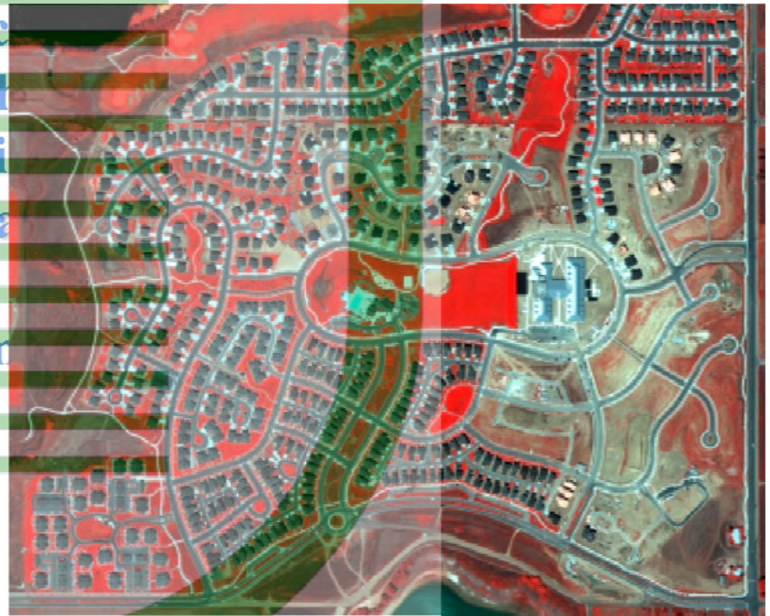
Change detection of urban areas with very high resolution optical imagery

Effectiveness : Test Area 2

Why this site?

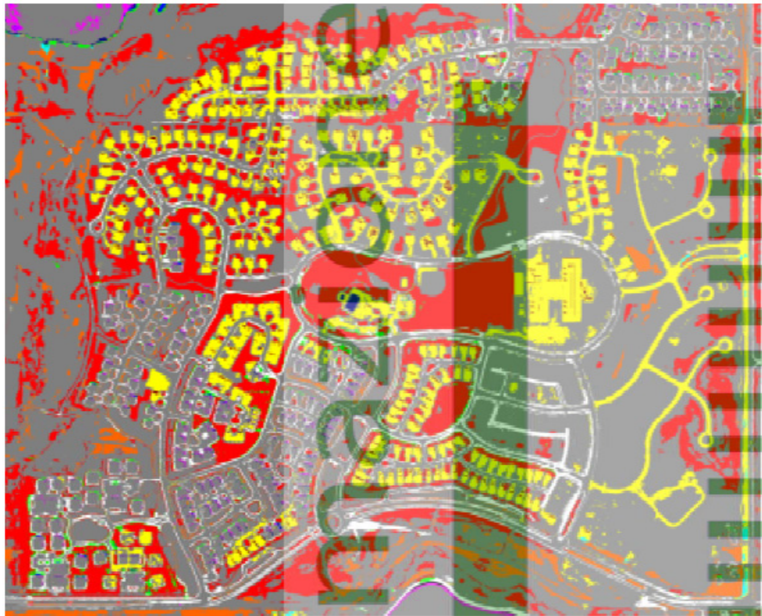
the c
ily
uildi
spatia
a m

Test Area 2 (2002)



Test Area 2 (2004)

