UNIVERSITY
OF TRENTO

DEPARTMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY

38050 Povo – Trento (Italy), Via Sommarive 14
http://www.dit.unitn.it

A CLASSIFICATION APPROACH BASED ON SVM FOR ELECTROMAGNETIC SUB-SURFACE SENSING

Andrea Massa, Emanuela Bermani, Andrea Boni, and Massimo Donelli

August 2004

Technical Report DIT-04-069

A Classification Approach based on SVM for Electromagnetic Sub-Surface Sensing

Andrea Massa, *Member*, *IEEE*, Emanuela Bermani, *Member*, *IEEE*, Andrea Boni, and Massimo Donelli, *Member*, *IEEE*

Department of Information and Communication Technologies,

University of Trento, Via Sommarive 14, 38050 Trento - Italy

Tel. +39 0461 882057, Fax +39 0461 882093, E-mail: *andrea.massa@ing.unitn.it*

A Classification Approach based on SVM for Electromagnetic Sub-Surface Sensing

Andrea Massa, Andrea Boni, Massimo Donelli, and Emanuela Bermani

Abstract

In clearing terrains contamined or potentially contamined by landmines and/or unexploded ordnances (UXOs), a quick wide-area surveillance is often required. Nevertheless, the identification of dangerous areas (instead of the detection of each subsurface object) can be enough for some scenarios/applications, allowing a suitable level of security in a cost-saving way.

In such a framework, this paper describes a probabilistic approach for the definition of risk maps. Starting from the measurement of the scattered electromagnetic field, the probability of occurrence of dangerous targets in an investigated subsurface area is determined through a suitably defined classifier based on a Support Vector Machine (SVM). To assess the effectiveness of the proposed approach and to evaluate its robustness, selected numerical results related to a two-dimensional geometry are presented.

1 Introduction

In the world, there are many areas contaminated (or potentially contamined) with unexploded ordnance (UXO) and anti-tank/anti-personnel landmines. Published reports indicate that the approximate number of UXOs is of about 110 millions in 70 countries [1]. Moreover, recent missions in Bosnia, Afghanistan and Iraq probably (and unfortunately) will further increase such an estimate. To return these zones to a civilian use, the ordnances should be obviously removed. However, in several cases, the former bombing ranges have been unused for many years and the UXO locations are partially known or completely unknown. Thus, a wide-area surveillance is needed in order to circumscribe those regions where the dangerous targets reside. Such a process is inevitably time-expensive and involves complex acquisition procedures. Consequently, high costs (from few hundred dollars for acre in the case of surface or shallow targets up to a couple of million dollars for subsurface objects) occur. This is one of the main motivation of the growing research interest in developing unsupervised techniques able to effectively (in terms of time and resources) repair landmine/UXO contaminated areas. Several solutions have been proposed based on various methodological approaches (see, for instance, [2] and the references cited therein), which consider different sensor modalities such as ground-sensors (e.g., magnetometers, electromagnetic radars, sensors based on electromagnetic induction, etc.) or synthetic aperture radars. In general, these techniques are aimed at achieving the following goals: (a) correctly localizing a large number of dangerous targets, thus ensuring the future security of cleaned areas; (b) reducing the false-alarm rate, which strongly contributes to the costs of the clearing procedure; (c) reducing the

time devoted to the detection process, thus realizing a *quick* area surveillance.

In such a framework, electromagnetic approaches based on learning-by-examples (LBE) techniques [3]-[5] have been recently proposed for the on-line[1] detection of subsurface objects. The detection process is recast as a regression problem where the unknowns (i.e., the position as well as the geometric and dielectric characteristics of the target) are directly evaluated from the data (i.e., the values of the scattered field) by approximating the data-unknowns relation through an off-line data fitting process (*training phase*). LBE regression-based approaches demonstrated their effectiveness in dealing with detection processes where a limited number of unknowns is considered. However, because of the complexity of the underlying architecture, some difficulties occur when a larger number of unknowns is taken into account. As a consequence, LBE regression-based approaches turn out to be very effective for the detection of a single (or few) buried object, whereas they are not-so-suitable in dealing with the detection of multiple targets. On the other hand, it should be pointed out that the identification of free-areas and an estimate of the concentration of subsurface objects (instead of the localization of each buried scatterer) might be enough in several situations. Then, the goal of a subsurface sensing technique could be moved from the "object detection" to the "definition of a risk map". Consequently, a classification approach, instead of a regression one, should be employed.

In this paper, a classification approach based on a LBE technique is proposed for an on-line sub-surface sensing. Starting from the knowledge of the scattered field values collected above the surface, the method is aimed at defining a risk map of the domain under test. By considering a SVM-based classifier, the proposed method estimates

[1]After the *learning process* (or *training phase*) performed once and off-line.

the *a-posteriori* probability of the presence of subsurface dangerous objects.

The paper is organized as follows. The electromagnetic problem is formulated and the SVM-based approach is outlined in Sections 2 and 3. For the numerical assessment, a selected set of results is presented to assess the effectiveness of the proposed approach (Sect. 4). Towards this end, a two-dimensional problem is dealt with. Both noiseless and corrupted measurement data are considered to check the robustness of the proposed approach. Finally, some conclusions and final remarks are presented in Sect. 5.

2 Problem Formulation

Let us consider a typical two-dimensional buried-object scenario as shown in Fig. 1. The upper region presents the same characteristics of the vacuum ($\varepsilon_{r1} = 1.0$, $\sigma_1 = 0.0$). The lossy subsurface region, which models the soil, is characterized by a conductivity σ_2 and by a relative dielectric permittivity ε_{r2} . Moreover, let us assume that the investigation domain D_I lies entirely in the subsurface medium, $D_I = \left\{-\frac{L}{2} \leq x \leq \frac{L}{2}, -\frac{L}{2} \leq y \leq \frac{L}{2}\right\}$. A set of targets (either dielectric or lossy) are supposed to belong to D_I and illuminated by T transmitters located at known positions (x_t, y_t), $t = 1, ..., T$ above the air-ground interface. Let $\underline{E}_{inc}$ be the so-called "incident field", i.e. the electric field distribution due to the transmitters in the absence of buried scatterers. The "scattered field" $\underline{E}_{scat}$ is collected by a set of sensors placed at given positions (x_r, y_r), $r = 1, ..., R$ close to the air-ground interface.

