A Comparative Study of NN and SVM-Based Electromagnetic Inverse Scattering Approaches to On-Line Detection of Buried Objects

Salvatore Caorsi[1], Davide Anguita[2], Emanuela Bermani[3], Andrea Boni[3],
Massimo Donelli[3] and Andrea Massa[3]

[1] Dept. of Electronics, University of Pavia, Via Ferrata 1, I-27100 Pavia,Italy
[2] Dept. of Biophysical and Electronic Eng., University of Genoa, Via Opera Pia 11A, I-16145, Genova,Italy
[3] Dept. of Information and Communication Tech.,University of Trento, Via Sommarive 14, I-38050 Trento,Italy

Abstract—**Microwave-based sensing techniques constitute an important tool for the detection of buried targets. In this framework, a key issue is represented by real-time scatterer localization. As far as such a topic is concerned, this paper presents a comparative evaluation of the performances provided by a conventional NN-based inverse scattering technique and by a new SVM-based electromagnetic approach. In order to estimate the effectiveness values of the two methods, realistic configurations and noisy environments are considered and current capabilities, as well as potential limitations, are pointed out. Finally, possible future research work is outlined.**

I. INTRODUCTION

The detection of buried objects by means of interrogating electromagnetic waves is usually a very difficult task. The addressed problem is nonlinear, due to the relation between unknown quantities (object parameters and field distribution) and problem data, it is ill-posed and, generally, only aspect-limited measures are available.

In the past few years, considerable efforts have been devoted to dealing with detection or, more generally, reconstruction problems, and several approaches have been proposed. As far as weak scatterers are concerned, linearized procedures have been applied (see [1], [2], [3] and references cited therein). The use of closed forms of the scattering equations (based on the diffraction theorem) and of the Fast Fourier Transform (FFT) has made it possible to obtain faster processing rates and real-time imaging. Moreover, numerical procedures based on higher-order Born approximations have demonstrated their effectiveness [4], [5].

On the contrary, nonlinear algorithms must be used when strong scatterers are considered. The retrieval problem is usually recast into an optimization one and is effectively solved with minimization techniques [6]-[10]. Unfortunately, large computational resources and a high computational load are needed, thus making these techniques impracticable (particularly when serial implementations are realized) if real-time performances are required.

In order to speed up the detection process, a key point is the reduction in the number of unknowns. Toward this end, *a-priori* information (if available) on the scenario under test can be very useful. This concept has been successfully exploited in inverse methodologies based on artificial neural networks (NNs) (see [11] (pp. 475-479) and references cited therein). As far as detection problems are concerned, methods based on both multilayered-perceptron [12], [13] and radial-basis-function [14] neural networks have demonstrated their capabilities for on-line retrieving of buried cylindrical scatterers.

Though NN-based approaches have generally offered good performances in solving inverse-scattering problems, they still suffer from several drawbacks not completely solved up to now. From the inductive-theory point of view, the main drawback is the difficult control of the complexity of underlying NN models. By the term *complexity* it is usually meant the capacity of a learning machine to fit the input data. Briefly, if a machine is too complex, it will typically overfit the data, thus losing the property of generalization for new measures not included in the training set. If complexity is too low, the machine will fail to correctly interpret the underlying relations among training samples. The complexity of a learning machine depends on many factors. In the case of NNs, the numbers of hidden layers and neurons, the number of interconnections, and the learning algorithm used for the training process [15] are the predominant parameters. Unfortunately, NNs lack an effective theory suggesting the most suitable NN topologies and/or calibration parameters. An NN adapts its internal parameters (i.e., the weights) automatically in order to best approximate the available training data, but the topology, the transfer function and the other parameters are heuristically selected. At present, there is no good way to determine how many hidden layers or how many hidden nodes each layer

should contain, given the sample data with which to train the NN for the solution of a given problem. From the computational and optimization points of view, NNs exhibit drawbacks as well. The learning process of an NN consists in solving a nonlinear optimization process. Therefore, any conventional optimization algorithm, including the widely used back-propagation procedure, leads to reach a solution that corresponds to one of the local minima of the target cost function. An empirical way to face such a problem is to train several NNs with different starting points, thus overloading the optimization process. However, this choice might result in inability to unambiguously evaluate statistical and systematic errors on neural computations.

A possibility to overcome these drawbacks is based on new results in Statistical Learning Theory (SLT) [16] which lead to new algorithmic paradigms and new computational architectures that, though still based on the NN model, entirely relinquish their biological plausibility to achieve a firm theoretical background. SLT allows one to derive statistical and algorithmic properties that can limit or avoid altogether the NN problems previously described. One of the main contributions in this field has been provided by Vladimir Vapnik [16], who has formulated and formalized the inductive rules that regulate the learning process by example principles. On the basis of these fundamentals, a new learning paradigm, called Support Vector Machine (SVM), has been developed. After initial studies, SVMs are now successfully applied in several fields ranging from pattern recognition to function approximation tasks. From a theoretical point of view, SVMs turn out to be very appealing, as compared with conventional NNs, for the following two basic reasons:

- the constrained-quadratic structure of the optimization problem solved for the learning process;
- the solid statistical theory on which SVMs are based.

In this paper, the detection of buried objects by means of interrogating electromagnetic waves is addressed by using an inductive approach. Within the framework of electromagnetic retrieval, the effectiveness and limitations of the SVM-based strategy are evaluated and a comparative study versus conventional NN-based methods is made. Finally, selected numerical results on realistic configurations and noisy environments are reported and discussed.

II. Mathematical Formulation

Let us consider the problem of determining the unknown parameters of an object buried in a homogeneous soil. With reference to a two-dimensional geometry, let D_S be a lossy region with complex contrast, $\tau_S = [\varepsilon_S - 1] - j\frac{\sigma_S}{2\pi f \varepsilon_0}$, enclosing a circular cylindrical scatterer of radius ρ_B. The dielectric properties of the
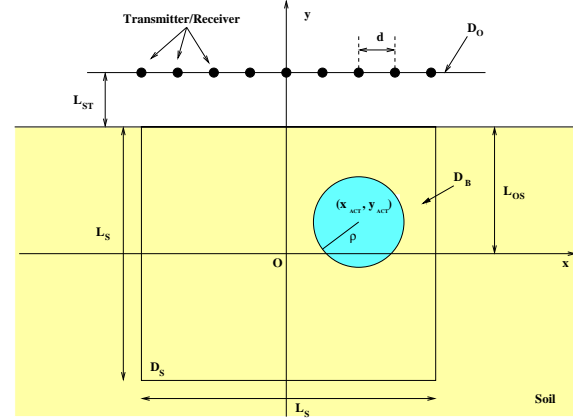

Fig. 1. Geometry of the problem

object are homogeneous, τ_B, and the dielectric profile of the geometry under test (Fig. 1) can be described as follows:

$$\tau(x,y) = \begin{cases} \tau_0 & if\ y > L_{0S} \\ \tau_B & if\ \begin{cases} 0 \le x \le x_B + \rho_B cos\theta \\ 0 \le y \le y_B + \rho_B sin\theta \\ 0 \le \theta < 2\pi \end{cases} \\ \tau_S & otherwise \end{cases}$$

(1)

This scenario is illuminated by multiple transmitters lying on D_0 and located, in the upper half space, at the positions $\{(x_t, y_t);\ t = 1, ..., T\}$. The probing fields, $E^{irr}(x,y)$, are radiated in the free space and at a fixed frequency by a known distribution of current filaments parallel to the z-axis. The same probes work as receivers for the anomalous field.