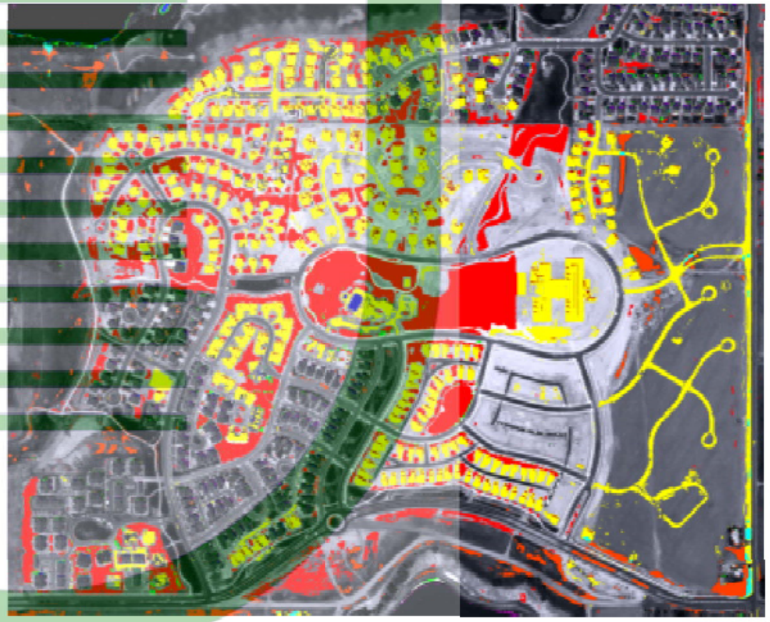
Effectiveness : Test Area 2

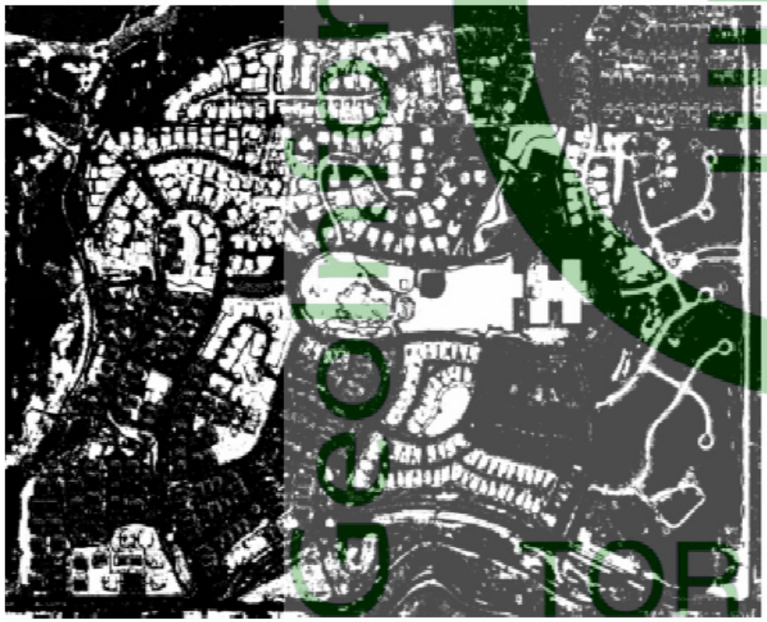


Change Map

	2004			
2002	Vegetation	Man-made	Water	Soil
Vegetation	Gray	Cyan	Dark Green	Orange
Man-made	Green	Gray	Dark Blue	White
Water	Blue	Magenta	Gray	Black
Soil	Red	Yellow	Brown	Gray

NAHIRI CD





Change Mask

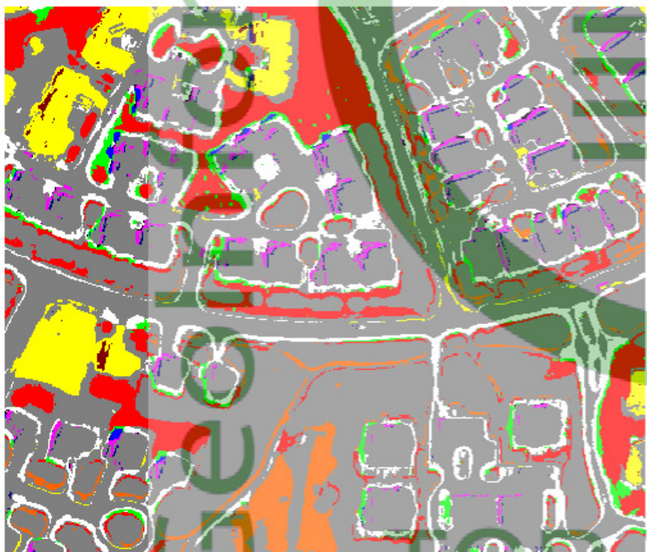
This area has an extension of 1.3 km^2 with 0.6 m of resolution.

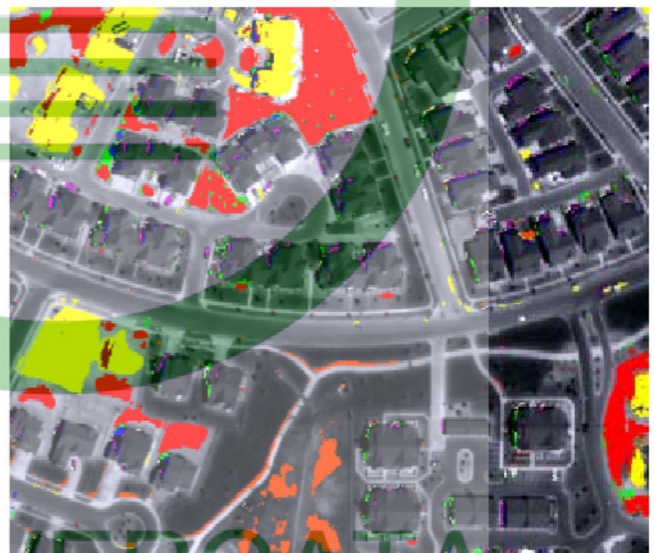
Effectiveness : Test Area 2

Boulder Area (2002)

Boulder Area (2004)

