Use of Neural Networks for Automatic Classification From High-Resolution Images

Fabio Del Frate, *Member, IEEE*, Fabio Pacifici, *Student Member, IEEE*, Giovanni Schiavon, and Chiara Solimini

Abstract—The effectiveness of multilayer perceptron (MLP) networks as a tool for the classification of remotely sensed images has been already proven in past years. However, most of the studies consider images characterized by high spatial resolution (around 15–30 m) while a detailed analysis of the performance of this type of classifier on very high resolution images (around 1–2 m) such as those provided by the Quickbird satellite is still lacking. Moreover, the classification problem is normally understood as the classification of a single image while the capabilities of a single network of performing automatic classification and feature extraction over a collection of archived images has not been explored so far. In this paper, besides assessing the performance of MLP for the classification of very high resolution images, we investigate on the generalization capabilities of this type of algorithms with the purpose of using them as a tool for fully automatic classification of collections of satellite images, either at very high or at high-resolution. In particular, applications to urban area monitoring have been addressed.

Index Terms—Features extraction, high-resolution imagery, information mining, neural networks (NNs).

I. INTRODUCTION

NEURAL networks (NNs) started playing a significant role in the field of remote sensing after which a new learning algorithm was proposed. The idea of the backpropagation algorithm was originally developed by Werbos [1] and reintroduced by Rumelhart *et al.* [2]. Since the early nineties, several studies aimed at evaluating the performance of NNs by comparison with traditional statistical methods to remote sensing applications, and in particular to image classification. Benediktsson *et al.* [3] considered the two approaches in classification of multisource remote sensing data. They concluded that in multisource classification, where we do not always know the distribution functions, NNs can be more appropriate than statistical algorithms. Bishof *et al.* [4] as well as Paola and Schowengerdt [5] compared methods for multispectral (MS) classification of Landsat Thematic Mapper (TM) data, and both found that with proper training, a NN was able to perform better than the maximum-likelihood classification. However, even if these studies seem to show that NN performance is comparable or better than that provided by other techniques, they are mainly focused on high-resolution Landsat images and on the use of a single NN for classifying and/or extracting

specific features from a single image, namely the image from which the examples training the network are taken. Conversely, a detailed analysis of the pixel-based classification yielded by this type of algorithms on very high resolution images such as those provided by the Quickbird or Ikonos platforms is still lacking. Moreover, the potentialities of a single NN as a tool for automatic and sequential processing of images contained in archives have been scarcely investigated until now. With processing here, we mean that the network 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. In other words, the network allows the identification of high-level (object or region of interest) spatial features from the low-level (pixel) representation contained in a raw image or image sequence, hence addressing scientific issues characteristic of the image information-mining field [6], [7].

In this paper, as a first step, we want to assess and optimize the NN approach for the pixel-based classification of a single very high resolution image, such as one of those provided by the Quickbird satellite. Later, we move to the conceptually most innovative part of this paper which is to investigate on the capabilities of supervised NN in providing automatic classification on a collection of images, therefore their potentialities from an image information-mining point of view. This means to stress their generalization capabilities, that is the capabilities to obtain good generalization to new input patterns from the patterns on which the nets have been trained. Several issues interfere with the objective of designing NN able to generalize on images not used in the training phase. The robustness of the spectral information despite such problems has to be investigated and such an analysis needs to concur with the design of the NN. Addressing this point, in this paper, we consider both very high (Quickbird) and high (Landsat) resolution images and a specific application domain which is the feature extraction and information discovery on urban areas. In fact, monitoring changes and urban growth over time is one of the major areas of scientific research in remote sensing that have a strong interaction with the policy cycles and that would improve environment and security monitoring [8]–[10]. A large volume of satellite data for such purposes is available, but despite there being many competing automatic approaches, it is difficult to fully and automatically address the problems raised by the different application scenarios. In this paper, the aim of the classification is to distinguish among areas made of artificial coverage (sealed surfaces) including asphalt or buildings, and open spaces such as bare soil or vegetation.

II. NEURAL NETWORK ALGORITHM

NN models are mainly specified by the net topology and training rules [11]. The term topology refers to the structure of the network as a whole: the number of its input, output, and hidden units and how they are interconnected. Among various topologies, multilayer perceptrons (MLPs) have been found to have the best-suited topology for classification and inversion problems [12]. These are feedforward networks where the input flows only in one direction to the output, and each neuron of a layer is connected to all neurons of the successive layer but has no feedback to neurons in the previous layers. As far as the numbers of hidden layers and of their units are concerned, the topology providing the optimal performance should be selected. In fact, if the number of neurons is too small, the input-output associative capabilities of the net are too weak. On the other hand, this number should not be too large; otherwise, these capabilities might show a lack of generality being tailored too much on the training set, and the computational complexity of the algorithm would be increased in vain. It turns out that a fair compromise has to be found. The number of hidden layers is another issue to be considered. It has been shown that networks having two layers of weights, i.e., one hidden layer of neurons, and sigmoidal hidden units can approximate arbitrarily well any functional continuous mapping, provided the number of hidden units is sufficiently large [13], [14]. However, how much the inclusion of an additional hidden layer might improve the classification performance is still an open issue. In this paper, we followed a rather heuristic approach. We systematically analyzed the performance of the network varying either the number of hidden layers (one or two) or the number of hidden units and selecting the best topology on the base of the accuracy results obtained on a set of examples not considered for the training. The weight or strength of each connection has to be determined via learning rules to approximate an unknown input-output relation. These rules indicate how to pursue minimization of the error function measuring the quality of the network's approximation on the restricted domain covered by a training set (i.e., a set of input-output examples). A typical error function which can be considered in this context is the sum-of-squares error function (SSE) [13], given by a sum over all patterns, and over all outputs, of the form

$$\text{SSE} = \sum_{n=1}^{N} \sum_{k=1}^{c} \{y_k(\mathbf{x}^n; \mathbf{w}) - t_k^n\}^2 \tag{1}$$