Under these hypotheses, the addressed inverse scattering problem can be mathematically stated as follows. Starting from the knowledge of the anomalous field, E^{tot}, collected at the receiver positions $\{(x_r, y_r);\ r = 1, ..., R\}$

$$E^{tot}(x_r, y_r|x_t, y_t) = E^{inc}(x_r, y_r|x_t, y_t) + \\ +k^2 \int_{D_S} E_S(x, y|x_t, y_t) G_S(x_r, y_r; x, y)\tau(x,y)\,dxdy$$

(2)

determine the set of unknown parameters $\{(x_B, y_B), \rho_B, \tau_B\}$ defining the scatterer under test. In eq. (2):

- $E^{inc}(x_r, y_r|x_t, y_t) = E^{irr}(x_r, y_r|x_t, y_t) + E^{ref}(x_r, y_r|x_t, y_t)$ is the electric field at the receivers in the absence of the object;
- $E^{ref}(x_r, y_r|x_t, y_t)$ is the electric field reflected by the planar interface at the receivers;
- $E_S(x, y|x_t, y_t)$ is the electric field induced inside the reconstruction domain D_S when it contains the object;
- $G_S(x_r, y_r; x, y)$ is the Sommerfeld-Green function for the half-space geometry [6].

Then the solution of the addressed inverse scattering problem requires the determination of the nonlinear function, Φ, defined as follows

$$\underline{\chi} = \Phi\left\{\underline{E}^{tot}\right\} \tag{3}$$

where $\underline{\chi}$ is the "scatterer array" ($\underline{\chi} = [\chi_p; p = 1, ..., P]$ $= [(x_B, y_B), \rho_B, \tau_B]$, P being the number of unknown parameters) and $\underline{E}^{tot}$ is the data array defined as $\underline{E}^{tot} = [E^{tot}(x_r, y_r|x_t, y_t);$ $r = 1, ..., R; \ t = 1, ..., T]$. This is a *regression problem* in which the unknown function (Φ) must be approximated by the knowledge of a number of known input-output pairs of vectors $\left\{(\underline{\chi})_n, (\underline{E}^{tot})_n\right\}; \ n = 1, ..., N$.

III. LEARNING-BY-EXAMPLES STRATEGIES FOR INVERSE SCATTERING PROBLEMS

The inverse-scattering problem described in Section II can be addressed in several ways. From a mathematical point of view, the key issue is to find an approximation, $\hat{\Phi}$, for the unknown function Φ on the basis of a set of samples $\{(\underline{\nu}_n, \underline{e}_n); \ n = 1, ..., N\}$, $\underline{\nu}_n$ and $\underline{e}_n$ being an input pattern (i.e., a data array $\underline{\nu}_n \equiv \underline{E}^{tot}$) and the corresponding target (i.e., a scatterer array $\underline{e}_n \equiv (\underline{\chi})_n$), respectively. This is a typical *learning-by-examples* problem, which is usually faced in the presence of unknown systems with measurable input/output signals. In the following, two approaches based on a multilayer perceptron (MLP) neural network and on an SVM, respectively, will be presented.

A. MLP-NN Approach

Neural networks are distributed computational systems characterized by a multi-layered structure of neurons fully interconnected by weighted links. MLP-NNs can be considered as universal approximators for any function $\Phi : \Re^{R \times T} \to \Re^P$ [17]. Therefore, they are suitable for facing with regression problems characterized by complex nonlinear relations between data and unknowns, such as inverse scattering or buried-object detection problems. In this context, $\Re^{R \times T}$ is the space of arrays representing measurement data, and $\Re^P$ is the space of unknown parameters describing a buried object.

MLP-NN theory [11] suggests approximating Φ by a nonlinear function of the weighted measurement data

$$\hat{\Phi}(\underline{\nu}) = \Psi\left\{\underline{\xi}_{L-1} \underline{\underline{w}}^{(L-1, L)} + \underline{b}_L\right\} \tag{4}$$

where L is the number of layers; $\underline{\xi}_l = \Psi\left\{\underline{\xi}_{l-1}\underline{\underline{w}}^{(l-1, l)} + \underline{b}_l\right\}$, $l = 1, ..., L$ being $\underline{\xi}_0 = \underline{\nu}$; Ψ is the nonlinear activation function (e.g., a sigmoid); $\underline{\underline{w}}^{(l-1, l)}$ and $\underline{b}_l$ are the weight matrix and the bias array of the l-th layer, respectively. Given known input-output pairs of vectors (called *training set*),

$\Gamma_{training} = \{(\underline{\nu}_n, \underline{e}_n); \ n = 1, ..., N\}$, and according to a backpropagation algorithm, weights and biases are computed by minimizing the error function ψ

$$\psi\left\{\underline{\underline{w}}^{(l-1, l)}, \underline{b}_l; \ l = 1, ..., L\right\} = \sum_{n=1}^{N}\left\|\underline{e}_n - \hat{\Phi}(\underline{\nu}_n)\right\| \tag{5}$$

by a gradient descent procedure.

Therefore, the direct solution of the inverse-scattering problem is avoided, and real-time (after the training phase) solutions to buried-object localization are obtained [13]. However, as the error function (5) is non-convex, one of the main drawbacks of the approach is the presence of local minima where the optimization algorithm might be trapped and the solution of which would have no physical significance.

B. SVM-Based Approach

In order to avoid the drawbacks of the NN-based inverse scattering method related to the nonlinear fitting of the training samples, an SVM-based approach is presented. The underlying idea of the SVM procedure is to split the approximation for the nonlinear function Φ into two steps. Instead of performing a nonlinear fitting in the input space, a nonlinear mapping of the training samples from the input space into a larger (possibly infinite) space (i.e, the *feature space*, $\Re^{\tilde{T}}$) is first performed. Then, a simple linear fitting is carried out in the new space, thus avoiding typical nonlinear-fitting drawbacks and keeping the advantages of a linear approach. Moreover, by exploiting some mathematical properties of nonlinear mappings, the evaluation of the data in the feature space is not required, as the SVM does not have to explicitly work in this space.