PCC



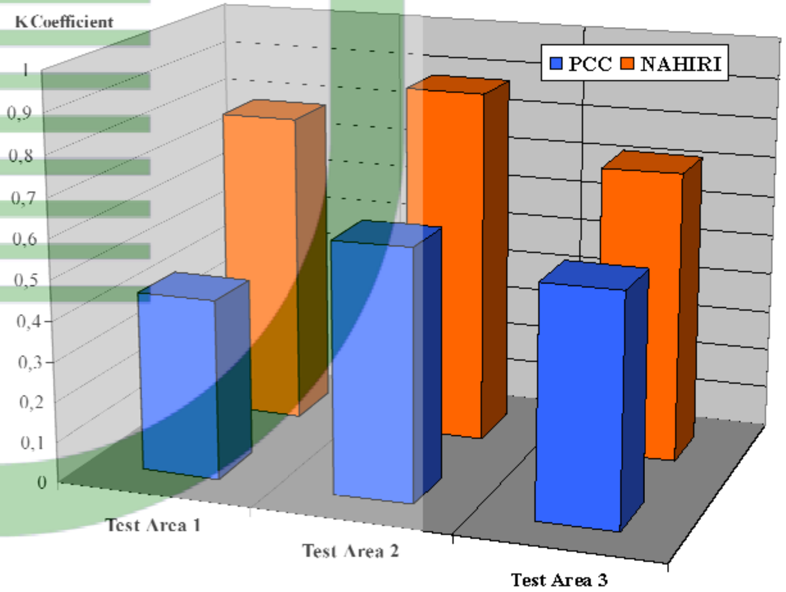


NAHIRI

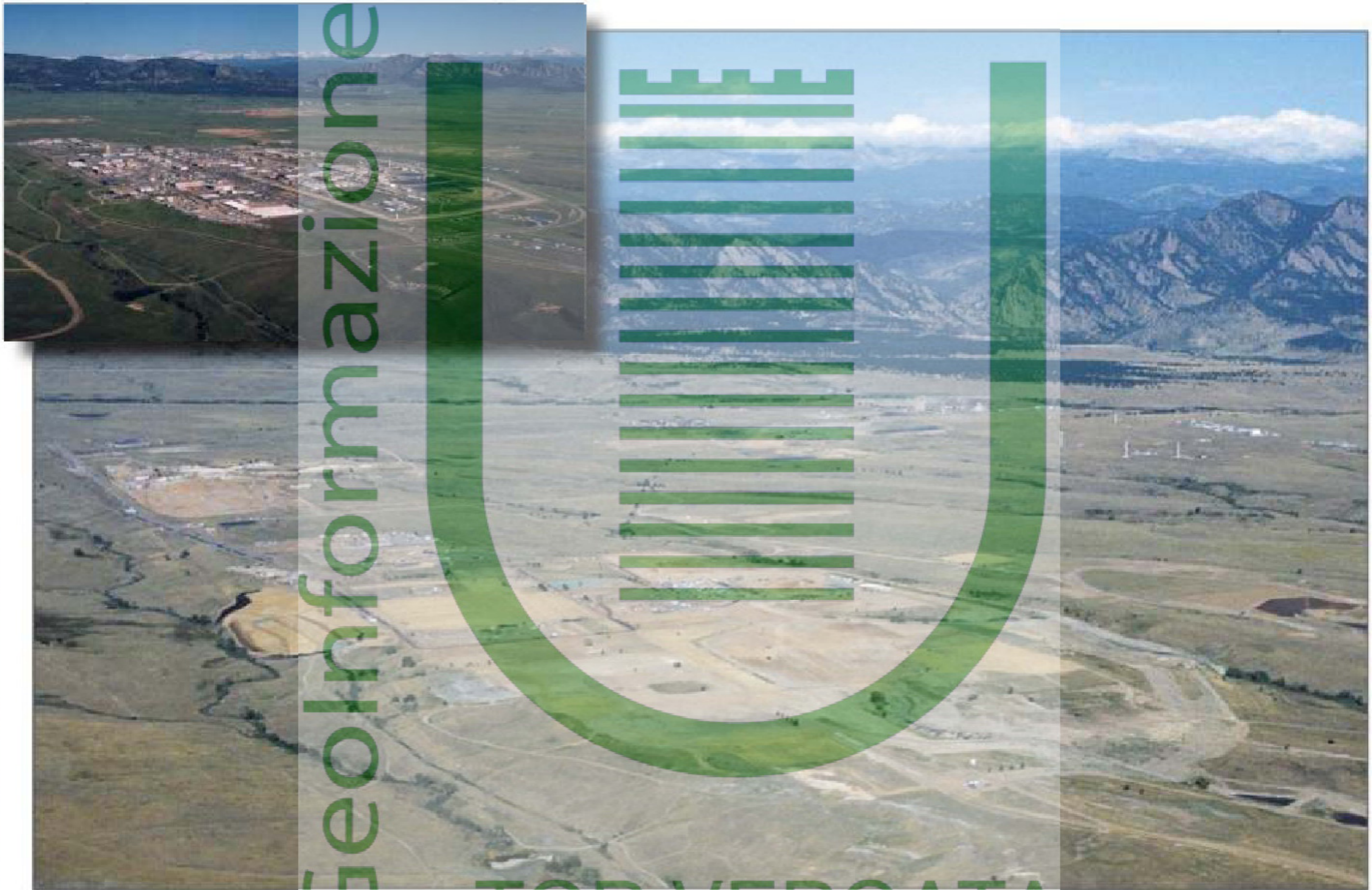
Effectiveness: Accuracy Assessment

The mean of the *k-Coefficient* ranges from 0.544 in the case of PCC to 0.795 (NAHIRI) over very high and high resolution optical imagery.

Location	Spatial Res. (m)	NAHIRI	PCC
		k-Coefficient	
Tor Vergata Campus, Rome, Italy	2.8	0.783	0.444
Superior, Colorado, U.S.A.	30	0.881	0.619
Superior, Colorado, U.S.A.	0.6	0.722	0.568
	Mean	0.795	0.544