where $y_k(\mathbf{x}^n; \mathbf{w})$ represents the output of unit k as a function of the input vector $\mathbf{x}^n$ and the weight vector $\mathbf{w}$, N is the number of testing patterns, and c is the number of outputs. The quantity t_k^n represents the target value for output unit k when the input vector is $\mathbf{x}^n$. In our case, the minimization of the error function has been pursued by a scaled conjugate gradient algorithm [15]. This is a member of the class of conjugate gradient methods, general-purpose second-order techniques that help to minimize goal functions of several variables. Second-order indicates that such methods use the second derivatives of the error function, while a first-order technique, like standard backpropagation, only uses the first derivatives. It should be mentioned that most of the neural

simulations were provided by the Stuttgart Neural Network Simulator package [16]. For the specific purpose of the image classification, a training set with a statistically significant number of pixels for each class has been generated. The learning of the NN has then been carried out by feeding it with pairs of vectors (patterns): the input vector contains the reflectances of the different channels of the MS image, and the output vector contains the corresponding known class of surface. To avoid saturation within the network it has been necessary to scale all the values of the input vector in the range between -1 and 1. The scaling has always been carried consistently on the entire dataset available. At the same time, the component of the output vector corresponding to the true class has been set to 1 while the others to 0. Once the NNs have been trained, they have been used for the classification of new data not considered in the training set. In the test phase, a competitive approach (winner-and-take) has been considered to decide on the final classification response.

III. SINGLE-IMAGE CLASSIFICATION

The QuickBird commercial remote sensing satellite provides images consisting of four MS channels with 2.4-m resolution and a single panchromatic band with 0.62-m resolution. The four MS bands collect data at the red, green, blue, and near-infrared wavelengths, and the data in each band are stored with 11-bit quantization. A QuickBird image taken over the Tor Vergata University campus, located in Italy, southeast of Rome, on March 13, 2003, has been initially considered. In the remainder of this paper, we will refer to this image with the name QB1. A view of the area is shown in Fig. 1. Besides the buildings in the campus, different residential areas belonging to the outskirts of the southeast side of the city can be distinguished in the image.

Our first purpose was to design an optimum NN able to classify the MS image. The considered land cover classes were buildings, roads, vegetated areas, and bare soil where the latter class includes non-asphalted road and artificial excavations. The inclusion of additional classes was discarded for several reasons: the considered classes are those that better describe the area under observation and are in themselves sufficient to detect significant features; the choice of a small number of classes enables an easier quantitative comparison of the performance obtained using a single net for a single image classification, with the one obtained using a single net for multiple images classification. In this latter case, we think that the choice of a number of four classes represents a rather ambitious target. It also has to be noted that a recent study analyzing satellite image classification experiments of 15 years pointed out that the idea postulating that the higher the number of classes used in a classification experiment, the more difficult the classification becomes, is not supported by the experimental results shown in this paper [17]. Once the classification problem has been configured, a first investigation consisted in analyzing the spectral behavior of the different considered surfaces. The selected pixels characterizing one class belong to polygons manually drawn in the image. It should be noted that, at the very high resolution of the images, the edges or boundaries between individual land cover objects were fairly sharp and it

Fig. 1.   Quickbird image of the Tor Vergata University Campus, Rome, and its surrounding (© DigitalGlobe, distributed by TELESPAZIO).
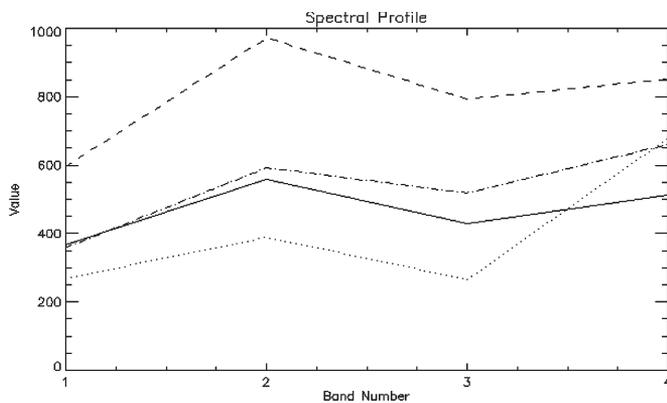

Fig. 2.   Spectral analysis from image QB1 for the classes buildings (dashed line), asphalted surface (solid line), bare soil (dash–dotted line), vegetation (dotted line).
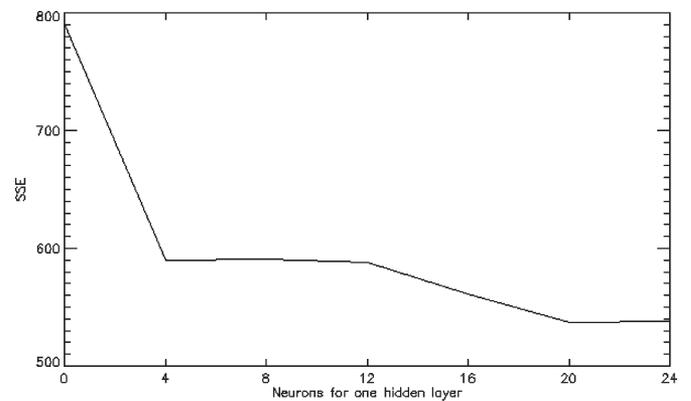

Fig. 3.   SSE values calculated over the test set changing the number of hidden neurons in a two-hidden-layers topology. The number of units is the same in both layers.

was usually easy to locate and assign a specific pixel to a land cover class. The mean values of the spectral signatures of the four categories are shown in Fig. 2. The figure clearly shows potentiality in discriminating between the classes. These stem from the spectral properties related to the different molecular resonance mechanisms which characterize the materials. With the same data considered for the sensitivity analysis we were able to generate a training set with a statistically significant number of pixels for each of the four categories. The training datasets were generated considering about 24 400 pixels. The design of the network was made putting particular care in the selection of the number of hidden units to be considered in the net. To this purpose, the plot illustrated in Fig. 3 was produced, where the SSE value over a test set of more than