Under the hypotheses of an isotropic background medium and that the electromag-

netic sources be z-directed electric line currents, both incident and scattered electric fields are also z-directed ($\underline{E}_{inc} = E_{inc}(x,y)\widehat{z}$, $\underline{E}^{(t)}_{scat} = E^{(t)}_{scat}(x,y)\widehat{z}$). To define a risk map, let us model the investigation domain through a two-dimensional lattice of M rectangular cells of linear dimension l and h whose center coordinates are $(x_m,\, y_m)$, $m = 1, ..., M$. The state χ_m of the m-th cell can be either empty (if any scatterer belongs to the cell) $\chi_m = -1$ or occupied $\chi_m = +1$. Then, the problem can be formulated as follows: "*starting from the scattered field measurements* $\underline{\Gamma}_E = \left\{ E^{(t)}_{scat}(x_r, y_r);\ r = 1, ..., R;\ t = 1, ..., T \right\}$, *find the probability* q_m *that the* $m-th$ *cell is occupied* $(m = 1, ..., M)$". That is, *determine the probability array* $\underline{Q}$, which is a function $\Im$ of the scattering data $\underline{\Gamma}_E$

$$\underline{Q} = Pr\left\{ \underline{\chi} = \underline{1} \,|\underline{\Gamma}_E \right\} = \Im\left(\underline{\Gamma}_E\right) \tag{1}$$

where $\underline{Q} = \{q_m,\, m = 1, ..., M\}$ and $\underline{\chi} = \{\chi_m,\, m = 1, ..., M\}$. Such a statement defines a classification problem. A solution based on the SVM will be detailed in the following by assuming the knowledge of a set of known examples (i.e., input-output relations $\left\{ (\underline{\Gamma}_E,\, m,\, \chi_m;\, m = 1, ..., M)^{(n)} ;\ n = 1, ..., N \right\}$ called *training set*).

3 SVM-based Classification Approach

The proposed SVM-based classification approach is formulated as a two-step procedure

- *Step 1:* to determine a decision function $\hat{\Phi}$ that correctly classifies an input pattern $(\underline{\Gamma}_E,\, m)$ (not-necessarily belonging to the training set);

- *Step 2:* to map the decision function $\hat{\Phi}\{(\underline{\Gamma}_E,\, m)\}$ into an *a-posteriori* probability $Pr\left\{\underline{\chi} = \underline{1}\,|\underline{\Gamma}_E\right\}$

3.1 Definition of the Decision Function

At this step, the status χ_m of each cell of the lattice has to be determined. Mathematically, such a problem formulates in the definition of a suitable discriminant function $\hat{\Phi}$ separating the two classes $\chi = 1$ and $\chi = -1$. Since these classes are non-linearly separable, the definition of a non-linear (in terms of the original data $\underline{\Gamma}_E$) discriminant function is usually required. As a matter of fact, such a solution is implemented when Artificial Neural Networks (ANN) are considered (see [6] and the references cited therein).

Conversely, SVM defines a linear decision function corresponding to a hyperplane that maximizes the separating margin between the classes. Such a linear data-fitting is not carried out in the original input space $\Re\{\underline{\Gamma}_E\}$, but in a higher dimensional space $\aleph\left\{\underline{\varphi}\left(\underline{\Gamma}_E\right)\right\}$(called *feature space*) where the original examples are mapped through a non-linear operator $\underline{\varphi}(\bullet)$. The nonlinear SVM classifier so obtained is defined as

$$\hat{\Phi}\left(\underline{\varphi}\left(\underline{\Gamma}_E,\, m\right)\right) = \underline{w} \cdot \underline{\varphi}\left(\underline{\Gamma}_E,\, m\right) + b \qquad m = 1, ..., M \tag{2}$$

where $\underline{w}$ and b are the parameters of $\hat{\Phi}$ to be determined during the training phase. The hyperplane so-defined causes the largest separation between the decision function values for the "borderline" training examples from the two classes. Mathemati-

cally, such a hyperplane can be found by minimizing the following cost function

$$\Omega\left(\underline{w}\right) = \frac{1}{2}\left\|\underline{w}\right\|^2 \tag{3}$$

subject to the separability constraints

$$\begin{aligned} \underline{w}\cdot\underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right) + b \geq 1 \qquad for \qquad \chi_m^{(n)} = 1 \quad m = 1, ..., M \\ \underline{w}\cdot\underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right) + b \leq -1 \qquad for \qquad \chi_m^{(n)} = -1 \quad n = 1, ..., N \end{aligned} \tag{4}$$

However, since the training data in the feature space are generally non-completely separable by a hyperplane, *slack variables* (denoted by $\xi_{(m)}^{(n)}$) are introduced to relax the separability constraints in (4) as follows

$$\begin{aligned} \underline{w}\cdot\underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right) + b \geq 1 - \xi_{(m)+}^{(n)} \qquad for \qquad \chi_m^{(n)} = 1 \quad m = 1, ..., M \\ \underline{w}\cdot\underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right) + b \leq \xi_{(m)-}^{(n)} - 1 \qquad for \qquad \chi_m^{(n)} = -1 \quad n = 1, ..., N \end{aligned} \tag{5}$$

Accordingly, the cost function in (3) modifies

$$\Omega\left(\underline{w}\right) = \frac{\left\|\underline{w}\right\|^2}{2} + \lambda_+ \sum_{m=1}^{M}\sum_{n=1}^{N_{(m)}^+} \xi_{(m)+}^{(n)} + \lambda_- \sum_{m=1}^{M}\sum_{n=1}^{N_{(m)}^-} \xi_{(m)-}^{(n)} \tag{6}$$

where $N_{(m)}^+$ and $N_{(m)}^-$ indicate the number of training examples for which $\chi_m^{(n)} = 1$ and $\chi_m^{(n)} = -1$, respectively; $\lambda_+ = \frac{C}{\sum_{m=1}^{M} N_{(m)}^+}$ and $\lambda_- = \frac{C}{\sum_{m=1}^{M} N_{(m)}^-}$ [7]. The user-defined hyperparameter C controls the trade-off between the *empirical risk* (i.e., the training errors) and the model complexity (the first term in (6)) to avoid the *overfitting*. In that case, the decision boundary too precisely corresponds to the training data. Thereby, the method is unable to deal with data outside the training

set [8].