In more detail, each data array $\underline{\nu}_n$ is mapped into the feature space through a nonlinear transformation $\underline{\varphi} : \Re^{R \times T} \to \Re^{\tilde{T}}$ with $\tilde{T} \gg R \times T$. Then, the samples in the feature space are linearly interpolated by defining a hyperplane with a normal vector $\underline{w}$. Thus, the approximating function is given by

$$\hat{\Phi}(\underline{\nu}) = \underline{w} \cdot \underline{\varphi}(\underline{\nu}) + b \tag{6}$$

Among all possible hyperplanes, SVMs find the one that corresponds to a function $\hat{\Phi}$ having at most a deviation ϵ from each target $e_n^{(p)}$ (*), for all the measures $\underline{\nu}_n$, and that, at the same time, is as "flat" as possible. As it is impossible for all the points to lie inside the ϵ band, some errors (ξ_n, ξ_n^*, also called *slack variables*) are allowed and linearly weighted. Mathematically, this description leads to a constrained

---

(*) As up to now it has been possible to synthesize only single-output SVM, we refer to the estimation of a single scatterer array component $e_n^{(p)} = (\chi_p)_n; \ p = 1, ..., P$.

quadratic optimization problem (CQP) where the regularized cost function γ

$$\gamma \left\{ \underline{w}, b \right\} = \left\{ \frac{1}{2} \left\| \underline{w} \right\|^2 + C \sum_{n=1}^{N} \left(\xi_n + \xi_n^* \right) \right\} \quad (7)$$

is minimized over $\underline{w}$ and b, subject to the following constraints:

$$\begin{cases} e_n^{(p)} - \underline{w} \cdot \varphi \left(\underline{v} \right) - b \leq \epsilon + \xi_n \\ \underline{w} \cdot \varphi \left(\underline{v} \right) + b - e_n^{(p)} \leq \epsilon + \xi_n^* \quad \forall n \\ \xi_n, \xi_n^* \geq 0 \end{cases} \quad (8)$$

The function γ is composed of two terms. The first forces the hyperplane to be as flat as possible, and the second penalizes the deviation of each target from the function $\hat{\Phi}$. The constant C measures the tradeoff between the two terms. It can be shown that this approach can be used to control the complexity of the learning machine, according to the Structural Risk Minimization principle [16]. This principle guarantees a considerable generalization ability of the model, and provides upper bounds to such ability, albeit in a statistical framework. It is also interesting to note that this formulation, which derives from SLT, resembles closely the regularization approach that is usually exploited when dealing with ill-posed problems, like inverse ones [18].

The problem defined by eqs. (7)-(8) is then rewritten in dual form by using the Lagrange multiplier theory. By introducing $2N$ Lagrange multipliers, $\alpha_n, \alpha_n^*, n = 1, \ldots N$, a dual functional, γ_d, to be maximized is obtained (see [19] or [16] for more mathematical details):

$$\gamma_d \left\{ \underline{\alpha}, \underline{\alpha}^* \right\} = \\ \left\{ -\frac{1}{2} \sum_{i,j=1}^{N} \left(\alpha_i - \alpha_i^* \right) \left(\alpha_j - \alpha_j^* \right) \varphi \left(\underline{v}_i \right) \cdot \varphi \left(\underline{v}_j \right) + \\ -\epsilon \sum_{n=1}^{N} \left(\alpha_n + \alpha_n^* \right) + \sum_{n=1}^{N} e_n^{(p)} \left(\alpha_n - \alpha_n^* \right) \right\} \quad (9)$$

subject to

$$\sum_{n=1}^{N} \left(\alpha_n - \alpha_n^* \right) = 0 \quad \alpha_n, \alpha_n^* \in [0, C] \quad (10)$$

as

$$\underline{w} = \sum_{n=1}^{N} \left(\alpha_n - \alpha_n^* \right) \varphi \left(\underline{v}_n \right) \quad (11)$$

Consequently, $\hat{\Phi}$ is equal to

$$\hat{\Phi} \left(\underline{v} \right) = \sum_{n=1}^{N} \left(\alpha_n - \alpha_n^* \right) \varphi \left(\underline{v}_n \right) \cdot \varphi \left(\underline{v} \right) + b \quad (12)$$

where only the inner product of the nonlinear mapping function (and not the function itself) appears. This is the well-known *kernel trick* that allows one to deal implicitly with nonlinear mappings through the use of Kernel functions

$$k \left(\underline{v}_i, \underline{v}_j \right) = \varphi \left(\underline{v}_i \right) \cdot \varphi \left(\underline{v}_j \right) \quad (13)$$

The theory of kernels, that is, the conditions under which equation (13) holds, has been known since the beginning of the last century; it is based on Mercer's theorem [16] and has been applied to pattern recognition tasks since the '60s [20]. However, only recently has the connection with learning machines been well formalized [18]. Kernel functions are positive semidefinite functionals. Therefore, using this property and the fact that the constraints of the above optimization problem are "affine", any local minimum is also a global one, and algorithms exist by which the solution can be found in a finite number of steps [21]. Furthermore, if the kernel is strictly positive definite (that is always the case, except in pathological situations), the solution is also unique. These properties overcome many typical drawbacks of classical neural-network approaches, such as the determination of a suitable minimum, the choice of the starting point, the optimal stopping criteria, and so on.

Since the publication of early seminal works on kernel functions, many functionals have been found that satisfy Mercer's theorem. As far as inverse-scattering problems are concerned, a Gaussian kernel

$$k \left(\underline{v}_i, \underline{v}_j \right) = exp \left\{ - \left[\frac{\left\| \underline{v}_i - \underline{v}_j \right\|^2}{2\sigma^2} \right] \right\} \quad (14)$$

performing a mapping in an infinite-dimensional feature space [18] and preliminarily used in [19], has demonstrated its effectiveness.

Concerning the SVM parameters, the threshold b is computed by means of the Karush-Khun-Tucker conditions of the CQP at optimality [19], and the hyper-parameters of the problem (σ, q, C and ϵ) are deduced by accomplishing the model-selection task proposed in [22].

Finally, the CQP is solved by a standard optimization algorithm, namely, Platt's SMO algorithm for regression [23].

IV. NUMERICAL RESULTS

In this work, a comparative study of NN and SVM-based approaches is made concerning the localization of a scatterer buried in the soil. Let us consider a square investigation domain $L_S = \lambda$-sided (λ being the free-space wavelength) completely embedded in the ground. The relative permittivity of the soil is $\varepsilon_S = 20.0$ and the conductivity is $\sigma_S = 0.01 \frac{S}{m}$. The center of the region under test is $L_{OS} = \frac{7}{12}\lambda$ deep. The buried object is a lossless circular cylinder of radius $\rho_B = \frac{1}{12}\lambda$ and the relative permittivity of the ground is equal to $\varepsilon_B = 5.0$. This scenario is illuminated by
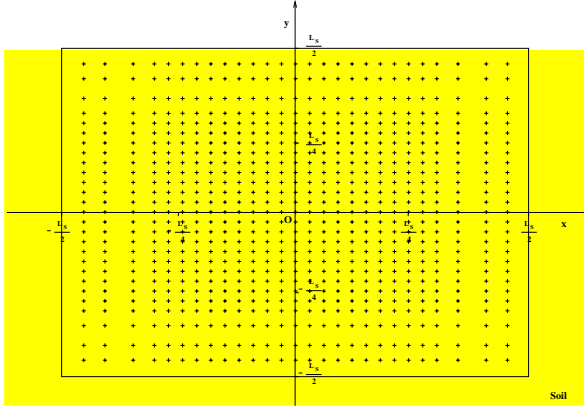
Fig. 2. Training set. Geometrical arrangement of the center of the cylinder under test
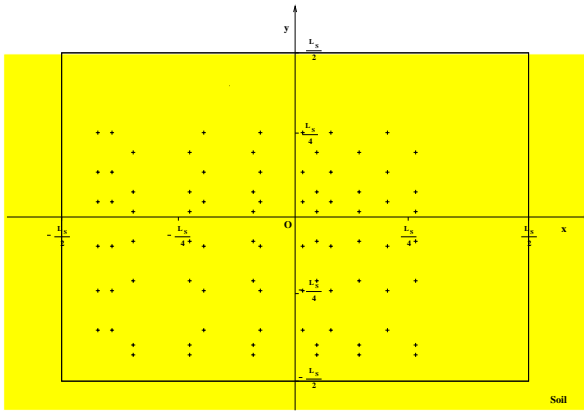