3000 patterns and corresponding to different numbers of hidden units is reported. It can be seen that, if we consider both the SSE error and the network complexity, the best results were obtained with a 4-20-20-4 topology. Indeed, the increase of the number of hidden units did not change significantly the SSE error. A similar plot is reported in Fig. 4 where now a single hidden layer is considered. Again the best result are obtained putting around 20 neurons in the hidden layer; however, this topology is slightly worse if compared with the two-hidden-layers topology. This indicates that the second layer can be able to extract additional information from what already elaborated by the first one. The topology 4-20-20-4 was then finally selected and used to classify the entire image (3 506 832 pixels). Fig. 5 shows the classification map derived with the described
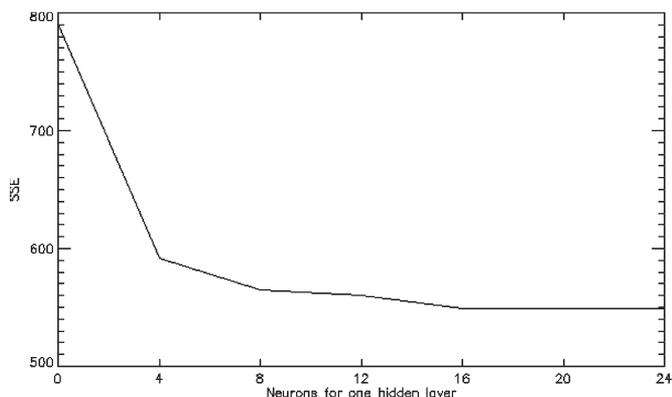
Fig. 4. SSE values calculated over the test set changing the number of hidden neurons in a one-hidden-layer topology.

Fig. 5. Classification map of the image QB1 using the optimized topology. Black: asphalted surfaces; white: buildings; dark gray: bare soil; light gray: vegetation.

TABLE I
CONFUSION MATRIX OBTAINED FOR IMAGE QB1 WITH THE 4-20-20-4 TOPOLOGY. OVERALL NUMBER OF PIXELS: 81 510. OVERALL ERROR 5998 (7.36%)

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

TABLE II
CHARACTERISTICS OF THE QUICKBIRD IMAGES USED IN THE WORK.
ALL THE ACQUISITION TIMES ARE AROUND 10:00–10:30 A.M.

Code	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 city
QB5	07/22/2005	2223x1450	12 Degrees	Nettuno town

procedure. The classification accuracy has been assessed by visual comparison with the original high-resolution image and by direct inspections on site. We stress the fact that our working area is located in the Tor Vergata University campus, which is almost in the center of image QB1, so direct inspection on site could be rather accurate. More in detail, a ground truth map, corresponding to a subset of the image, has been manually elaborated. We observed that the classification provided by the network is rather accurate and with a high level of resolution. In particular, we reached a 93% level of accuracy in the considered subimage. The whole confusion matrix is reported in Table I. Once the network topology for this kind of problem has been

optimized and the performance assessed, we move to investigate 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 choice of this new image should follow some similarity criteria with respect to the already classified one. For example, it would not be very meaningful to consider a new image characterized by land cover classes, such as water, not appearing in the QB1 image, hence not memorized at all by the network during its training process. The failure of the NN in this case can be given for granted, and this test would not provide much information in the evaluation of the network generalization capability. Therefore, we decided the other direction of selection and chose as a test image a QB image quite similar to QB1. Indeed, the new QB image (QB2) is taken on the same area of the first one, but in a different season and at a slightly different incident angle. In Table II, the basic information of the two images analyzed so far and of those that will be considered in the following of this paper are summarized. If the already trained network fails in generalizing over this image it will be very probably unsuccessful with many other QB images, even if taken on similar urban scenarios. In Fig. 8(a), we show the result of the classification of the QB2 image by using the net trained on patterns retrieved from image QB1. For the sake of completeness and for a better interpretation of the results, we also produced the classification, reported in Fig. 8(b), that would be obtained replying on the image QB2 the single-image classification methodology considered for the image QB1, therefore relying on a network (4-20-20-4), trained with examples belonging to the same image that one wants to classify. The classification map shown in Fig. 8(b) seems, as expected, rather accurate. Indeed, the misclassification percentage computed over the same image subset considered for QB1 is 95% thus resembling the one obtained in the former case. The classification result shown in Fig. 8(a) is completely different. Although the network recognizes many patterns and assigns the correct class to the corresponding pixels, entire objects are misclassified: the bare soil class and the built areas class are definitely overestimated, and the general noise level produced by the classification is significantly increased. From a quantitative point of view, the misclassification rate computed over the subset test image is 56%. Fig. 6 may contribute to understand the classification performance. We can observe that even if the shapes of the signatures resemble those plotted in
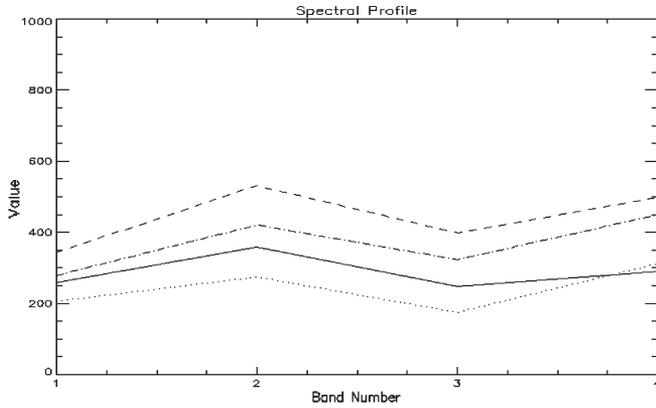
Fig. 6. Spectral analysis from image QB2 for the classes buildings (dashed line), asphalted surface (solid line), bare soil (dash–dotted line), vegetation (dotted line).
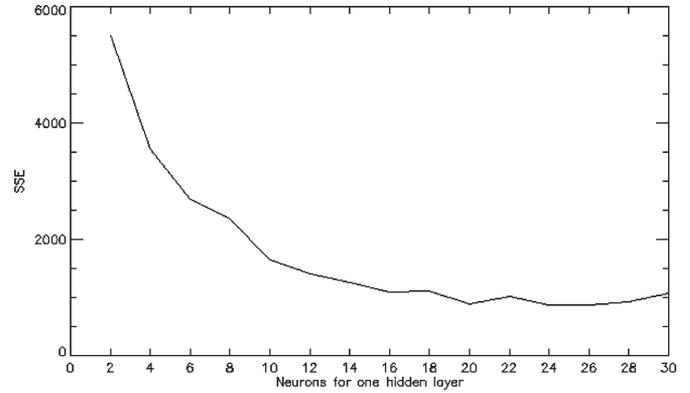