Robustness: Test Area 3



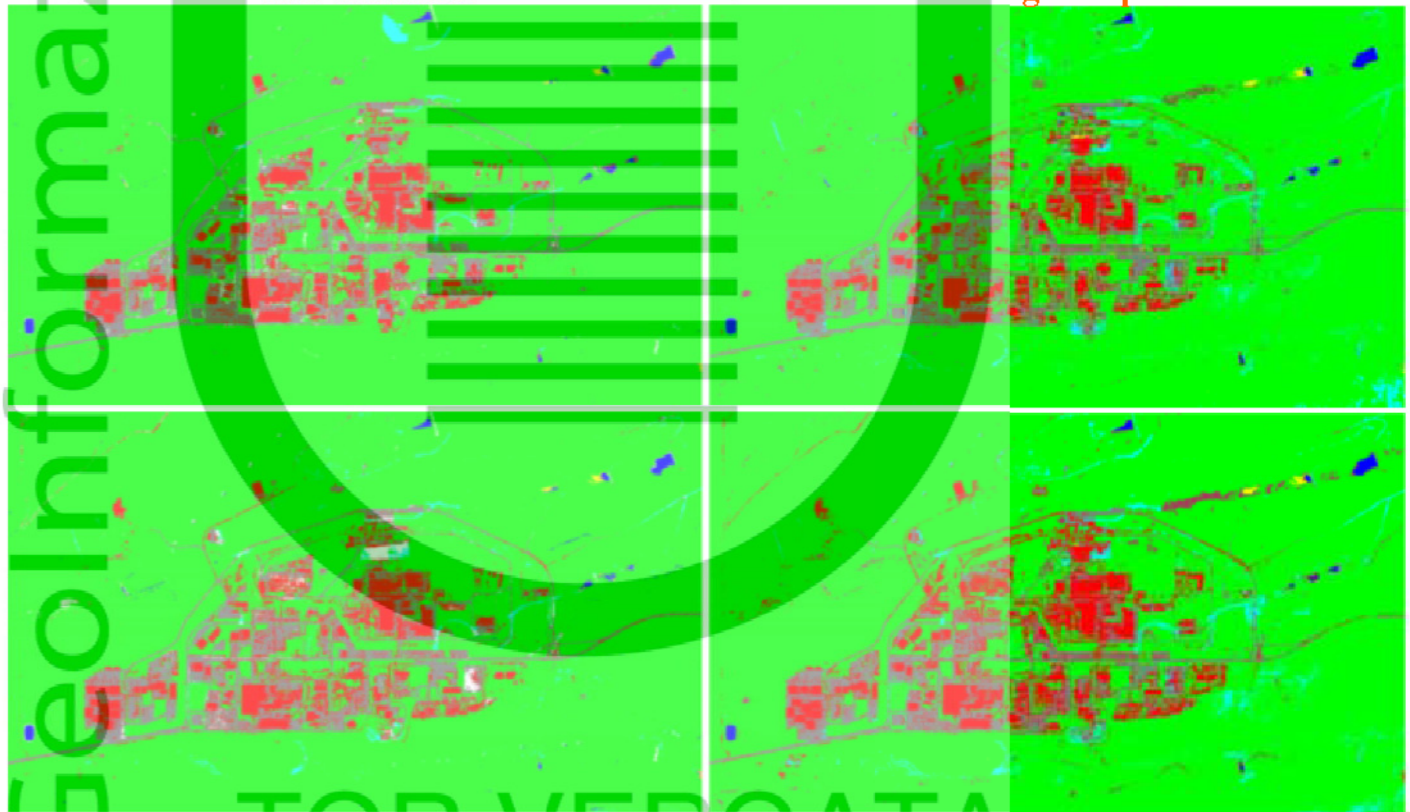
Test Area 3 (about 2008)

Comparative analysis between Neural Networks and Maximum Likelihood methods within the NAHIRI framework to demonstrate the robustness of this method when different classifiers are used.

Method	*k*-coeff	Improv.
NN-single step	0.666	14%
NN-NAHIRI	0.758	
ML-single step	0.585	12%
ML-NAHIRI	0.654	

Class		
1	Buildings to Soil	
2	Steady Soil	
3	Steady Water	
4	Water to Soil	
5	Soil to Asphalt	
6	Asphalt to Soil	
7	Soil to Water	
8	No Relevant	

NN-NAHIRI NN-single step



ML-NAHIRI ML-single step

Conclusions

We developed a novel Neural Network-based architecture for change detection which is able to process simultaneously multi-temporal and multi-spectral information.

Effectiveness:

The mean of the *k-Coefficient* ranges from 0.544 in the case of PCC to 0.795 (NAHIRI) over very high and high resolution optical imagery.

Robustness:

The comparative analysis between Neural Networks and Maximum Likelihood demonstrated an enhancement of about 12% between the single step ML-classification and the ML-based NAHIRI implementation (from 0.585 to 0.654, respectively) confirming the robustness of the method when different classifiers are used.

Refereed publications of NAHIRI

Books:

1. F. Pacifici, F. Del Frate, C. Solimini, W. J. Emery, "Neural Networks for Land Cover Applications", in Computational Intelligence for Remote Sensing, Manuel Grana, Richard Duro, Eds. Studies in Computational Intelligence, Spinger, in Press

Journals:

1. F. Pacifici, F. Del Frate, C. Solimini, W. J. Emery, "An Innovative Neural-Net Method to Detect Temporal Changes in High-Resolution Optical Satellite Imagery", IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 9, pp. 2940-2952, Sept. 2007

2. M. Chini, F. Pacifici, W. J. Emery, N. Pierdicca, F. Del Frate, "Comparing statistical and neural network methods applied to very high resolution satellite images showing changes in man-made structures at Rocky Flats", IEEE Transactions on Geoscience and Remote Sensing vol. 46, no. 6, Jun. 2008

Conference Proceedings:

1. F. Del Frate, F. Pacifici, C. Solimini, W. J. Emery, "Urban Change Detection by Using Neural Networks Algorithms and High Resolution Optical Satellite Imagery", International Geoscience and Remote Sensing Symposium 2006, Denver, Colorado, U.S.A., August 2006