To minimize (6), it can be observed that a necessary condition is that $\underline{w}$ is a linear combination of the mapped vectors $\underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right)$

$$\underline{w} = \sum_{m=1}^{M} \sum_{n=1}^{N} \left\{ \alpha_m^{(n)} \chi_m^{(n)} \underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right) \right\} \tag{7}$$

where $\alpha_m^{(n)} \geq 0$, $n = 1, ..., N$, $m = 1, ..., M$ are Lagrange multipliers to be determined. Moreover, from the Karush-Khun-Tucker conditions at the optimality [9], b turns out to be expressed as follows

$$b = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N_{sv}} \left\{ \chi_m^{(n)} - \sum_{p=1}^{N} \left\{ \alpha_m^{(p)} \underline{\varphi}\left(\underline{\Gamma}_E^{(n)},\, m\right) \cdot \underline{\varphi}\left(\underline{\Gamma}_E^{(p)},\, m\right) \right\} \right\}}{N_{sv}} \tag{8}$$

N_{sv} being the number of patterns $\left(\underline{\Gamma}_E^{(n)},\, m\right)$ for which $\alpha_m^{(n)} \neq 0$ (called *support vectors*).

Substituting (7) and (8) in (2) yields

$$\begin{array}{c} \hat{\Phi}\left(\underline{\varphi}\left(\underline{\Gamma}_E\right),\, m\right) = \sum_{m=1}^{M} \sum_{n=1}^{N} \left\{ \alpha_m^{(n)} \chi_m^{(n)} \Theta\left(\underline{\Gamma}_E^{(n)},\, \underline{\Gamma}_E,\, m\right) \right\} \\ + \frac{\sum_{m=1}^{M} \sum_{n=1}^{N_{sv}} \left\{ \chi_m^{(n)} - \sum_{p=1}^{N} \left\{ \alpha_m^{(p)} \Theta\left(\underline{\Gamma}_E^{(n)}, \underline{\Gamma}_E^{(p)}, m\right) \right\} \right\}}{N_{sv}} \end{array} \tag{9}$$

where $\Theta\left(\underline{\Gamma}_E^{(i)},\, \underline{\Gamma}_E^{(j)},\, m\right) = \underline{\varphi}\left(\underline{\Gamma}_E^{(i)},\, m\right) \cdot \underline{\varphi}\left(\underline{\Gamma}_E^{(j)},\, m\right)$ is a suitable *kernel function*[2]. Then, the decision function is completely determined when the Lagrange multipliers are specified. Towards this end, the constrained optimization problem formulated

[2]For a detailed discussion (out-of-the scope of this paper), the interested reader may refer to [10]

in (6) and (5) is reformulated in a more practical dual form

$$max_{\underline{\alpha}} \left\{ \Omega_{Dual} \left(\underline{\alpha} \right) \right\} = \\ max_{\underline{\alpha}} \left\{ \frac{\sum_{n=1}^{N} \sum_{p=1}^{N} \sum_{m=1}^{M} \left[\alpha_m^{(n)} \alpha_m^{(p)} \chi_m^{(n)} \chi_m^{(p)} \Theta \left(\underline{\Gamma}_E^{(n)}, \underline{\Gamma}_E^{(p)}, m \right) \right]}{2} - \sum_{n=1}^{N} \sum_{m=1}^{M} \alpha_m^{(n)} \right\} \tag{10}$$

subject to $\sum_{n=1}^{N} \sum_{m=1}^{M} \alpha_m^{(n)} \chi_m^{(n)} = 0$, $\alpha_m^{(n)} \in [0, \lambda_-]$ if $\chi_m^{(n)} = -1$ and $\alpha_m^{(n)} \in [0, \lambda_+]$ otherwise.

Finally, since $\Omega_{Dual} \left(\underline{\alpha} \right)$ is a convex and quadratic function of the unknown parameters $\alpha_m^{(n)}$, it is solved numerically by means of a standard quadratic programming technique (e.g., the Platt's SMO algorithm for classification [11]).

3.2 Mapping of the Decision Function into the *A-Posteriori* Probability

Concerning standard classification, the SVM classifier labels an input pattern according to the following rule [12]

$$\chi_m = sign \left\{ \hat{\Phi} \left(\underline{\varphi} \left(\underline{\Gamma}_E, \, m \right) \right) \right\} \qquad m = 1, ..., M \tag{11}$$

Unlike standard approaches, the proposed method is aimed at defining an *a-posteriori* probability. Consequently, some modifications to the standard SVM-based classification approach are needed. Towards this aim, a set of efficient solutions has been recently proposed (see, for instance, [10],[13]-[15]) either based on a direct training of the SVM with a logistic link function and a regularized maximum likelihood score or based on *a-posterior* fitting probability process.

The first class of approaches usually leads to non-sparse kernel machines and requires a significant modification of the SVM structure. In this paper, the *a-posteriori* probability fitting method [15] is adopted since the use of a parametric model allows a direct fitting of the *a-posteriori* probability $Pr\left\{\underline{\chi} = \underline{1} \,|\underline{\Gamma}_E\right\}$. More in detail, such a model approximates the *a-posteriori* probability through a sigmoid function

$$Pr\left\{\chi_m = 1 \,|(\underline{\Gamma}_E,\, m)\right\} = \frac{1}{1 + exp\left\{\gamma\hat{\Phi}\left(\underline{\varphi}\left(\underline{\Gamma}_E,\, m\right)\right) + \delta\right\}} \qquad m = 1, ..., M \tag{12}$$

where γ and δ are unknown parameters.