Fig. 7. SSE values calculated over the test set changing the number of hidden neurons in a two-hidden-layers topology designed for the classification of a collection of Quickbird images. The number of units is the same in both layers.

Fig. 2, which still enables some possibility of distinction among classes, the ranges of the digital number values are significantly different, generating confusion when the network gives out its classification response. Thus, the classification of the QB2 image obtained using a network trained on another image, even if taken on the same scenario, is not satisfactory. This means that to design a network able to provide good accuracy over images not considered in the training phase is an ambitious goal, even if the classification is performed on a limited number of classes.

IV. CLASSIFICATION ON A COLLECTION OF IMAGES

A. Quickbird Images

Three more Quickbird images have been considered in this case for an overall number of five images. As shown in Table II, the five images are of similar size but include different years, different sites, and different seasons. Besides the QB1 and QB2 images centered on the Tor Vergata University campus, we have one image (QB3) looking at northeast suburbs, a fourth image quite close to the old town (QB4), and a fifth image (QB5) which has been taken on a small urban area a few kilometers away from Rome. A pixel-based classification algorithm has again been implemented to distinguish among the four main classes: buildings, asphalted surfaces, vegetated areas, and bare soils. In previous section, we showed that a successful classification performance relies on a proper training and design of the network. In particular, it is important that the patterns included in the training set could significantly represent all potential scenarios that might be encountered during the application phase, in other words resemble the statistics of the classification problem. To this purpose, a larger archive of spectral signatures has been generated. Images QB1, QB3, and QB5 have been considered for the training, and about 26 000 examples have been collected for the generation of the network learning set.

The optimal performance both in terms of classification accuracy and of training time has been again determined by an extensive search whose results are illustrated in Fig. 7. With regard to the number of hidden layers, we relied on the previous result indicating a topology with two hidden layers was more effective, so the final selected topology was again 4-20-20-4.

Indeed, with respect to the single-image processing case, most of the physics characterizing the classification problem has not changed, which involved minor implications in terms of the topology to be selected. In Fig. 8(c) and (d), we report the classification maps obtained by applying the trained NN to the images QB2 and QB4 which did not contain any of the pixels included in the training set. From both visual inspection on the original images and direct inspection on site, we observed a general good agreement with the map generated automatically. All main features such as big roads and buildings are individuated with good precision even though some inaccuracies can be noted in the objects edge detection, possible causes of disturb being represented by shadow effects. A more quantitative analysis, computed on the same subarea of image QB2 considered in Section II, gave an overall accuracy rate of about 87%. Considering the encouraging results, and given the availability of two images (QB1 and QB2) over the same site, we tried to extend the described methodology to a typical change detection exercise. The two images have been coregistered using a set of about 30 ground control points and considering the older image as a master. We remind that the time interval between the two images is of one year. The two corresponding classification maps, obtained by means of the same network, have been used for the production of change detection maps. In particular, the change detection was evaluated in terms of the pixels that migrated from vegetation, bare soil or asphalted surface class to the building class in the considered time window. As in this case, we are more interested in an object-based result, the final change detection mask was obtained after a postprocessing which removed all clusters of pixels detecting changes but containing less than 20 elements. The ground-truth confirmed that the changes corresponding to the main detected structures were buildings constructed in the considered time interval. An example of detection result is shown in Fig. 9 where the previous corresponding classification maps are also reported. The corresponding confusion matrix, reported in Table III and computed on the base of the ground-truth, gives a high percentage of pixels in the diagonal. On the other hand, most of the pixels out of the diagonal, more than a real failure of the classification algorithm, may be a consequence of an imperfect coregistration of two images.
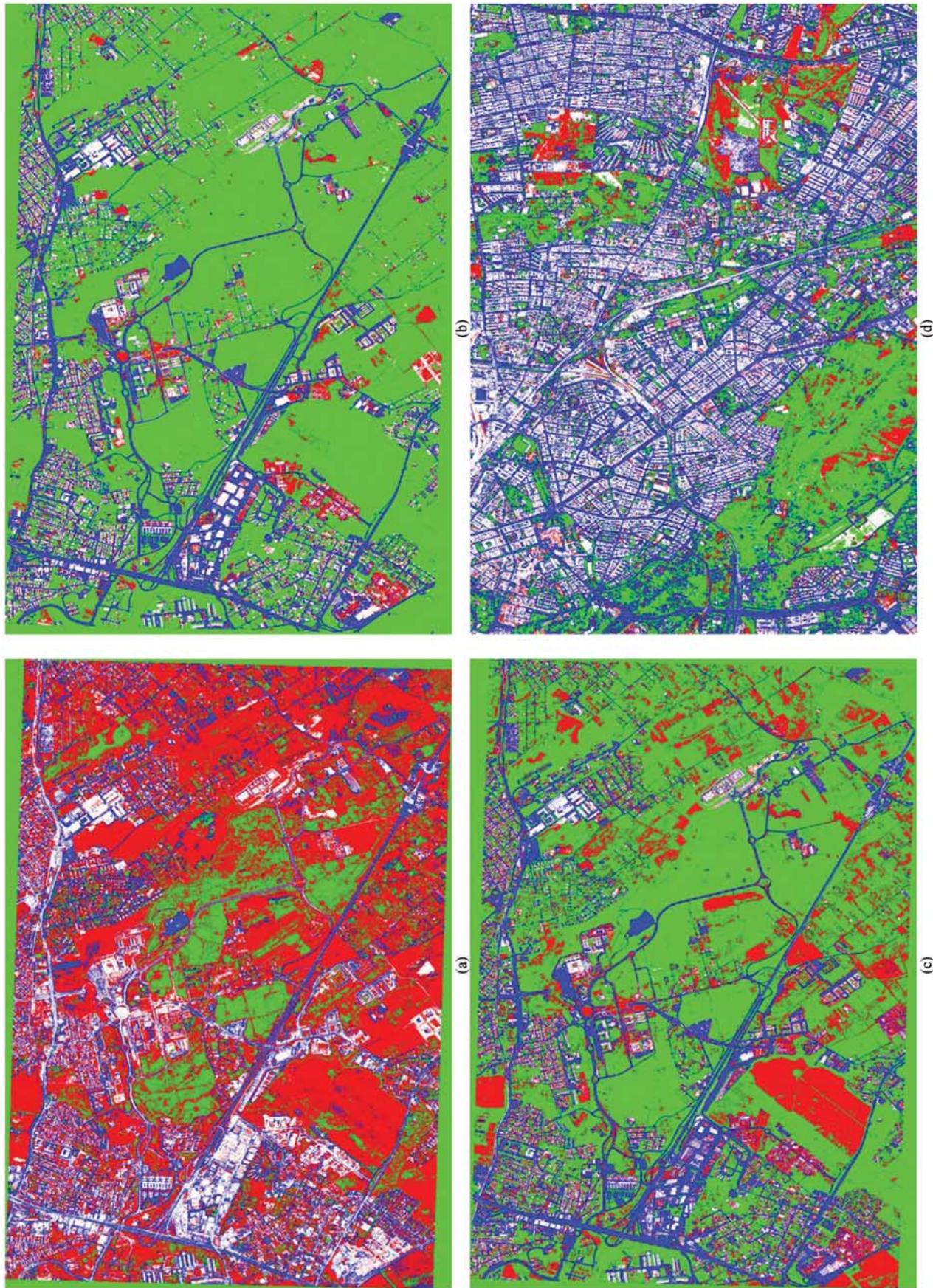
Fig. 8. Automatic classification map from: image QB2 with a net trained with examples taken from image QB1 (a), image QB2 with a net trained with examples taken from image QB2 (b), image QB2 (c), and image QB4 (d) with a net trained with examples taken from other images. Red: bare soil, blue: asphalted surface, white: buildings, green: vegetation.