2. F. Del Frate, F. Pacifici, C. Solimini, W. J. Emery, "A new neural network architecture for automatic Urban Change Detection From Satellite Imagery", ESA-EUSC 2006: Image Information Mining for Security and Intelligence, EUSC, Torrejon Air Base, Madrid, Spain, November 2006

3. F. Pacifici, F. Del Frate, C. Solimini, W. J. Emery, "A New Neural Architecture for Detecting Urban Changes in Quickbird Imagery", Proceedings of Urban Remote Sensing Joint Event 2007, Paris, France, 11-13 April 2007

4. M. Chini, W. J. Emey, F. Pacifici, "Change mapping in the Rocky Flats area as test bed for damage detection algorithms", Geophysical Research Abstracts, Vol. 9, 06607, 2007, Vienna, Austria, 15-20 April 2007

5. F. Pacifici, F. Del Frate, C. Solimini, W. J. Emery, "A robust neural network design for detecting changes from multispectral satellite imagery", Proceedings of International Geoscience and Remote Sensing Symposium 2007, Barcelona, Spain, 23-27 July 2007

6. M. Chini, W. J. Emey, F. Pacifici, "Satellite mapping of the demolition of the Rocky Flats nuclear weapons plant", Proceedings of International Geoscience and Remote Sensing Symposium 2007, Barcelona, Spain, 23-27 July 2007

Seminaries:

1. Neural Network algorithms for urban change detection, Université de Lausanne, Lausanne, Switzerland, 7 June 2007
2. NAHIRI framework for urban change detection, DigitalGlobe, Longmont, Colorado, USA, July 24, 2007

Change detection of urban areas with aerial optical imagery

Pulse-coupled neural network (PCNN)

The pulse-coupled neural network (PCNN) is a relatively new technique based on the implementation of the mechanisms underlying the visual cortex of small mammals.

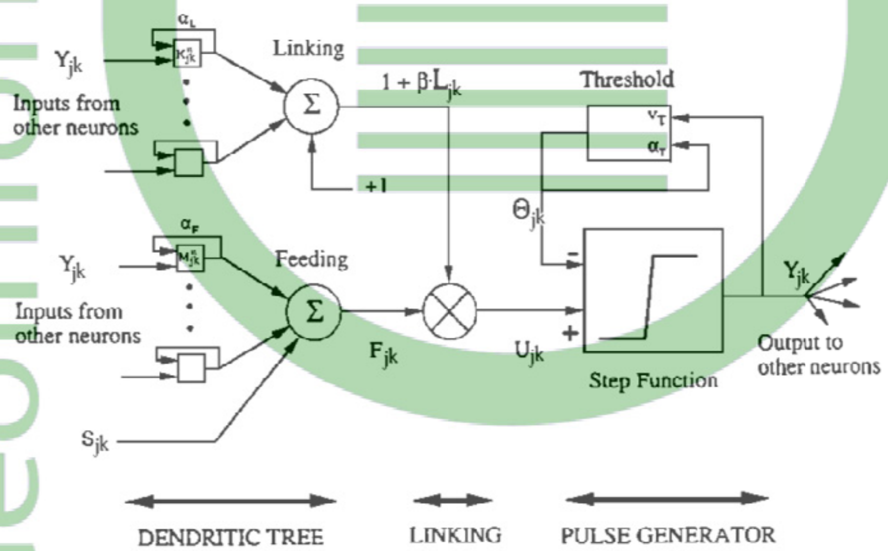
PCNN is an algorithm that produces a series of binary pulse images when stimulated with a gray scale or color input.

It is different from what we generally mean with artificial neural networks in the sense that it does not need to be trained.

The Neuron Model (1/2)

The network consists of multiple nodes coupled together with their neighbors within a definite radius, forming a grid (2Dvector).

The PCNN neuron has two input compartments: linking and feeding. The feeding compartment receives both an external and a local stimulus, whereas the linking compartment only receives a local stimulus.



The internal activity rises until it becomes larger than an active threshold value. Then the neuron fires and the threshold will decay until once again the internal activity becomes larger. Such a process gives rise to the pulsing nature of the PCNN.

$$F_{ij}[n] = e^{\alpha_F \delta_n} F_{ij}[n-1] + S_{ij} + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1]$$

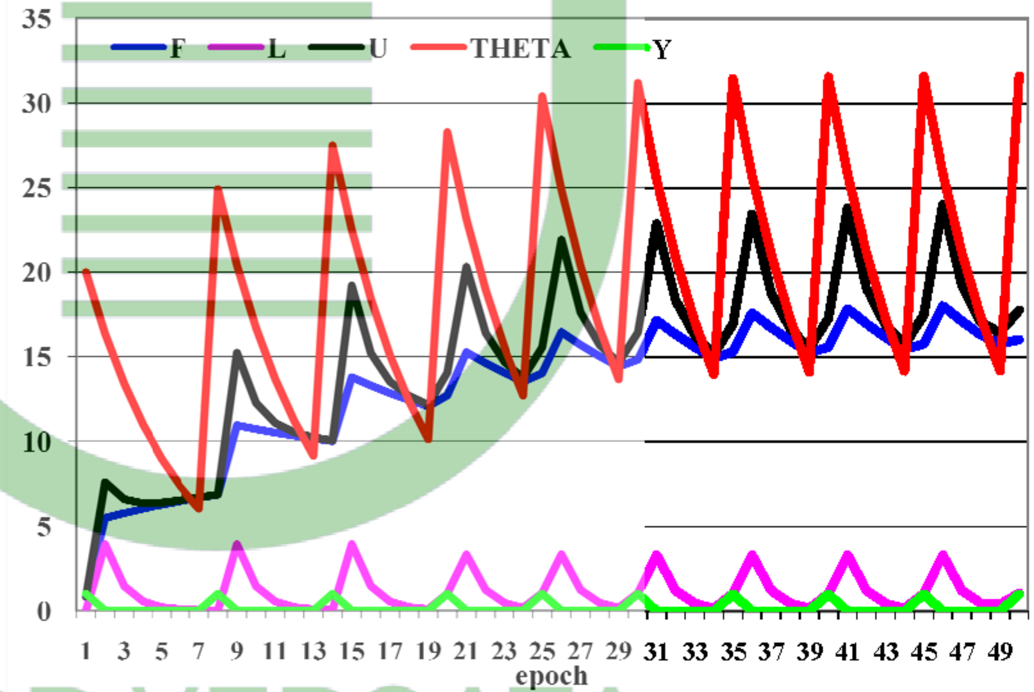
$$L_{ij}[n] = e^{\alpha_F \delta_n} L_{ij}[n-1] + V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1]$$