To estimate the optimal values for the parameters of the sigmoid function, a fitting process is performed. A subset of the input patterns of the training set is chosen $\left\{(\underline{\Gamma}_E,\, m,\, \chi_m;\, m = 1, ..., M)^{(s)}\,;\, s = 1, ..., S\right\}$, where $\widehat{\Phi}_m^{(s)} = \hat{\Phi}\left(\underline{\varphi}\left(\underline{\Gamma}_E^{(s)},\, m\right)\right)$. Then, the following cost function is defined

$$\begin{gathered}\Upsilon\left\{\gamma, \delta\right\} = \\ -\sum_{s=1}^{S}\sum_{m=1}^{M}\left\{\frac{\chi_m^{(s)}+1}{2} log\left[\frac{1}{1+exp\left(\gamma\widehat{\Phi}_m^{(s)}+\delta\right)}\right] + \left(\frac{1-\chi_m^{(s)}}{2}\right) log\left[\frac{exp\left(\gamma\widehat{\Phi}_m^{(s)}+\delta\right)}{1+exp\left(\gamma\widehat{\Phi}_m^{(s)}+\delta\right)}\right]\right\}\end{gathered} \tag{13}$$

and successively minimized according to the numerical procedure proposed in [15].

4 Numerical Results

In this section, the results of a set of numerical examples are analyzed to assess the effectiveness, but also current limitations, of the proposed approach.

With reference to the geometry shown in Fig. 1, the following geometric and dielectric parameters are considered. The relative permittivity and the conductivity of the

homogeneous subsurface region are $\varepsilon_{r2} = 4.0$ and $\sigma_2 = 1 \times 10^{-3}\, S/m$, respectively. The investigation domain D_I is a $2.0\,\lambda \times 2.0\,\lambda$ region and it was discretized in a two-dimensional lattice of $M = 36$ square cells. The buried objects, modeling UXOs or landmines, are lossless circular cylinders of diameter $d_{cil} = \frac{\lambda}{6}$ and characterized by a relative permittivity $\varepsilon_{cil} = 5.0$. Concerning the measurement system, $R = 16$ ideal receivers are equally spaced along an observation line $2.0\,\lambda$-long and parallel to the air-ground interface at a distance $h_r = 0.6\,\lambda$ from the surface. The probing source ($t = 1$) is located at $x_t = 0.0$ and $y_t = \frac{7}{6}\,\lambda$.

The measurement data were synthetically computed by using a finite-element-based simulator and a PML truncation technique [17]. As far as the *training set* is concerned, various examples $N = N^{(2)} + N^{(3)}$ ($N^{(2)}$ and $N^{(3)}$ being the numbers of patterns related to two and three-targets configurations, respectively) were considered. In particular, $N^{(2)} = 35$ and $N^{(3)} = 34$. Moreover, these patterns have been also used during the *validation test* aimed at defining the *a-posteriori* fitting model. The optimal values of the fitting parameters turned out to be $\gamma = -0.533$ and $\delta = 1.272$.

Concerning the SVM structure, Gaussian kernel functions were adopted and their parameters selected according to [16].

In order to analytically evaluate the effectiveness of the classification method in correctly locating the dangerous areas, a *dangerous-area-location error* ς is then defined

$$\varsigma = \frac{\sqrt{(x_{cil} - \widetilde{x}_{cil})^2 - (y_{cil} - \widetilde{y}_{cil})^2}}{\lambda} \tag{14}$$

where $(x_{cil},\, y_{cil})$ and $(\widetilde{x}_{cil},\, \widetilde{y}_{cil})$ are the actual and estimated coordinates of the center

of a dangerous zone, respectively. Being

$$\widetilde{x}_{cil} = \frac{\sum_{m=1}^{M_{(j)}} \{x_m Pr\left(\chi_m = 1 \,|\underline{\Gamma}_E\right)\}}{\sum_{m=1}^{M_{(j)}} \{Pr\left(\chi_m = 1 \,|\underline{\Gamma}_E\right)\}} \qquad \widetilde{y}_{cil} = \frac{\sum_{m=1}^{M_{(j)}} \{y_m Pr\left(\chi_m = 1 \,|\underline{\Gamma}_E\right)\}}{\sum_{m=1}^{M_{(j)}} \{Pr\left(\chi_m = 1 \,|\underline{\Gamma}_E\right)\}} \tag{15}$$

where $M_{(j)}$ indicates the number of connected cells where $Pr\left(\chi_m = 1 \,|\underline{\Gamma}_E\right) \neq 0$. Moreover, the extension of the estimated dangerous zone $\widetilde{\Delta}$ is defined as follows

$$\widetilde{\Delta} = \pi \left\{ \frac{\sum_{m=1}^{M_{(j)}} \left\{ \frac{\rho_m Pr(\chi_m=1|\underline{\Gamma}_E)}{max_m\{Pr(\chi_m=1|\underline{\Gamma}_E)\}} \right\}}{\sum_{m=1}^{M_{(j)}} \left\{ \frac{Pr(\chi_m=1|\underline{\Gamma}_E)}{max_m\{Pr(\chi_m=1|\underline{\Gamma}_E)\}} \right\}} \right\}^2 \tag{16}$$

where $\rho_m = \sqrt{\left(x_m - \widetilde{x}_{cil}\right)^2 - \left(y_m - \widetilde{y}_{cil}\right)^2}$.

Within the numerical validation, the first experiment deals with a *test set* of $P = 69$ patterns (related to examples different from those of the training phase and concerned with two- and three-scatterers configurations, $P^{(2)} = 35$ and $P^{(3)} = 34$) and noiseless conditions. Figs. 2 and 3 show the risk maps obtained for two examples of the test set. The first example (Fig. 2) refers to a two-targets configuration where the scatterers are located at $(x_{cil}^{(1)} = -\frac{5}{6}\lambda,\ y_{cil}^{(1)} = \frac{\lambda}{2})$ and $(x_{cil}^{(2)} = y_{cil}^{(2)} = \frac{5}{6}\lambda)$. In such a case, the values of the error figures turn out to be equal to $\varsigma^{(1)} = 0.291$ and $\varsigma^{(2)} = 0.389$. Moreover, the highest values of the occurrence probability are very close to the actual positions of the scatterers. As far as the dimensions of the two dangerous zones (the objects being not-adjacent) are concerned, they are slightly over-estimated ($\frac{\widetilde{\Delta}^{(1)}}{\Delta^{(1)}} = 2.023$ and $\frac{\widetilde{\Delta}^{(2)}}{\Delta^{(2)}} = 2.760$, $\Delta^{(1)} = \Delta^{(2)} = \frac{1}{9}\lambda^2$).