Fig. 9.   Change detection results. (Left) 2002 classification map results. (Center) 2003 classification map results. (Right) Change detection.

TABLE III
CONFUSION MATRIX OBTAINED FOR THE CHANGE DETECTION EXERCISE.
OVERALL NUMBER OF CONSIDERED PIXELS: 148 538.
OVERALL ACCURACY 93.2%

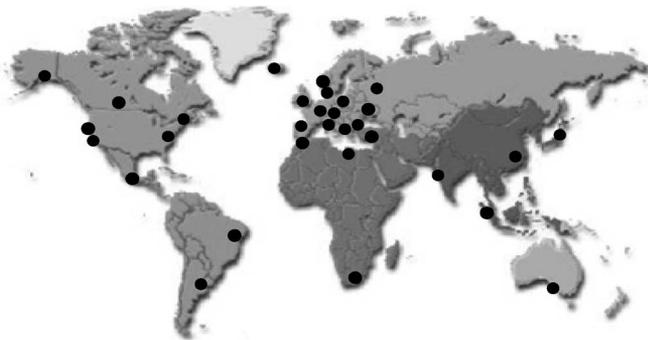| Classified as	True	
	Changed	*Unchanged*
Changed	14864	33
Unchanged	132	44785

Fig. 10.   Dataset geographic distribution of Landsat data.

B. Landsat Images

The objective of designing a single net enabling automatic classification of large archives of datasets has been extended to the case of Landsat imagery. The Landsat dataset consisted of a collection of images containing urban areas and located throughout the all five continents (Fig. 10). In this case, the inputs to the algorithm are taken from six bands, measurements corresponding to band 6 being discarded for its worse spatial resolution. Again, 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. For example in Figs. 11–13, we report the analysis for the classes high-density residential, forest, and water, respectively. For the three classes, the spectral behavior is calculated starting from an overall number of about 25 000 pixels, distributed over nine considered different geographic areas (see Table IV). We see that within the same class the shapes of the signature are in general rather similar, and only a bias value seems to characterize the different plots. On the other hand, different classes have quite dissimilar spectral shapes.