$$U_{ij}[n] = F_{ij}[n]\{1 + \beta L_{ij}[n]\}$$

$$Y_{ij}[n] = \begin{cases} 1 \; if \; U_{ij}[n] > \Theta_{ij}[n-1] \\ 0 \quad Otherwise \end{cases}$$

$$\Theta_{ij}[n] = e^{\alpha_\Theta \delta_n} \Theta_{ij}[n-1] + V_\Theta Y_{ij}[n]$$

$$G[n] = \sum_{i,j} Y_{ij}[n]$$



Change Detection

The waves generated in a moving window by each iteration of the algorithm create specific signatures of the scene which are successively compared for the change detection.

We considered very-high resolution aerial imagery (20 cm). False alarms due to different geometrical views (misregistration or different acquisition angle) are typical for this kind of data.

SIGNATURE ANALYSIS – The signal associated to the PCNN is invariant to changes in rotation, scale, shift, or skew of an object within the scene. These features make PCNN a promising tool for sub-metric change detection applications. Three different test cases are investigated: No/Light changes, Changes and Huge Changes.

Change Detection

No/ Light Changes

Changes

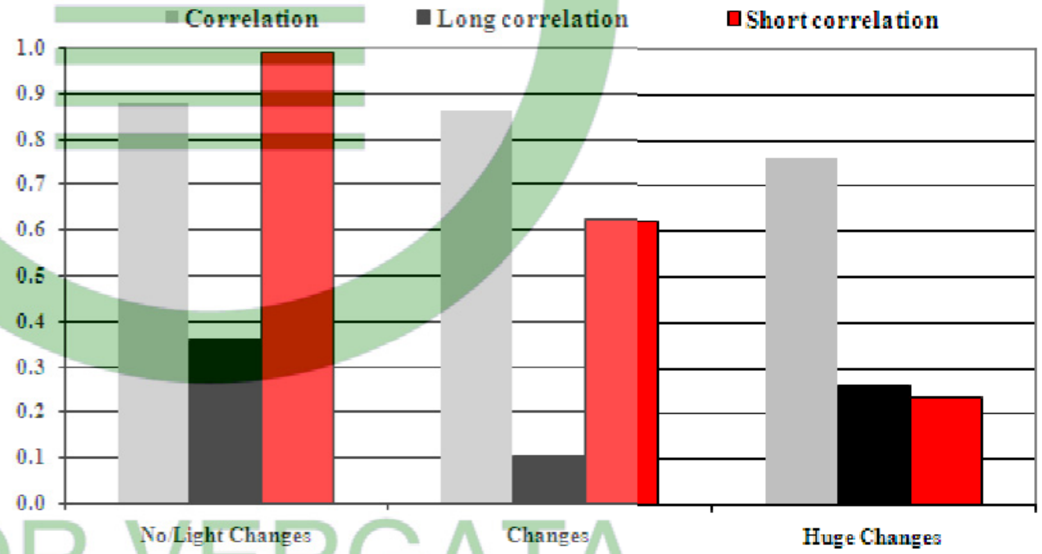
Huge Changes

Change detection of urban areas with aerial optical imagery

It takes several iterations before the threshold values decay enough to allow the neuron to fire. Experimental results for the considered cases showed that neurons fired after 7 iterations (i.e., similar signatures in the interval 1-6).

Therefore, the correlation between the different signatures can be further investigated considering all iterations or part of them. In particular, we considered three intervals: 1-50 for Correlation, 21-50 for Long Correlation and 7-20 for Short Correlation.

Among these features, Short Correlation only reflects the levels of change, ranging from 0.992 to 0.238.