The second example (Fig. 3) is related to a three-scatterers configuration. The objects are adjacent and lie at the bottom of the investigation domain: $(x_{cil}^{(1)} = -\frac{\lambda}{6}$, $y_{cil}^{(1)} = -\frac{5}{6}\lambda)$, $(x_{cil}^{(2)} = \frac{\lambda}{6},\ y_{cil}^{(2)} = -\frac{5}{6}\lambda)$, and $(x_{cil}^{(3)} = \frac{\lambda}{2},\ y_{cil}^{(3)} = -\frac{5}{6}\lambda)$. As expected, when the targets are buried far from the surface, the localization of the "dangerous

zones" is more difficult. In spite of this, the approach is still able to localize these areas with an acceptable degree of accuracy ($\varsigma^{(1)} = 0.359$, $\varsigma^{(2)} = 0.385$, and $\varsigma^{(3)} = 0.502$).

For completeness, by considering the whole test-set, the statistics of the *dangerous-area-location error* ς are given in the first column of Tab. I.

The second numerical experiment considers a more critical scenario where a single target is supposed to be located in the investigation domain ($P^{(1)} = 36$). It should be pointed out that such a configuration does not belong to the training set. Concerning the effectiveness of the approach in dealing with this kind of test set, the localization statistics are reported in Tab. I. As expected, the error figure increases as compared to the first experiment and the average value changes from $av_p\left\{\varsigma^{(p)}\right\}\big|_{Exp.1} = 0.312$ to $av_p\left\{\varsigma^{(p)}\right\}\big|_{Exp.2} = 0.385$.

As an example, the risk map for a sample of the test set ($x_{cil}^{(1)} = y_{cil}^{(1)} = -\frac{\lambda}{6}$) is shown in Fig. 4 ($\varsigma^{(1)} = 0.364$ and $\frac{\tilde{\Delta}^{(1)}}{\Delta^{(1)}} = 4.718$).

Finally, to evaluate the robustness of the proposed approach, a noisy environment was considered. Corrupted measurement data were simulated by adding an uniform Gaussian noise to synthetic data [5] of the test set used in the first experiment. The risk maps for the two-targets configuration (where the actual scatterers are located as in Fig. 2) and for different signal-to-noise ratios ($SNRs$) are shown in Fig. 5 as representative examples. It can be observed that, when $SNR \geq 20\,dB$ the "contamined zones" are quite correctly detected and located ($\varsigma^{(1)}\big|_{SNR=35\,dB} = 0.383$, $\varsigma^{(2)}\big|_{SNR=35\,dB} = 0.462$ and $\varsigma^{(1)}\big|_{SNR=20\,dB} = 0.389$, $\varsigma^{(2)}\big|_{SNR=20\,dB} = 0.464$). Otherwise, the performance of the approach reduces ($\varsigma^{(1)}\big|_{SNR=10\,dB} = 0.526$, $\varsigma^{(1)}\big|_{SNR=10\,dB} =$

0.580) even though the highest values of the occurrence probability are situated in correspondence with and close to the target positions.

As far as the prediction of the extension of the dangerous zones is concerned, Tab. II shows that the effectiveness of the approach reduces with the increasing of the SNR. This causes an over-estimate of the dangerous areas.

Finally, for completeness, the results of an exhaustive analysis on the whole data set are reported in terms of statistical values of the localization error (Tab. I).

5 Conclusions

In this paper, a classification approach for the real-time sub-surface sensing of multiple buried targets has been proposed. A suitable SVM-based strategy has been developed for determining the probability of occurrence of buried targets and to define a "risk map" of the investigation domain.

The effectiveness of the approach has been preliminarily assessed by considering a two-dimensional geometry and noiseless as well as noisy conditions. The obtained results confirmed the ability of the method in detecting and locating multiple targets as well as in estimating the extension of the dangerous zones.

Future works, current under development, will be devoted to fully exploit the key-features of the approach and to deal with three-dimensional scenarios.

References

[1] "Mine facts," CD ROM-based encyclopedia of current world-wide mine technology, 1997 (available from OASD (SO/LIC) Acquisition, 2500 Defence Pentagon, Wash., D.C. 20301-2500).

[2] IEEE Trans. on Geoscience and Remote Sensing, Special Issue on: "*New Advances in Subsurface Sensing: Systems, Modeling and Signal Processing,*" vol. 39, no. 6, June 2001.

[3] S. Caorsi, D. Anguita, E. Bermani, A. Boni, M. Donelli, and A. Massa, "A comparative study of NN and SVM-based electromagnetic inverse scattering approaches to on-line detection of buried objects," *Journal of the Applied Computational Electromagnetics Society*, Special Issue on "*Neural Network Applications in Electromagnetics*," vol. 18, no. 2, pp. 1-11, 2003.

[4] I. T. Rekanos, "Inverse scattering of dielectric cylinders by using radial basis function neural networks," *Radio Science*, vol. 36, no. 5, pp. 841-849, 2001.

[5] E. Bermani, A. Boni, S. Caorsi, and A. Massa, "An innovative real-time technique for buried object detection" *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, no. 4, pp. 927-931, 2003.

[6] Ch. Christodoulou and M. Georgiopoulos. *Application of Neural Networks in Electromagnetics*, Artech House, Norwood, 2001.

[7] K. Morik, P. Brockhausen, and T. Joachims, "Combining statistical learning with a knowledge-based approach: a case study in intensive care monitoring", *Proc. of the 16th Int. Conf. on Machine Learning*, MIT Press., 1999.

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

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

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

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

[12] K. -R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Schölkopf, "An introduction to Kernel-Based Learning Algorithms," *IEEE Trans. on Neural Networks*, vol. 12, no. 2, pp. 181-201, March 2001.

[13] G. Wabba, "Support vector machines, reproducing kernel hilbert spaces and the randomized GACV," in *Advances in Kernel methods - Support Vector Learning*, B. Scholkopf, C. J. C. Burges, and A. J. Smola (Eds.), MIT Press, 1999.

[14] T. Hastie and R. Tibshirani, "Classification by pairwise coupling". Technical Report, Stanford University and University of Toronto, 1996 (http://www.stat.stanford.edu/trevor/Papers/2class.ps).

[15] J. Platt, "Probabilistic outputs for support vector machines and comparison to regularized likelihood methods," in *Advances in Large Marging Classifiers*, A. J. Smola, P. Bartlett, B. Scholkopf, D. Schuurmans (Eds.), MIT Press, 1999.