The analysis carried out on other classes typical of urban and suburban land cover confirmed the discrimination possibilities, especially if similar classes, such as forest and short vegetation areas, or high-density and low-density residential areas, were grouped together. In this case, the final classification problem was to discriminate among three classes: sealed, not sealed, and water. The sealed fraction of an urban areas is indeed one of the primary index for monitoring the urbanization process. However, many big cities are characterized by large amount of water surfaces, belonging to rivers, lakes, or sea. Therefore, the addition of the class water could be significant to obtain a better monitoring. More than 56 000 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, and 4 continents. A description of the training set in terms of the images and classes considered for each image is summarized in Table V. Given the variety of geographical sites taken into account, this classification problem is inherently more complex with respect to the classification of Quickbird data. In order to avoid overfitting, this shrinks the size of the optimum network topology which, for this case, has been found to be 6-9-9-3 (see Fig. 14). With this selected topology and considering medium-speed CPU computing platforms we obtained an average rate of processing of 800 pixels per second. This means less than 20 min for an image of 1000×1000 pixels, so basically we can speak of near-real-time processing. In Fig. 15, we show some examples of the results. The yielded accuracy seems to be rather satisfactory at careful visual inspection. Water bodies are detected rather precisely as the major parts of the urban lattice. On the other hand, we noted some inaccuracies on areas which appear as low residential areas at image visual inspection but are labeled as unsealed areas in the classification map. Similar results have been obtained selecting other images from the available dataset. In any case, a more quantitative validation exercise could be performed on a limited area of the city of Rome, where we could use the Quickbird very high resolution image as ground-truth. The area chosen for the validation exercise does not contain pixels used for the training of the final Landsat network. In Table VI, we report the obtained corresponding confusion matrix where we did not include pixels, such as pixels on edges, whose real class could be not stated with certainty. Given the totally automatic procedure, the overall accuracy of about 82% might be recognized as encouraging and establishes a benchmark for this kind of application.
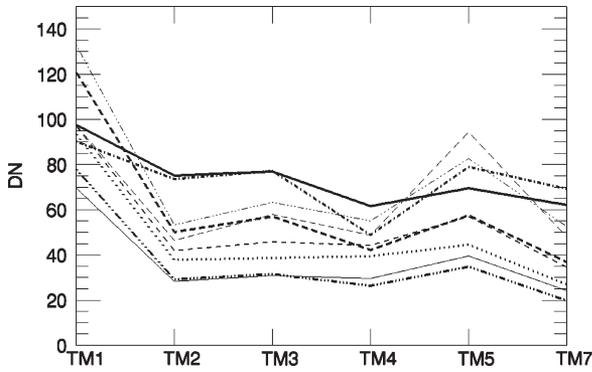
Fig. 11. Spectral analysis from Landsat measurements of the class high-density residential for different cities in the world.
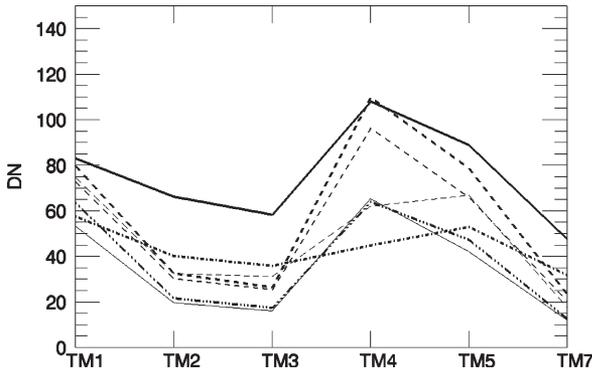

Fig. 12. Spectral analysis from Landsat measurements of the class forest for different cities in the world.
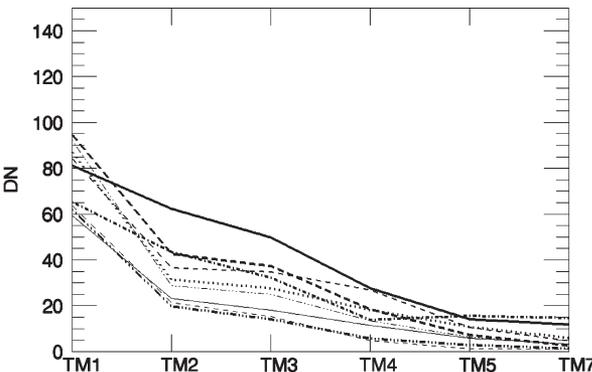

Fig. 13. Spectral analysis from Landsat measurements of the class water for different cities in the world.

TABLE IV
LEGEND OF FIGS. 11–13

TABLE V
LOCATION AND DATES OF THE LANDSAT IMAGES USED
FOR THE GENERATION OF THE TRAINING SET

CITY	Sensor	Acquisition date	Classes
AMSTERDAM	TM	22/05/1992	all
BARCELONA	ETM+	10/08/2000	unsealed
BERLIN	ETM+	14/08/2000	unsealed
BUDAPEST	ETM+	09/08/1999	all
LONDON	TM	20/05/1992	all
MELBOURNE	ETM+	05/10/2000	all
NEW YORK	TM	28/09/1989	all
PARIS	TM	09/05/1987	all
RIO DE JANEIRO	TM	18/01/1988	all
ROME	ETM+	03/08/2001	all
TOKIO	TM	21/05/1987	all
WIEN	TM	10/09/1991	all
ROME_2	ETM+	16/01/2001	all
UDINE	ETM+	16/08/2000	sealed, water

The rightmost column indicates which classes have been considered for the specific image.
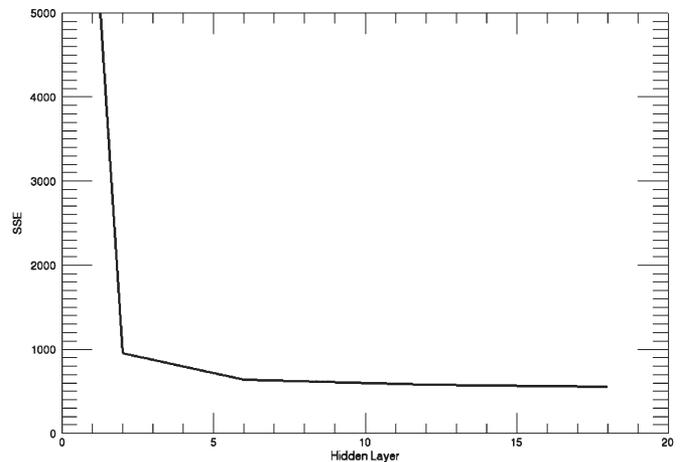

Fig. 14. SSE values calculated over the test set changing the number of hidden neurons in a two-hidden-layers topology for the classification of a collection of Landsat images. The number of units is the same in both layers.