Conference Proceedings:

1. F. Del Frate, F. Pacifici, D. Solimini, "Automatic change detection with Pulse Coupled Neural Networks", ESA-EUSC 2008: Image Information Mining: pursuing automation of geospatial intelligence for environment and security, ESA/ESRIN, Frascati, Rome, Italy, March 4-6, 2008

WorldView-1 colorization

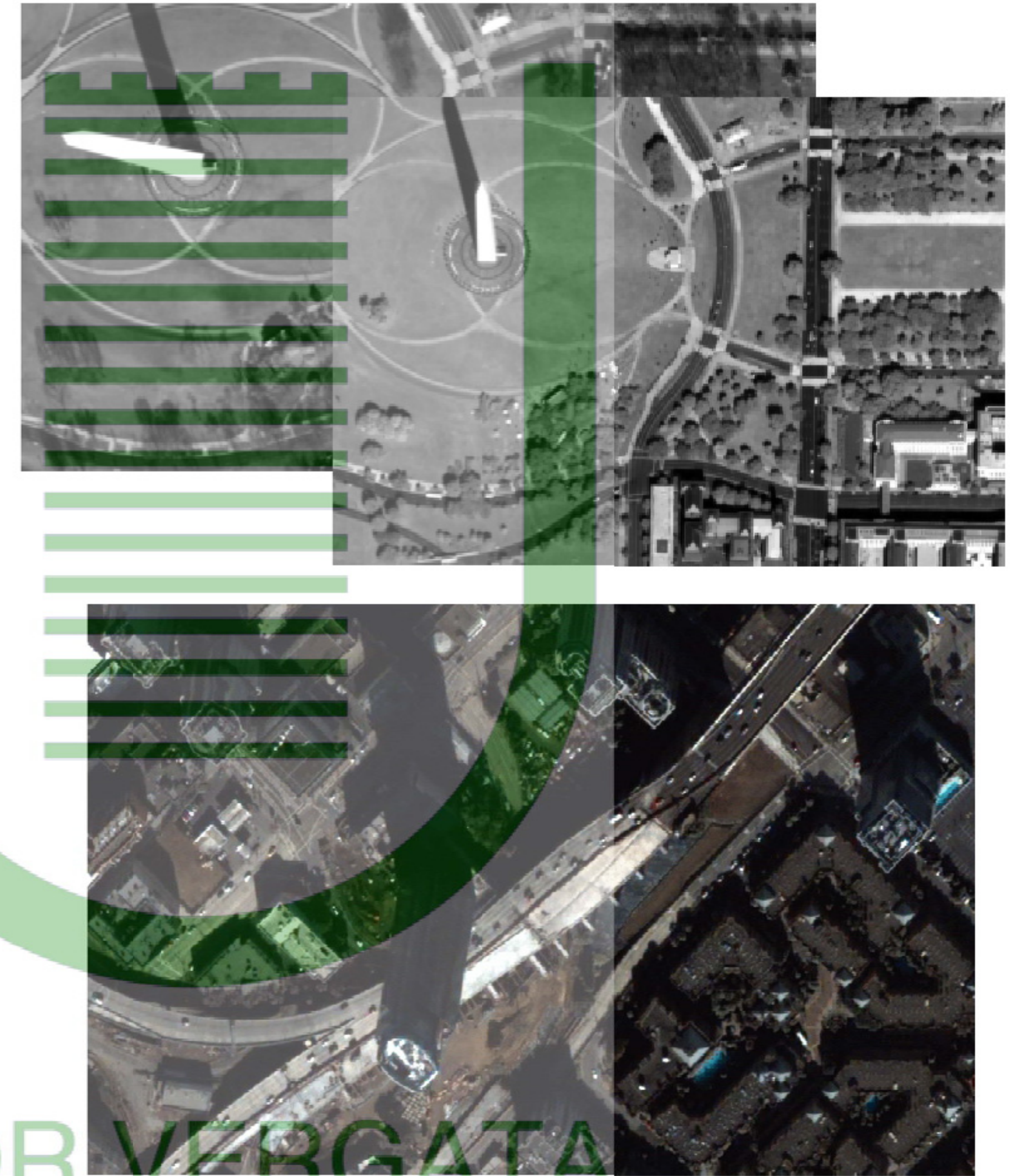
GeoInformazione

TOR VERGATA

Motivation (1/2)

The panchromatic nature of the new higher resolution WorldView1 satellite makes it attractive to consider colorizing the new imagery with other multispectral samples of the same target area.

It is not possible to overcome the mismatch between tall objects (such as towers or buildings) encountered in two different acquisitions using fast and automatic procedures.



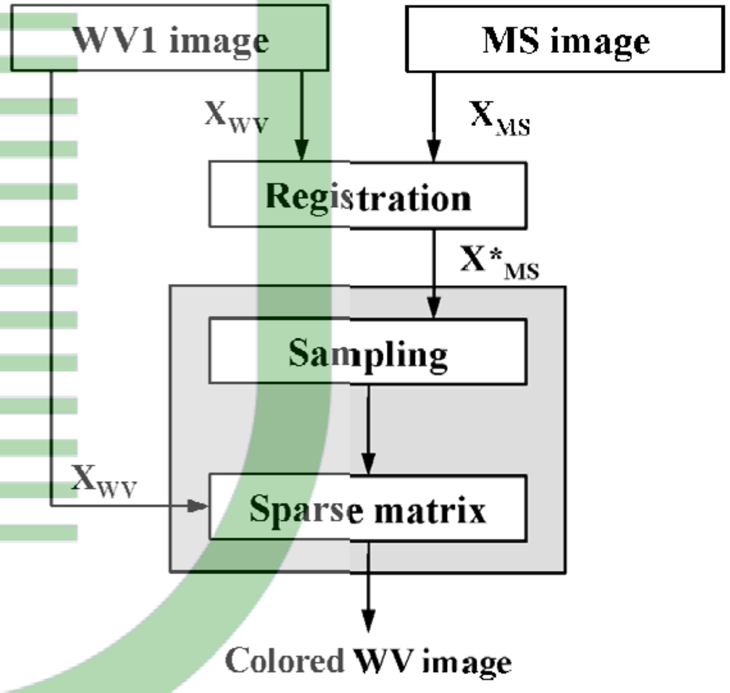
Motivation (1/2)

Moreover, an accurate image registration is not possible due to the likely differences in ground projections between sensors.