Fig. 3. Test set. Geometrical arrangement of the center of the cylinder under test

an electric line source, located in the upper region with the coordinates $x_t = 0$ and $y_t = \frac{2}{3}\lambda$, $t = T = 1$, and parallel to the air-soil interface. The anomalous field is collected at $R = 16$ measurement points equally spaced ($d = \frac{\lambda}{16}$) and located on a line placed close to the air-soil interface ($L_{ST} = \frac{\lambda}{12}$).

The data used to generate the training set and those used to test the SVM approach, as well as the MLP neural network, were obtained synthetically by a Finite Element code and a PML technique [24]. During the learning phase, the training set, $\{\Gamma_{training}, ; N = 729\}$, was obtained by moving the center of the cylinder inside D_S among the positions shown in Figure 2 and collecting the anomalous field at the receiver positions. As far as the test phase is concerned, $M = 84$ randomly chosen locations of the scatterer (Fig. 3) were considered in order to define the test set $\Gamma_{test} = \{(\underline{\nu}_m, \underline{e}_m); m = 1, ..., M\}$. An additive Gaussian noise, characterized by the signal-to-noise ratio (SNR)

$$SNR = 10 log_{10} \frac{\sum_{r=1}^{T} \sum_{t=1}^{T} |E^{tot}(x_r, y_r | x_t, y_t)|^2}{2T^2 \kappa_{noise}^2}$$

(15)

κ_{noise}^2 being the variance of noise, affected the measured signals.

A two-layer MLP-NN [12], characterized by 32 inputs, 32 hidden neurons, and 2 output neurons, was trained by using a delta-bar-delta back algorithm [25] in order to overcome the shortcomings of the gradient-descent procedure and to increase the convergence rate of the standard back-propagation learning algorithm.

Concerning the SVM-based approach, two SVMs were used to estimate the center coordinates of the buried object. Moreover, after the optimal selection procedure, the values of the SVM hyperparameters turned out to be constant quantities equal to $(C)_{x_B} = (C)_{y_B} = 100$ and $\epsilon = 0.001$. On the contrary, the variance values of the Gaussian kernels, $(\sigma^2)_{x_B}$ and $(\sigma^2)_{y_B}$, were determined independently of each scenario under test.

A. Definitions

In order to quantitatively estimate the effectiveness of the presented approaches, some error values are defined. Let us introduce the

- *local errors* on the center location, δ_x^u and δ_y^v:

$$\delta_x^u = \frac{|x_{act}^u - x_{rec}^{v(u)}|}{d_{max}} \qquad \begin{array}{l} u = 1, ..., U; \\ v(u) = 1, ..., V(u) \end{array}$$
$$\delta_y^v = \frac{|y_{act}^v - y_{rec}^{u(v)}|}{d_{max}} \qquad \begin{array}{l} v = 1, ..., V; \\ u(v) = 1, ..., U(v) \end{array}$$

(16)

- *local average errors* on the object localization, $\zeta_x = \{\zeta_x^u, u = 1, ..., U\}$ and $\zeta_y = \{\zeta_y^v, v = 1, ..., V\}$:

$$\zeta_x^u = \frac{|x_{act}^u - \frac{1}{V(u)} \sum_{v(u)=1}^{V(u)} x_{rec}^{v(u)}|}{d_{max}} \qquad u = 1, ..., U$$
$$\zeta_y^v = \frac{|y_{act}^v - \frac{1}{U(v)} \sum_{u(v)=1}^{U(v)} y_{rec}^{u(v)}|}{d_{max}} \qquad v = 1, ..., V$$

(17)

- *global average errors*, Θ_x and Θ_y:

$$\Theta_x = \frac{1}{d_{max}} \sqrt{\frac{1}{U} \sum_{u=1}^{U} \left[x_{act}^u - \frac{\sum_{v(u)=1}^{V(u)} x_{rec}^{v(u)}}{V(u)} \right]^2}$$
$$\Theta_y = \frac{1}{d_{max}} \sqrt{\frac{1}{V} \sum_{v=1}^{V} \left[y_{act}^v - \frac{\sum_{u(v)=1}^{U(v)} y_{rec}^{u(v)}}{U(v)} \right]^2}$$

(18)

where the subscripts rec and act refer to the estimated and actual coordinates of the scatterer, respectively; $d_{max} = L_S$ is the maximum error in defining the coordinates of the center of the circular scatterer when it is contained in the investigation domain, D_S.
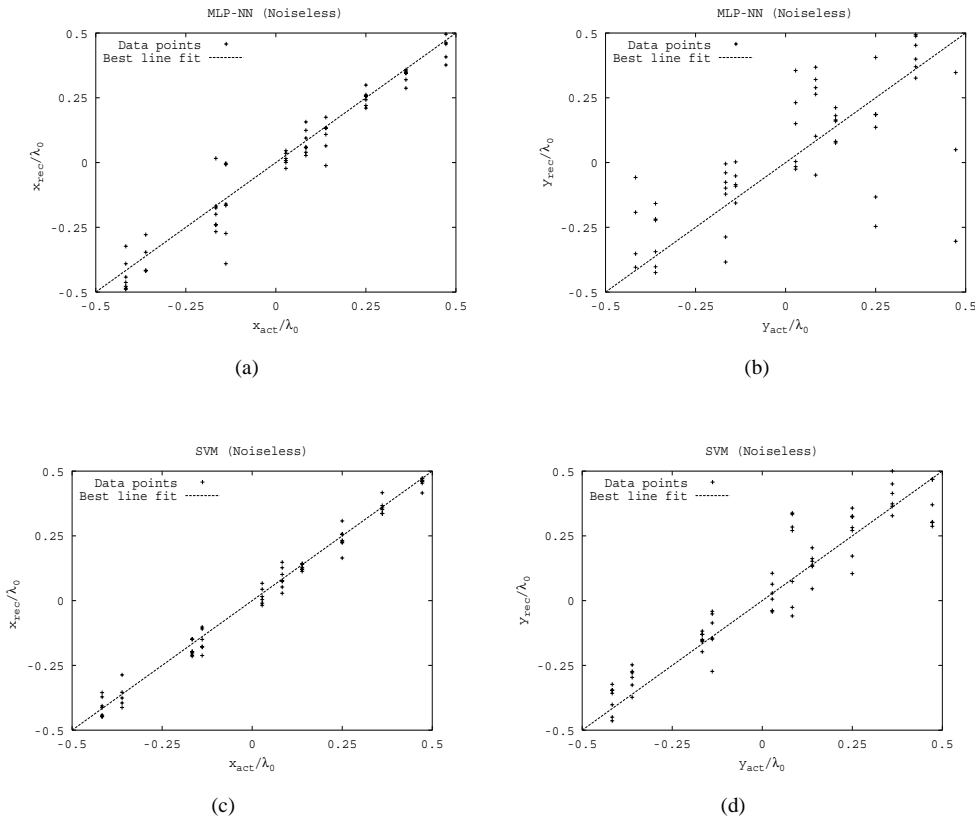
Fig. 4.   Reconstructed data versus actual data (Noiseless Case). (*a*)-(*b*) MLP-NN approach, (*c*)-(*d*) SVM-based approach. (*a*)-(*c*) $\frac{x_B}{\lambda_0}$ and (*b*)-(*d*) $\frac{y_B}{\lambda_0}$