[16] D. Anguita, S. Ridella, F. Rivieccio, and R. Zunino, "Hyperparameter Design Criteria for Support Vector Machines," *Neurocomputing*, vol. 55, pp. 109-134, 2003.

[17] S. Caorsi and M. Raffetto, "Perfectly matched layer for the truncation of finite element mashes in layered half-space geometries and applications to electromagnetic scattering by buried objects," *Microwave Opt. Technol. Lett.*, vol. 19, pp. 427-434, 1998.

Figure Captions

- Figure 1 - Problem geometry.

- Figure 2 - Noiseless data. Risk map for the two-targets scenario ($(x_{cil}^{(1)} = -\frac{5}{6}\,\lambda$, $y_{cil}^{(1)} = \frac{\lambda}{2}$) and ($x_{cil}^{(2)} = y_{cil}^{(2)} = \frac{5}{6}\,\lambda$)).

- Figure 3 - Noiseless data. Risk map for the three-targets scenario ($(x_{cil}^{(1)} = -\frac{\lambda}{6}$, $y_{cil}^{(1)} = -\frac{5}{6}\,\lambda$), ($x_{cil}^{(2)} = \frac{\lambda}{6}$, $y_{cil}^{(2)} = -\frac{5}{6}\,\lambda$), and ($x_{cil}^{(3)} = \frac{\lambda}{2}$, $y_{cil}^{(3)} = -\frac{5}{6}\,\lambda$)).

- Figure 4 - Noiseless data. Risk map for the single-target scenario ($x_{cil}^{(1)} = y_{cil}^{(1)} = -\frac{\lambda}{6}$).

- Figure 5 - Noisy data. Risk maps for the two-targets scenario ($(x_{cil}^{(1)} = -\frac{5}{6}\,\lambda$, $y_{cil}^{(1)} = \frac{\lambda}{2}$) and ($x_{cil}^{(2)} = y_{cil}^{(2)} = \frac{5}{6}\,\lambda$)) when (a) $SNR = 35\,dB$, (b) $SNR = 20\,dB$, and (c) $SNR = 10\,dB$.

Table Captions

- Table I - Numerical Experiment No. 1: Statistics of the dangerous-area-localization error, ς.

- Table II - Numerical Experiment No. 3: Estimate of the error in evaluating the area of the dangerous zone.

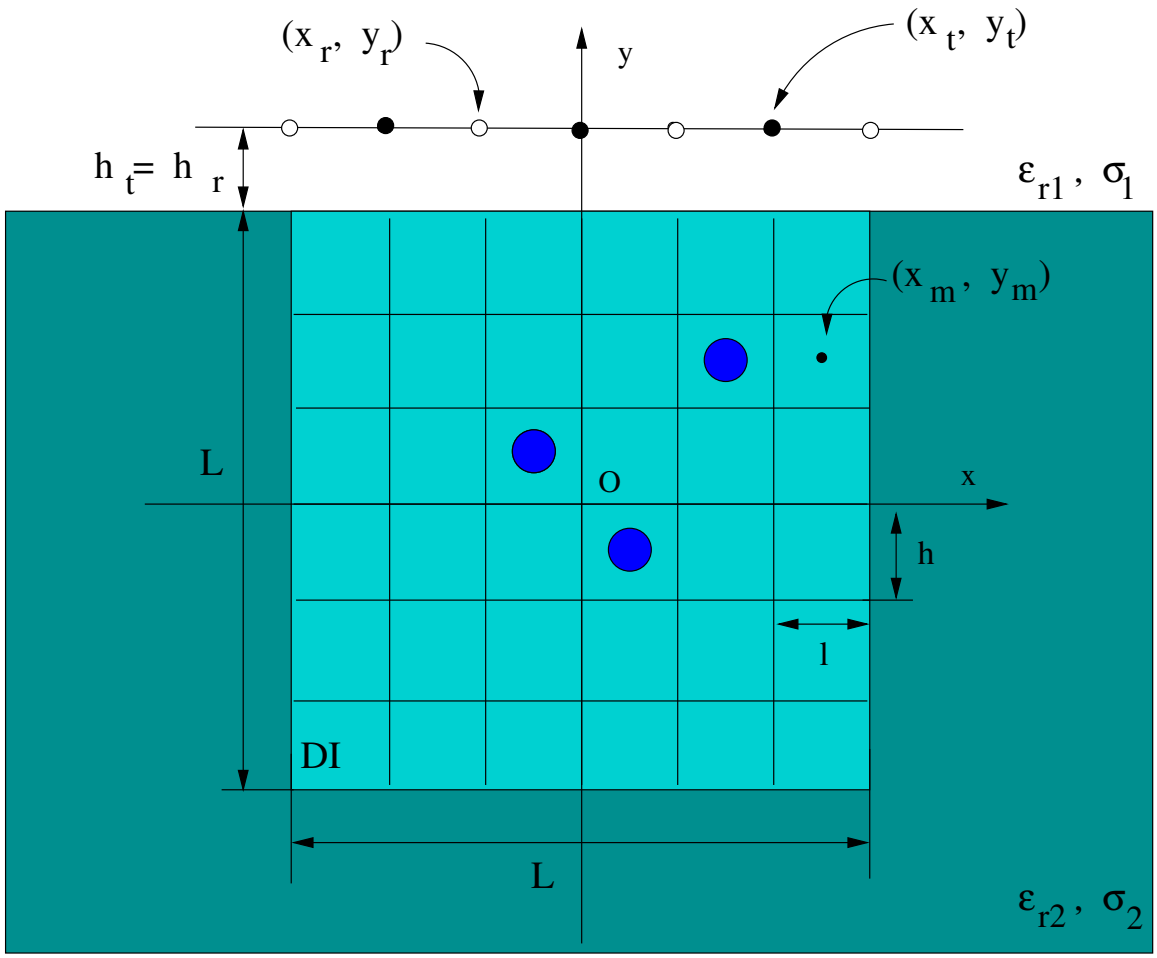


Figure 1 - A. Massa *et al.*, "A Classification Approach ..."