V. CONCLUSION

It is well recognized that one of the major advantages of NNs with respect to Bayesian and other statistically based classifiers is that NNs draw their own input-output discriminant relations directly from the data and do not require that a particular form of a probability density function be assumed [18]. In this paper, we exploited these characteristics of MLP networks for automatic processing of large datasets of satellite imagery and with particular interest for features extraction from urban areas. In fact, this paper can be considered as a first step in demonstrating how NNs can contribute to the development of image information mining (IIM) in Earth observation. We considered two types of satellite data: Quickbird data characterized by very high spatial resolution and the Landsat data characterized by

high spatial resolution. In both cases, the purpose was both to yield accurate classification maps and to train the networks in order to generalize out of the image dataset considered in the training phase so that the new images could be processed in near-real time. To that purpose careful spectral analysis over statistically significant datasets have been carried out, and the NN topologies have been designed avoiding possible effects of
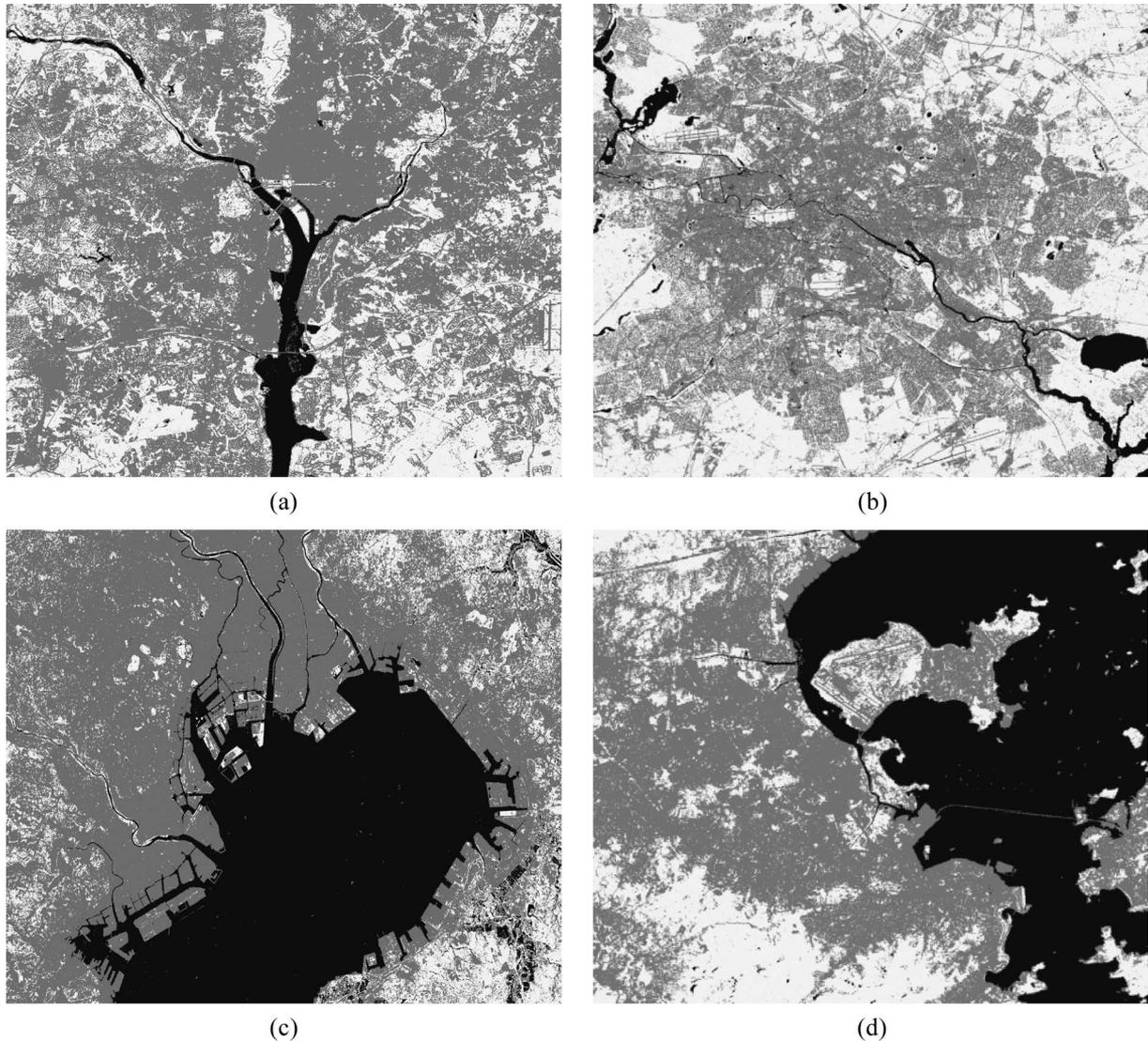
Fig. 15. Automatic classification map of the city of (a) Washington, DC (U.S.), (b) Berlin (Germany), (c) Tokyo (Japan), and (d) Rio de Janeiro (Brazil). Black: water surface, gray: sealed surface, white: unsealed surface (open space).

TABLE VI
CONFUSION MATRIX OBTAINED FOR LANDSAT NEURAL ALGORITHM: THE OVERALL ACCURACY IS OF ABOUT 82%

Classified as	True		
	Sealed	*Unsealed*	*Water*
Sealed	332	54	1
Unsealed	52	172	2
Water	2	1	15

overfitting. 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, which are 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. Even though, both for high and very high spatial resolutions, the experiments have been carried out on similar scenarios, the overall accuracies of 87% and 82% obtained for selected Quickbird and Landsat subareas, respectively, represent a benchmark for successive studies. Finally, if images of the same area are available at different times, the described technology seems also to be useful for an automatic discoverage of changes, such as new buildings, that occurred in the area under observation.

REFERENCES

[1] P. J. Werbos, "Beyond regression: New tools for prediction and analysis in the behavioural sciences," Ph.D. dissertation, Harvard Univ., Cambridge, MA, 1974.
[2] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA: MIT Press, 1986, pp. 318–362.

[3] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, "Neural network approaches versus statistical methods in classification of multi-source remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 4, pp. 540–552, Jul. 1990.

[4] H. Bishof, W. Schneider, and A. J. Pinz, "Multispectral classification of Landsat-images using neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 3, pp. 482–490, May 1992.