B. Numerical Assessment - Scenario 1

The first example is aimed at evaluating the possibility of locating the buried object starting from the knowledge of the measured electric field and assuming the knowledge of the soil characteristics to be a-priori information about the geometry under test. Consequently, the incident field is a known quantity and the data array is defined as follows: $\underline{E}^{scatt} = \left[E^{tot}(x_r, y_r | x_t, y_t) - E^{inc}(x_r, y_r | x_t, y_t); r = 1, ..., R; t = 1, ..., T \right]$. In this context, the SVM parameters have been chosen equal to $\left(\sigma^2 \right)_{x_B} = 0.64$ and $\left(\sigma^2 \right)_{y_B} = 0.32$.

Figure 4 shows the localization results for the examples making up the test set and obtained by using the MLP-NN (Fig. 4(*a*)-(*b*)) and the SVM-based procedure (Fig. 4(*c*)-(*d*)), respectively. As can be observed, a good accuracy concerning the center location is achieved along the two reference axes and for both the MLP-NN and SVM-based approaches. In particular, even if the detection accuracy decreases as the distance from the air-soil interface increases, good localizations are achieved in the whole domain, as confirmed by the statistics shown in Table I. In particular, as far as the scatterer depth estimation is concerned, the SVM

sharply reduces the error of the MLP-NN, reaching an average error equal to $< \delta_y^v >_{SVM} = 0.0584$ ($< \delta_y^v >_{MLP} = 0.1004$ being the average error made by the MLP-NN approach). Moreover, it should be noted that the time required for the SVM training is about one tenth of the one required by the MLP-NN, whereas there is no significant difference between the computation times taken by the two approaches for the object localization (i.e, after the learning phase).

TABLE I
SCENARIO 1 (*Noiseless Case*). LOCAL ERROR STATISTICS

	$< \delta_x^u >$	$\max\{\delta_x^u\}$	$\min\{\delta_x^u\}$
MLP	0.0347	0.2098	$7.14 \cdot 10^{-4}$
SVM	0.0177	0.1243	$7.30 \cdot 10^{-5}$

	$< \delta_y^v >$	$\max\{\delta_y^v\}$	$\min\{\delta_y^v\}$
MLP	0.1004	0.5631	$5.31 \cdot 10^{-3}$
SVM	0.0584	0.3487	$7.44 \cdot 10^{-4}$

In order to assess the robustness of the learning-based retrieval strategies, a noisy environment has been taken into account. For all the simulations, the buried cylinder and the electromagnetic scenario are unchanged and characterized by the same dielectric
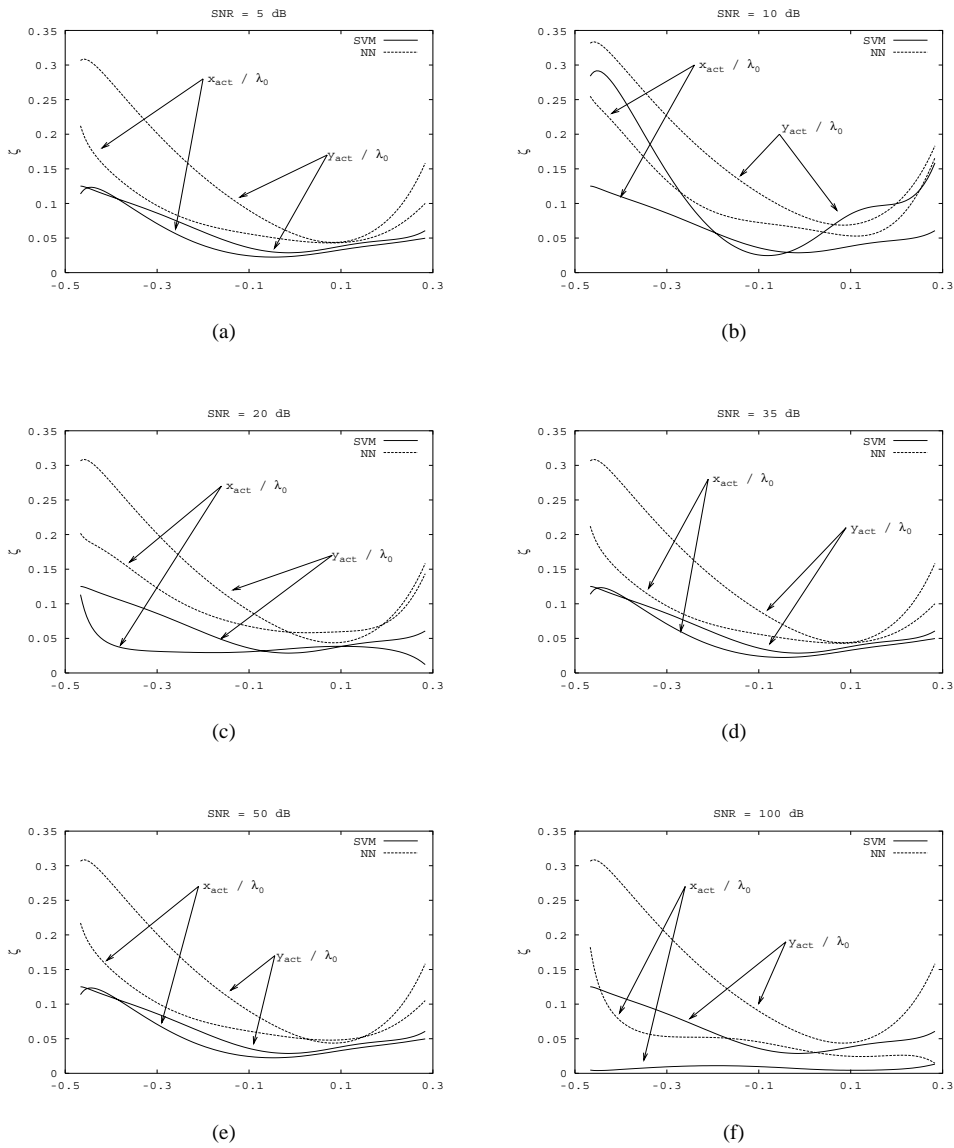
Fig. 5. Scenario 1 (*Standard Validation*). Local average errors of the MLP-NN and SVM-based procedures for different signal-to-noise ratios: (*a*) $SNR = 5\,dB$, (*b*) $SNR = 10\,dB$, (*c*) $SNR = 20\,dB$, (*d*) $SNR = 35\,dB$, (*e*) $SNR = 50\,dB$, and (*f*) $SNR = 100\,dB$