These aspects are particularly important for the WorldView 1 data colorization if we want to use other remotely sensed imagery, such as QuickBird, SPOT 5 or future WorldView 2 data, with different view angles, sun elevation and seasons.

The colorization method overcomes this limitation using a segmentation/sampling technique, which does not require accurate image registration.

Therefore, registration can be automatically done using standard software.



The algorithm (2/3)

The number of samples is a **critical choice** which has an effect on the time needed by the overall processing to solve the sparse matrix and on the colorization accuracy.

A **small number** of samples (for example 1% of the image) results in too few values used to solve the sparse system of linear equations. This results in two drawbacks:

1. the algorithm needs more time to solve the equations since few constraints are known
2. many objects in the scene are not sampled, therefore, these objects will not have the correct color information

The algorithm (3/3)

On the other hand, with a large number of samples (for example 70% of the image) the sparse system of linear equations is solved more rapidly since fewer variables have to be determined:

→ this results in a system that is more sensitive to view angle variations.

A trade off regarding the number of samples is necessary

Data Set

We tested the algorithm in urban areas since this is the most complex and challenging case, presenting a variety of housing, transportation systems, utilities, recreational areas, technical infrastructure and many temporary objects, such as cars, buses or daily markets. The test cities are San Francisco (USA), Beijing (China), Washington DC (USA) and Seville (Spain).

City	WorldView 1			QuickBird		
	Acquisition date	View angle ( )	Sun elev. ( )	Acquisition date	View angle ( )	Sun elev. ( )
San Francisco	November 26, 2007	19.6	29.6	November 11, 2007	26.6	34.4
Beijing	November 27, 2007	24.9	26.7	April 24, 2002	18.9	59.5
Washington	December 18, 2007	27.7	24.8	September 23, 2007	18.4	50.1
Seville	November 26, 2007	25.5	29.7	June 9, 2002	12.6	68.8

The San Francisco example

The city of San Francisco is mainly composed of large structures and tall buildings surrounded by many temporary objects such as cars and busses.

The height of buildings associated with different sun elevations during the acquisitions, creates canyons of shadow, which may result in difficulties in interpreting the scene.

San Francisco in False Colors (Bands 431)

Many DigitalGlobe costumers will be interested only in natural color scenes, but the proposed method is also able to produce false-color imagery.

We note the accuracy of the algorithm in colorizing the single trees along the highway in the right-hand part of the scene along with the shadows of these trees.

…another example

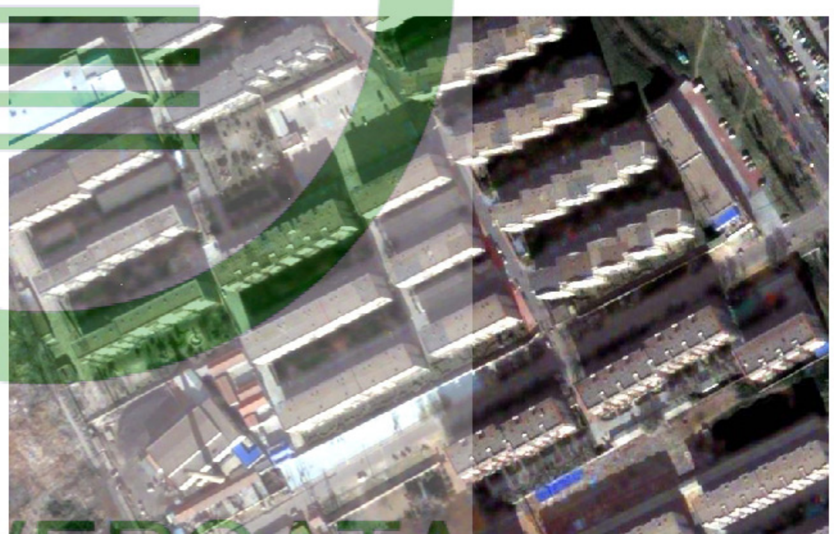
The city of Beijing, as many other cities in China, has shown a constant growth in terms of population. The resulting housing construction has resulted in the presence of many new apartment blocks having been built in the past few years.

WorldView-1 colorization

Comparison with the Gram-Schmidt algorithm

The contrast between buildings and areas of shadow is higher for the proposed method than the Gram-Schmidt algorithm, which produces an image dominated by brown due to the color of the surrounding buildings.

Compared with the previous landscape, this is more sensitive to the perspective mismatches due to the higher elevation of the buildings. Many unreal colors surround the buildings resulting in a fuzzyfication of the object`s boundaries.



Conclusions

The colorization method overcomes the limitation of the mismatch between tall objects using a segmentation/sampling technique, which does not require accurate image registration.

The number of samples is a critical choice which has an effect on the time needed by the overall processing to solve the sparse matrix and on the colorization accuracy.

The comparison with the Gram-Schmidt algorithm shown the higher robustness of the colorization method in producing accurate images using data acquired by different sensors, with different view angles, sun elevation and seasons.

Thank you for your attention!

Contacts:

Fabio Pacifici
f.pacifici@disp.uniroma2.it

Earth Observation Laboratory (EOLab)
Dottorandi GeoInformazione - room DT09
Dipartimento di Informatica Sistemi e Produzione - DISP
Via del Politecnico, 1 – 00133, Rome - ITALY
Office. +39 06 7259 7711