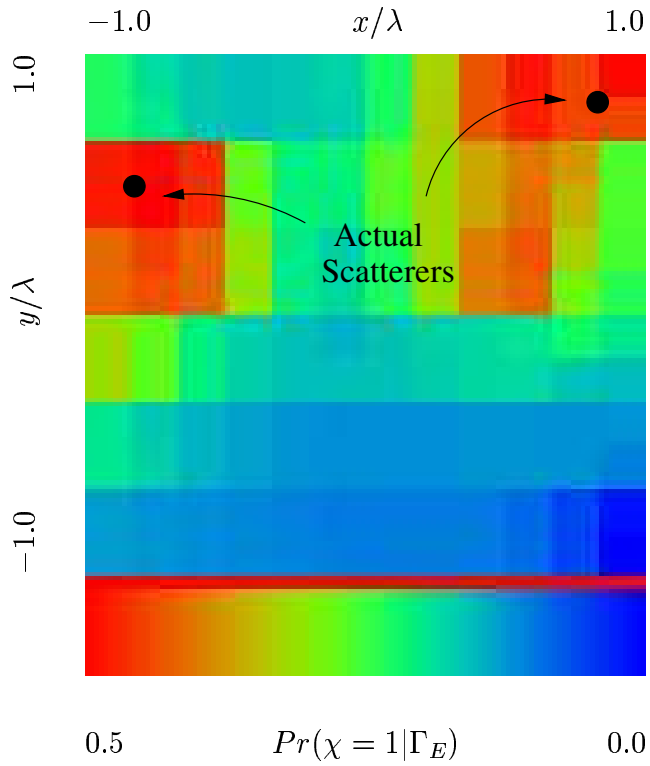


Figure 2 - A. Massa *et al.*, "A Classification Approach ..."

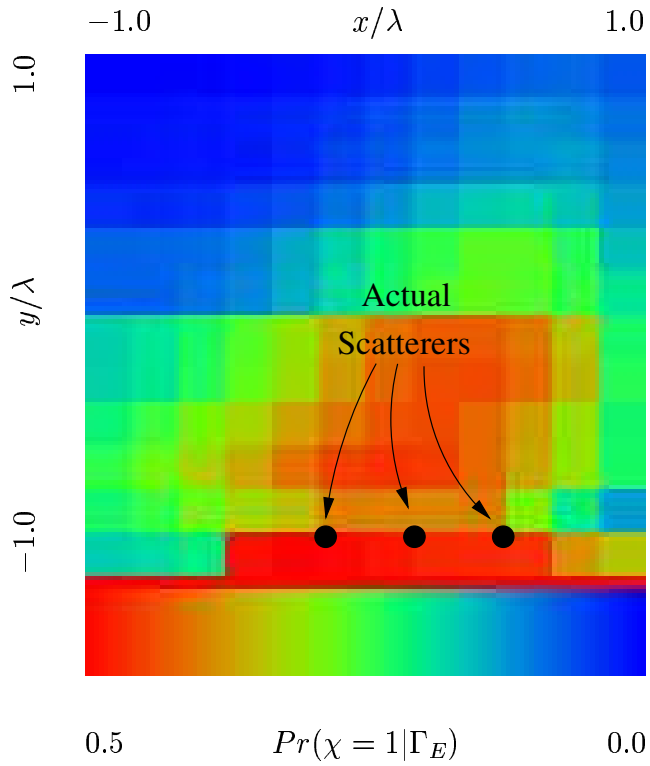


Figure 3 - A. Massa *et al.*, "A Classification Approach ..."

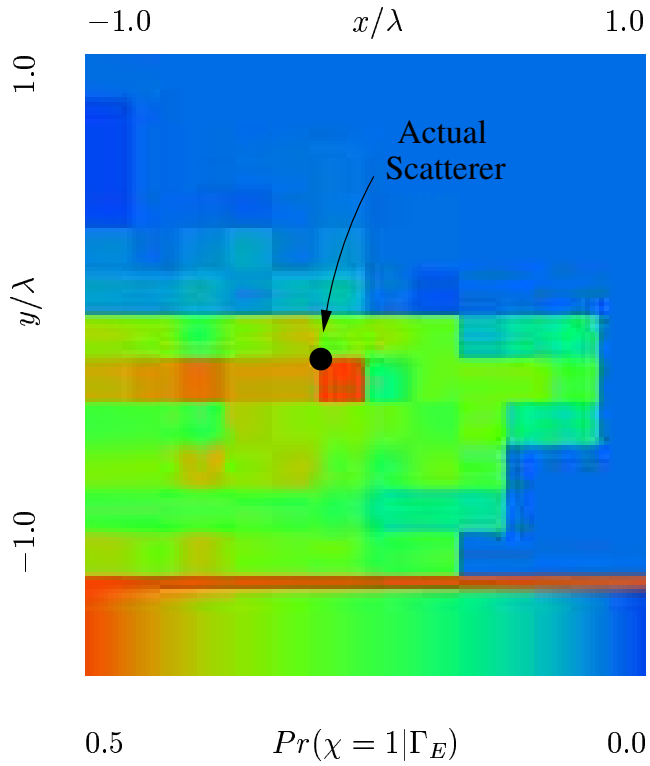
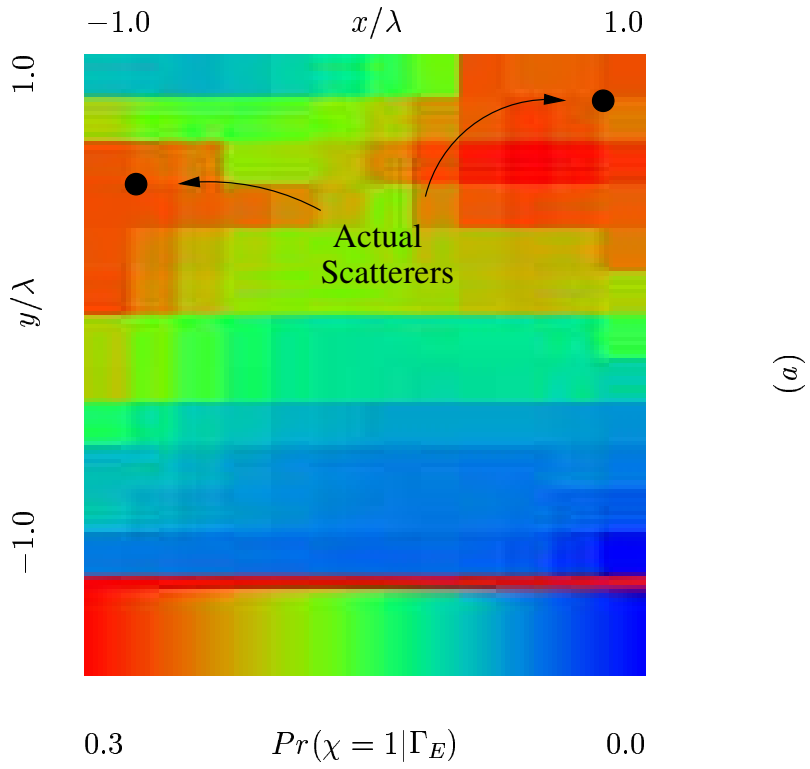