properties as during the training phase. The local average errors are given in Figure 5. For different signal-to-noise ratios, the plots of ζ_x and ζ_y related to both the MLP-NN and the SVM-based procedures are shown. As expected, the estimation of the scatterer depth turns out to be more difficult than the horizontal detection. However, the performances guaranteed by the SVM procedure generally outperform those achieved by the MLP-NN approach. Concerning the dependence of the reconstruction properties on the SNR value, the scatterer is located quite correctly, and $\zeta_x \leq 0.025$ whatever the noisy case considered. Moreover, ζ_y is greater than 0.05 only in the interface regions (i.e., near the air-soil interface and at the bottom of the

investigation area). This behavior is not surprising, as confirmed by the experimental results reported in [26], where the problem of the pollution of the useful signal due to the reflections of the air-ground interface is clearly pointed out.

Another evaluation of the robustness of the proposed approaches has also been obtained by carrying out the so-called *cross validation test*. The two methods have been trained with a noisy data set (i.e., a data set whose samples are related to an assigned signal-to-noise ratio $SNR_{Training}$) and tested with a test set computed in a different noisy environment (SNR_{Test}).
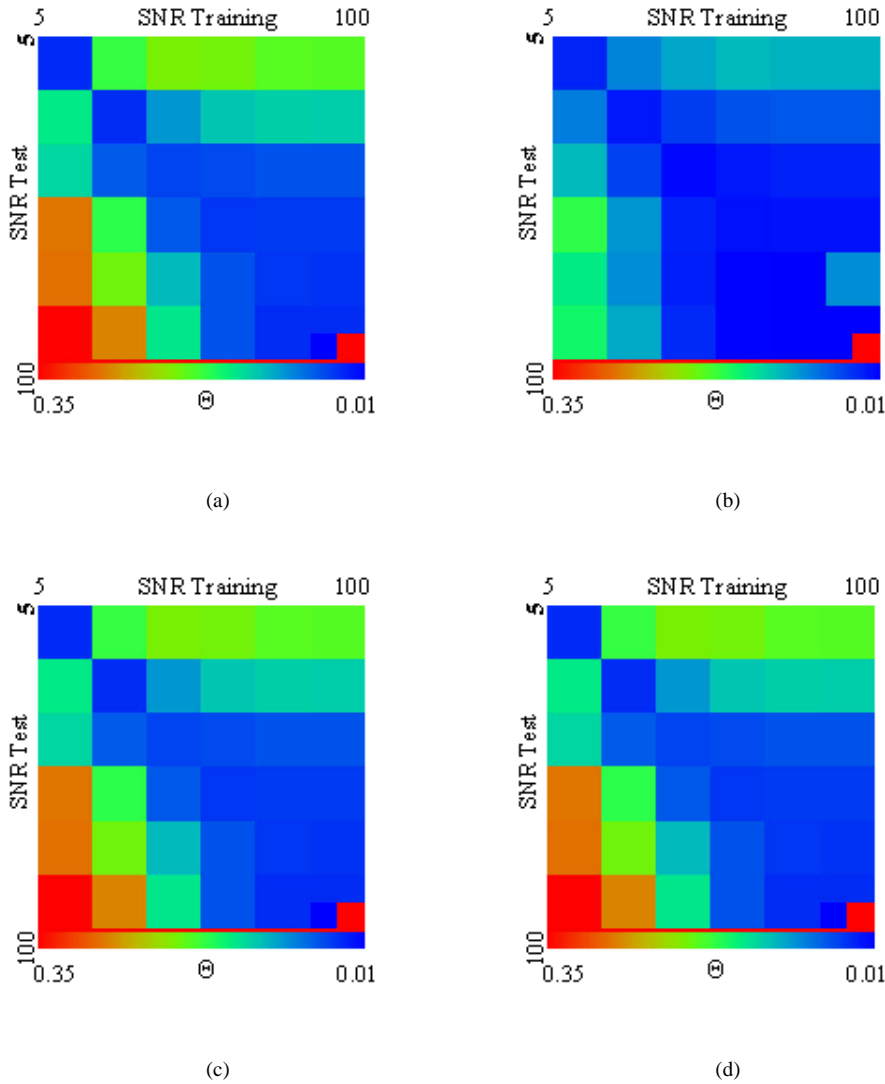
(a)

(b)

(c)

(d)

Fig. 6.   Scenario 1 (*Cross Validation*). Global average errors: (*a*) $\{\Theta_x\}_{MLP-NN}$, (*b*) $\{\Theta_x\}_{SVM}$, (*c*) $\{\Theta_y\}_{MLP-NN}$, and (*d*) $\{\Theta_y\}_{SVM}$

Figure 6 shows a color-level representation (*) of the global average errors for different values of signal-to-noise ratio of the training and test sets ranging from $5\,dB$ to $100\,dB$. Figures 6(*a*)-(*c*) and 6(*b*)-(*d*) refer to the MLP-NN approach and the SVM method, respectively. As expected, the smallest values of the global errors are reached when the same noisy environment is considered for both the training and test data-sets. Otherwise, the SVM method always outperforms the MLP-NN approach in the estimation of the horizontal coordinate of the scatterer (Θ_x). As far as the depth of the scatterer location is concerned, similar conclusions can be drawn for the region defined by the following ranges: $SNR_{Training} \geq 10\,dB$ and

_____

* The two pixels at the right-bottom angles of the plots indicate the minimum and maximum values of the global errors.

$SNR_{Test} \geq 10\,dB$. Otherwise, the comparative study does not provide any significant information.

C. Numerical Assessment - Scenario 2

In the second example, a more complex scenario has been preliminarily considered. No information about the soil is available and the problem data are the measures of the anomalous field, $\underline{E}^{tot} = [E^{tot}(x_r, y_r | x_t, y_t); \quad r = 1, ..., R; \; t = 1, ..., T]$. As far as the choice of the hyperparameters is concerned, the same value equal to 0.04 has been assumed for $(\sigma^2)_{x_B}$ and for $(\sigma^2)_{y_B}$.