[5] J. D. Paola and R. A. Schowengerdt, "A detailed comparison of back propagation neural network and maximum likelihood classifiers for urban land use classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 981–996, Jul. 1995.

[6] M. Datcu *et al.*, "Information mining in remote sensing image archives: System concepts," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2923–2936, Dec. 2003.

[7] W. Hsu, M. L. Lee, and J. Zhang, "Image mining: Trends and developments," *J. Intell. Inf. Syst.*, vol. 19, no. 1, pp. 7–23, Jul. 2002.

[8] J. R. Jensen and D. C. Cowen, "Remote sensing of urban/suburban infrastructure and socio-economic attributes," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 5, pp. 611–622, 1999.

[9] J. P. Donnay, M. J. Barnsley, and P. A. Longley, "Remote sensing and urban analysis," in *Remote Sensing and Urban Analysis*, J. P. Donnay, M. J. Barnsley, and P. A. Longley, Eds.   London, U.K.: Taylor & Francis, 2001, pp. 3–18.

[10] T. Carlson, "Applications of remote sensing to urban problems," *Remote Sens. Environ.*, vol. 86, no. 3, pp. 273–274, 2003.

[11] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE Acoust. Speech Signal Process. Mag.*, vol. 4, no. 2, pp. 4–22, Apr. 1987.

[12] S.-Y. Hsu, T. Masters, M. Olson, M. Tenorio, and T. Grogan, "Comparative analysis of five neural network models," *Remote Sens. Rev.*, vol. 6, no. 1, pp. 319–329, 1992.

[13] C. Bishop, *Neural Networks for Pattern Recognition*.   New York: Oxford Univ. Press, 1995.

[14] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.

[15] M. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Netw.*, vol. 6, no. 4, pp. 525–533, 1993.

[16] A. Zell *et al.*, "SNNS Stuttgart neural network simulator user manual," Univ. Stuttgart, Inst. Parallel and Distrib. High Perform. Syst., Stuttgart, Germany, Rep. N.6/95, 1995. [Online]. Available: www.ra.informatik.uni-tuebingen.de/SNNS

[17] G. G. Wilkinson, "Results and implications of a study of fifteen years of satellite image classification experiments," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 433–440, Mar. 2005.

[18] M. S. Dawson, "Applications of electromagnetic scattering models to parameter retrieval and classification," in *Microwave Scattering and Emission Models and Their Application*.   Norwood, MA: Artech House, 1994.

Fabio Pacifici (S'02) was born in Rome, Italy, in 1980. He received the Laurea (B.S.) and Laurea Specialistica (M.S.) degrees in telecommunication engineering (both *summa cum laude*) from Tor Vergata University, Rome, in 2003 and 2006, respectively. He is currently working toward the Ph.D. degree in geoinformation at the Earth Observation Laboratory (EOLab), Tor Vergata University and collaborates with the Aerospace Department of the University of Colorado at Boulder.

His main research activity is in the area of remote sensing image processing. In particular, his interests are related to the classification and change detection task of urban areas by using statistical approaches such as neural networks and support vector machines in very high spatial resolution optical imagery and SAR imagery. He is currently involved in remote sensing projects supported by the ESA. In 2005, he was a Visitor Student at the Colorado Center for Astrodynamics Research with the University of Colorado.

Mr. Pacifici is a member of the Italian Association for Remote Sensing (AIT) and a Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Giovanni Schiavon received the Laurea degree (*cum laude*) in electronic engineering from the University of Rome "La Sapienza," Rome, Italy, in 1982.

He is currently an Associate Professor with the University of Rome "Tor Vergata," Rome, where he has been a Researcher since 1984 to 2000. He has been teaching a course on remote sensing since 1996 and a course on electromagnetic fields since 2000. His research activity has mainly been concerned with remote sensing of the atmosphere and of the Earth's surface, propagation, and millimeter waves. He has managed several work packages of different contracts with the ESA. He has been involved in the international remote sensing projects AGRISAR (1986), AGRISCATT (1987, 1988), MAESTRO-1 (1989), MAC Europe (1991), SIR-C/X-SAR (1994), and in the ERA–ORA program (European Commission Fourth Framework Programme) on Earth observation for environmental monitoring. He has authored or coauthored more than 150 scientific papers, most of which are in international journals or proceedings. He has acted as Reviewer for the international journals *Radio Science*, *Natural Hazards*, and for several volumes of VSP publisher.

Dr. Schiavon is a Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Fabio Del Frate (M'03) received the Laurea degree in electronic engineering and the Ph.D. degree in computer science from Tor Vergata University, Rome, Italy, in 1992 and 1997, respectively.

From September 1995 to June 1996, he was a Visiting Scientist with the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, In 1998 and 1999, he was with the European Space Agency (ESA), ESA Centre for Earth Observation, Frascati, Italy, as a Research Fellow and was engaged in projects concerning end-to-end remote sensing applications. He is currently a Research Professor with Tor Vergata University, where he teaches courses on electromagnetics and neural networks (NN). He has acted and is Principal Investigator in several remote sensing projects supported by ESA. He his author or coauthor of more then 100 scientific publications, with a special focus on the applications of NNs to remote sensing inversion problems. His main research topics also include retrieval and classification algorithms for land cover from satellite data, oil spill detection in SAR imagery, retrieval of atmospheric variables with microwave radiometry, data-exploitation for the new missions Project for On-Board Autonomy and the Ozone Monitoring Instrument.

Dr. Del Frate serves as a Reviewer for different remote sensing journals and as Associate Editor for IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Chiara Solimini received the Laurea (M.S.) degree in environmental engineering from Tor Vergata University, Rome, Italy, in 2002, where she is currently working toward the Ph.D. degree in geoinformation, focusing on classification and change detection of urban areas using neural networks, transform algorithms, and high-resolution QuickBird images, at the EOLab in the Department of Information, Systems and Productions, Tor Vergata University.

In January 2005, she started a collaboration with the Aerospace Department of the University of Colorado at Boulder and with Company DigitalGlobe, aimed at high-resolution urban monitoring.