Figure 4 - A. Massa *et al.*, "A Classification Approach ..."



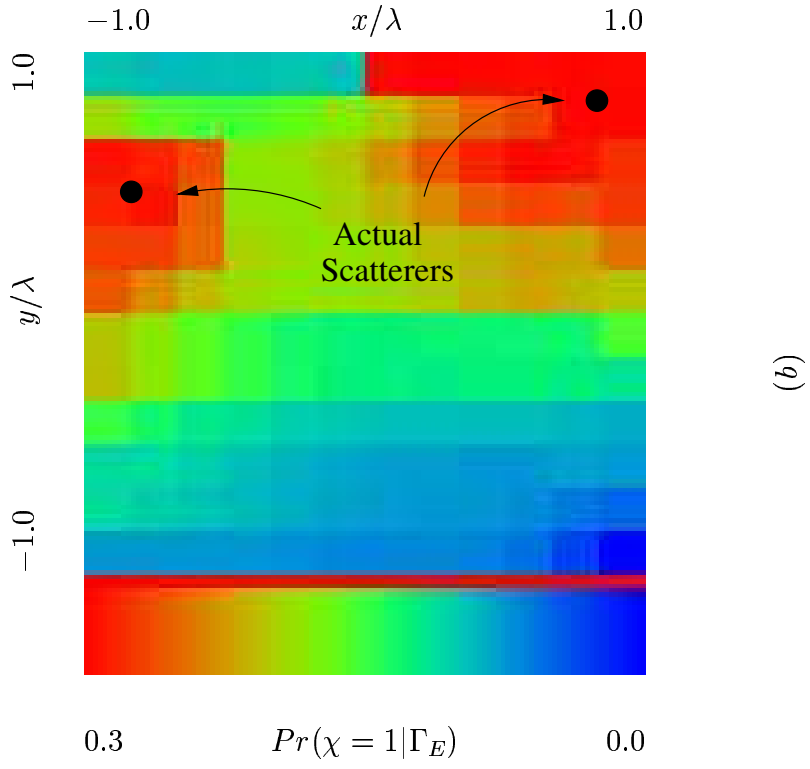


Figure 5(I) - A. Massa *et al.*, "A Classification Approach ..."

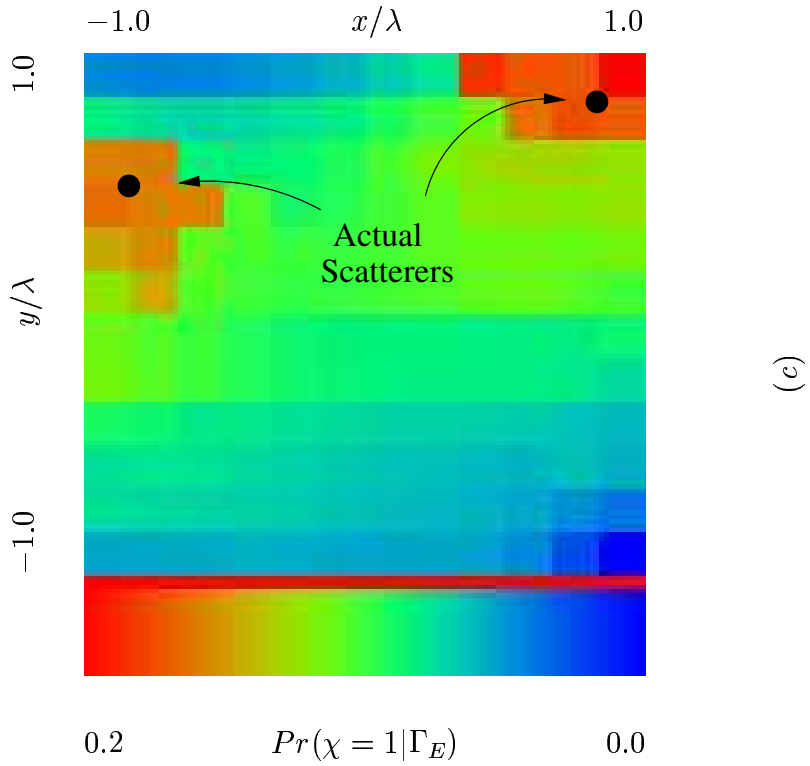


Figure 5(*II*) - A. Massa *et al.*, "A Classification Approach ..."

$Experiment$	$No.1$	$No.2$	$No.3$		
$SNR\ [dB]$			20	10	5
$min_p \left\{ \varsigma^{(p)} \right\}$	0.065	0.151	0.139	0.193	0.053
$max_p \left\{ \varsigma^{(p)} \right\}$	1.024	1.046	0.995	0.992	0.936
$av_p \left\{ \varsigma^{(p)} \right\}$	0.312	0.385	0.347	0.358	0.378
$var_p \left\{ \varsigma^{(p)} \right\}$	0.077	0.098	0.074	0.075	0.073

Table I - A. Massa *et al.*, "A Classification Approach ..."

$SNR\ [dB]$	∞	35	20	10
$\frac{\widetilde{\Delta}^{(1)}}{\Delta^{(1)}}$	2.023	3.230	3.249	4.456
$\frac{\widetilde{\Delta}^{(2)}}{\Delta^{(2)}}$	2.760	3.249	3.278	4.752

Table II - A. Massa *et al.*, "A Classification Approach ..."