As expected (Fig. 7), the performances of the learning-by-examples strategies considerably reduce, as compared with Scenario 1. However, the higher effectiveness of the SVM-based procedure than that
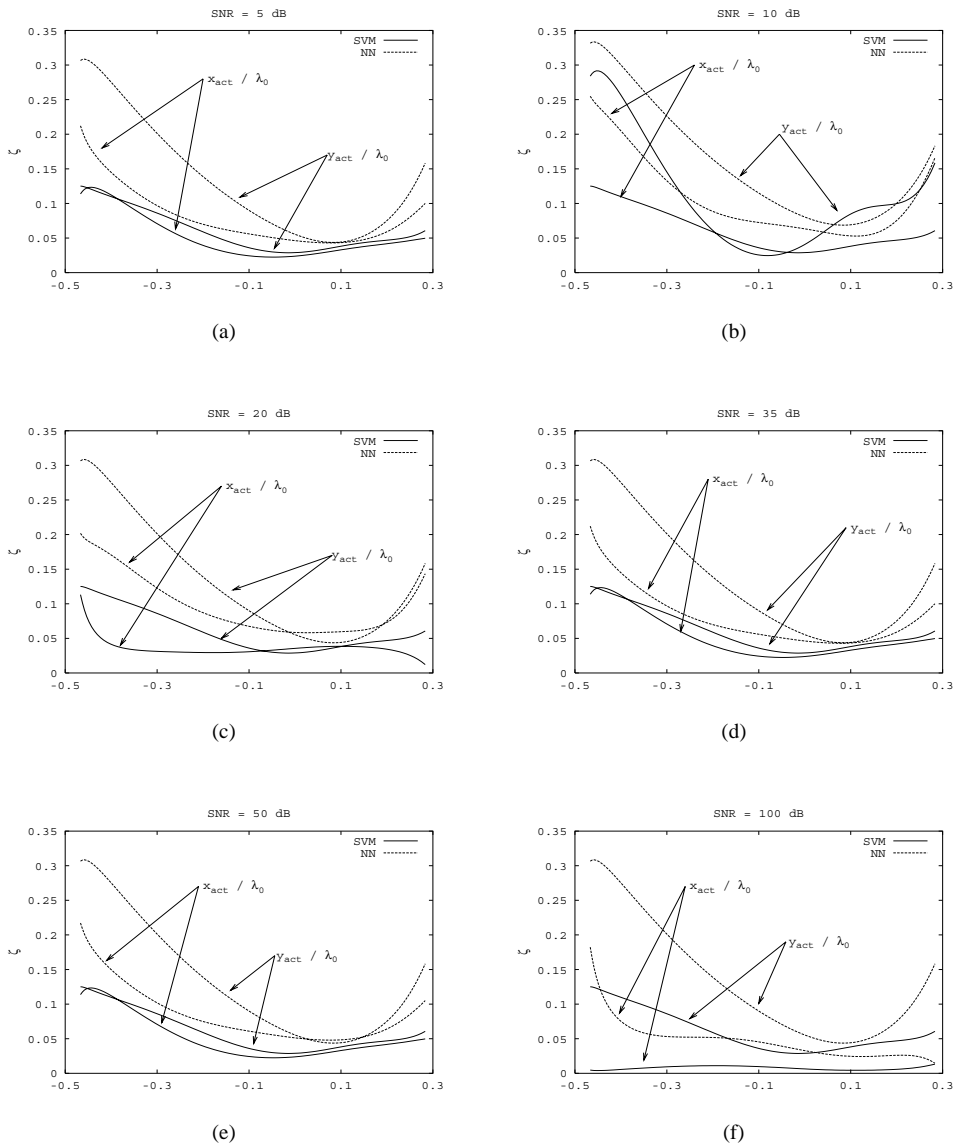
Fig. 7. Scenario 2 (*Standard Validation*). Local average errors of the MLP-NN and SVM-based procedures for different signal-to-noise ratios in the range between $SNR = 5\,dB$ and $SNR = 100\,dB$

of the MLP-NN method is confirmed. Starting from $SNR = 20\,dB$, the local error values turn out to be smaller than 0.15. On the contrary, the performances of the MLP-NN method strongly worsen, as indicated by the dashed lines in Figure 7.

V. CONCLUSIONS

In this paper, two inductive methods for the detection of buried objects have been extensively compared. Starting from an integral formulation of the scattering equations, the buried-object localization has been reformulated as a regression problem and successively solved by means of two learning-by-examples strategies, namely, the MLP-NN approach and the SVM-based procedure. The estimation of the effectiveness

of the proposed procedures has been carried out in different test cases that have clearly confirmed the higher robustness of the SVM-based approach in solving difficult approximation problems as compared with traditional neural networks. Several scenarios have been considered and the behaviors of the two inductive models have been illustrated for different operating conditions. The obtained results have demonstrated the successful application of the SVM-based procedure to solve inverse-scattering problems. Future work, currently under development, will be devoted to improving the performances of the SVM-based procedure and to determining customized kernel functions.

REFERENCES

[1] C. Dourthe, C. Pichot, J. C. Dauvignac, and J. Cariou, "Inversion algorithm and measurement system for microwave tomography of buried object," *Radio Science*, vol. 35, pp. 1097-1108, 2000.

[2] T. J. Cui and W. C. Chew, "Novel diffraction tomographic algorithm for imaging of two-dimensional targets buried under a lossy earth," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 38, pp. 2033-2041, 2000.

[3] A. Joisel and J. C. Bolomey, "Rapid microwave imaging of living tissues," *Proc. of the SPIE - The International Society for Optical Eng.*, vol. 39, pp. 320-330, 2000.

[4] P. Chaturvedi and R. G. Plumb, "Electromagnetic imaging of underground targets using constrained optimization," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 33, pp. 551-561, 1995.

[5] T. J. Cui, W. C. Chew, A. A. Aydiner, and S. Chen, "Inverse scattering of two-dimensional dielectric objects buried in a lossy earth using distorted Born iterative method," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, pp. 339-345, 2001.

[6] S. Caorsi, G. L. Gragnani, and M. Pastorino, "Numerical electromagnetic inverse-scattering solutions for two-dimensional infinite dielectric cylinders buried in a lossy half-space," *IEEE Trans. Microwave Theory and Technique*s, vol. 41, pp. 352-356, 1993.

[7] K. Belkebir, R. E. Kleinman, and C. Pichot, "Microwave imaging-location and shape reconstruction from multifrequency scattering data," *IEEE Trans. on Microwave Theory and Techniques*, vol. 45, pp. 469-76, 1997.

[8] I. T. Rekanos, T. V. Yioultsis, and T. D. Tsiboukis, "Inverse scattering using the finite-element method and a nonlinear optimization technique," *IEEE Trans. on Microwave Theory and Techniques*, vol. 47, pp. 336-344, 1999.

[9] S. Caorsi, A. Massa, M. Pastorino, "A computational technique based on a real-coded genetic algorithm for microwave imaging purposes," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 38, pp. 1697-1708, 2000.

[10] S. Caorsi, A. Massa, and M. Pastorino, " A microwave inverse scattering technique for image reconstruction based on a genetic algorithm," *IEEE Trans. on Instrumentation and Measurement*, vol. 49, pp. 573-578, 2000.

[11] C. Christodoulou and M. Georgiopoulos. *Applications of Neural Networks in Electromagnetics*. Boston, Artech House, 2001.

[12] S. Caorsi and P. Gamba, "Electromagnetic detection of dielectric cylinders by a neural network approach," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 37, pp. 820-827, 1999.

[13] E. Bermani, S. Caorsi, and M. Raffetto, "An inverse scattering approach based on a neural network technique for the detection of dielectric cylinders buried in a lossy half-space," in *Progress in Electromagnetic Research*, vol. 26, pp. 69-90, 2000.

[14] I. T. Rekanos, "On-line inverse scattering of conducting cylinders using radial-basis -function neural networks," *Microwave and Optical Technology Letters*, vol. 28, pp. 378-380, 2001.

[15] M. Anthony and P. Bartlett. *Neural Network Learning: Theoretical Foundations*. Cambridge, University Press, 1999.

[16] V. N. Vapnik. *The Nature of Statistical Learning Theory*. John Wiley & Sons, New York, 1999.

[17] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Macmillan, New York, 1994.

[18] B. Schölkopf and A. Smola. *Learning with Kernels*, The MIT Press, 2002.

[19] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines*. Cambridge, University Press, 2000.

[20] M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer, "Theoretical foundations of the potential function method in pattern recognition learning," *Automation and Remote Control*, vol. 25, pp. 821-837, 1964.

[21] C.-J. Lin, "Asymptotic convergence of an SMO algorithm without any assumptions," *IEEE Trans. on Neural Networks*, vol. 13, no. 1, pp. 248-250, 2002.

[22] D. Anguita, A. Boni, and S. Ridella, "Evaluating the generalization ability of Support Vector Machines through the bootstrap," *Neural Processing Letters*, vol. 11, pp. 1-8, 1999.

[23] J. Platt, "Fast training of support vector machines using sequential minimal optimization," in *Advances in Kernel Methods - Support Vector Learning*, The MIT Press (B. Schölkopf, C. Burges, and A. Smola, Eds.), 1999.

[24] S. Caorsi and M. Raffetto, "Perfectly matched layer for truncation of the finite element meshes in layered half-space geometries and applications to the electromagnetic scattering by buried objects," *Microwave and Optical Tech. Lett.*, vol. 19, pp. 427-434, 1999.

[25] R. A. Jacobs, "Increased rates of convergence through learning rate adaption," *Neural Networks*, vol. 1, no. 4, pp. 295-307, 1988.

[26] P. Millot, P.Borderies, E. Guillanton, E. Le Brusq, J. Y. Dauvignac, and Ch. Pichot, "Radar imaging of shallow buried objects using an ultra-wideband technique with specific antennas and synthetic antenna processing," in *Proc. PIERS 2000*, Cambridge, MA, p. 298, July 2000.

Salvatore Caorsi received the laurea degree in electronic engineering from the University of Genoa, Italy, in 1973. Since 1994, he has been a Full Professor of Electromagnetic Compatibility at the Department of Electronics, University of i Pavia, Italy. He is also teaching the course of Antennas at the University of Genoa. His primary activities focus in applications of electromagnetic field to telecommunications, artificial vision and remote sensing, biology and medicine. In particular, he is working on research project concerning human hazard to electromagnetic exposure, numerical methods for solving electromagnetic i problems, wave interaction in the presence of non linear media, inverse scattering and microwave imaging, and electromagnetic compatibility. Prof. Caorsi is the past president and founding member of the Inter-University Research Center for the Interactions Between Electromagnetic Fields and Biological Systems (ICEmB). Moreover, he is a member of the Elettrotecnica ed Elettronica Italiana (AEI), the European Bioelectromagnetic Association (EBEA), and the European Society for Hyperthermic Oncology (ESHO).

Davide Anguita graduated in Electronic Engineering in 1989 and obtained the Ph.D. in Computer Science and Electronic Engineering at the University of Genova, Italy, in 1993. After working as a research associate at the International Computer Science Institute, Berkeley, USA, on special-purpose processors for neurocomputing, he joined the Dept. of Biophysical and Electronic Engineering at the University of Genova, where he teaches digital electronics. His current research focuses on industrial applications of artificial neural networks and kernel methods and their implementation on digital and analog electronic devices. He is a member of IEEE and chair of the Smart Adaptive Systems committee of the European Network on Intelligent Technologies (EUNITE).

Emanuela Bermani received the "laurea" degree in Electronic Engineering, and the Ph.D. degree in "Electronic and Computer Sciences" from the University of Pavia, Pavia, Italy, in 1997 and in 2000, respectively. She is currently an Assistant Professor in the Department of Information and Communication Technology, University of Trento, Italy. Her main interests are in the fields of electromagnetic direct and inverse scattering, neural networks in electromagnetics, and biomedical applications of electromagnetic fields.

Andrea Massa received the "laurea" degree in Electronic Engineering from the University of Genoa, Genoa, Italy, in 1992 and Ph.D. degree in electronics and computer science from the same university in 1996. From 1997 to 1999 he was an Assistant Professor of Electromagnetic Fields at the Department of Biophysical and Electronic Engineering (University of Genoa) teaching the university course of Electromagnetic Fields 1. Since 2001, he has been Associate Professor of at the University of Trento where he teaches the courses of "Electromagnetic Fields 1", "Electromagnetic Techniques for Biomedical and Industrial Diagnostics", and "Wireless communications 1". At present, Prof. Massa is the director of the ELEDIALab and a member of the Inter-University Research Center for Interactions Between Electromagnetic Fields and Biological Systems (ICEmB). His research work since 1992 has been principally on electromagnetic direct and inverse scattering, optimization techniques for microwave imaging, wave propagation in presence of nonlinear media, applications of electromagnetic fields to telecommunications, medicine and biology.

Andrea Boni was born in Genova, Italy, in 1969 and graduated in Electronic Engineering in 1996. He received a Ph.D. degree in Electronic and Computer Science in 2000. After working as research consultant at DIBE, University of Genova, he joined the Department of Information and Communication Technologies, University of Trento, Italy, where he teaches digital electronics. His main scientific interests are on the study and development of digital circuits for advanced information processing, with particular attention to programmable logic devices, digital signal theory and analysis, statistical signal processing, statistical learning theory and support vector machines. The application of such interests focuses on identification and control of non-linear systems, pattern recognition, time series forecasting, signals processing.

Massimo Donelli received the "laurea" degree in Electronic Engineering from the University of Genoa, Italy, in 1998. He is currently a Ph. D. student in Space Science and Engineering and works as a consultant for the Department of Information and Communication Technology, University of Trento, Italy. His main interests are on electromagnetic inverse scattering, adaptive antennas synthesis, optimization techniques for microwave imaging, wave propagation in superconducting materials.