Talanta 110 (2013) 108–117

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Recognition of explosives fingerprints on objects for courier services using machine learning methods and laser-induced breakdown spectroscopy

J. Moros [a], J. Serrano [a], F.J. Gallego [b], J. Macías [b], J.J. Laserna [a,*]

[a] Department of Analytical Chemistry, University of Malaga, E-29071 Malaga, Spain
[b] Department of Mathematical Analysis, University of Malaga, E-29071 Malaga, Spain

ARTICLE INFO

Article history:
Received 7 November 2012
Received in revised form
1 February 2013
Accepted 11 February 2013
Available online 19 February 2013

Keywords:
LIBS
Fingerprints
Home-made explosives
Harmless
Machine learning
Decision tree

ABSTRACT

During recent years laser-induced breakdown spectroscopy (LIBS) has been considered one of the techniques with larger ability for trace detection of explosives. However, despite of the high sensitivity exhibited for this application, LIBS suffers from a limited selectivity due to difficulties in assigning the molecular origin of the spectral emissions observed. This circumstance makes the recognition of fingerprints a latent challenging problem. In the present manuscript the sorting of six explosives (*chloratite*, *ammonal*, *DNT*, *TNT*, *RDX* and *PETN*) against a broad list of potential harmless interferents (butter, fuel oil, hand cream, olive oil, …), all of them in the form of fingerprints deposited on the surfaces of objects for courier services, has been carried out. When LIBS information is processed through a multi-stage architecture algorithm built from a suitable combination of 3 learning classifiers, an unknown fingerprint may be labeled into a particular class. Neural network classifiers trained by the *Levenberg–Marquardt* rule were decided within *3D* scatter plots projected onto the subspace of the most useful features extracted from the LIBS spectra. Experimental results demonstrate that the presented algorithm sorts fingerprints according to their hazardous character, although its spectral information is virtually identical in appearance, with rates of false negatives and false positives not beyond of 10%. These reported achievements mean a step forward in the technology readiness level of LIBS for this complex application related to defense, homeland security and force protection.

1. Introduction

In recent times, laser-induced breakdown spectroscopy (LIBS) has been demonstrated to be a powerful analytical tool to cope the direct chemical detection of energetic materials and residues of explosives in real-time [1–3]. However, despite all the efforts and the gains achieved in recent years when using LIBS for this last topic [4,5], a particularly challenging problem still needs to be overcome for offering greater assurances in this matter. From the identification point of view, the recognition aspects of LIBS are considered one of its *Achilles' heels*, because compounds sharing a similar chemical composition also have virtually the same LIBS signature. Difficulties on discrimination issues are increasing when residues and also the support where they are left are from the same nature. The situation is further complicated due to the fact that for heterogeneous residues their particular emissions and their characteristic intensity ratios, which are usually used in classification of explosives by LIBS, may fluctuate in a random manner. Finally, to round off the problem, the varying influence of the surrounding atmosphere may also contribute to these cited fluctuations.

For all these reasons, a broad range of powerful chemometric tools have been applied for dealing with this issue of assigning an identity to a residue, or at least, for its sorting into as hazardous or harmless. Table 1 summarizes the state-of-the-art of statistical and mathematical tools applied to LIBS information in order to improve its discrimination capability. The most relevant aspects involved in the successful application of these multivariate analysis techniques, namely, the evaluated materials (not only explosives but also potential interferents), the supports where the residues are left, the way used to prepare the experimental targets, and supplied quantities, are provided. In any case, readers interested in detailed information are invited to consult the references cited.

Among the multivariate techniques demonstrated to be viable to classify an unknown sample as an explosive or a harmless product, the most widely used has been the principal component analysis (*PCA*) [6–9]. This technique works well when the variability within a group is much smaller than the variability among groups. However, this circumstance is unlikely when account is taken of the heterogeneous deposition of a fingerprint and the large *shot-to-shot* variability of LIBS spectra. For these reasons, the deposition of discrete volumes of acetone solutions and the

* Corresponding author. Tel.: +34 952 13 1881; fax: +34 952 13 2000.
E-mail address: laserna@uma.es (J.J. Laserna).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.02.026

Table 1
State-of-the-art in the combination of LIBS with chemometric tools for the detection of explosive residues.

Explosives	Potential interferents	Support/s	Sample preparation/form	Modeling approach	Ref.
• RDX (Cyclotrimethylenetrinitramine) • Composition-B (36% TNT, 63% RDX, and 1% wax)	• Diesel fuel	Aluminum foil	• Depositing of a discrete volume of solutions (powdered explosives in acetone) that then is allowed to evaporate. • Spreading of diesel fuel on the support surface as a thin film.	Principal component analysis (PCA)	[6]
• RDX • TNT (Trinitrotoluene) • Composition-B	• Arizona road dust • Commercial lubricant (WD-40) • Oil from the surface of the skin in a fingerprint pattern	Aluminum foil	• Depositing of a discrete volume of solutions (powdered explosives in acetone) that then is allowed to evaporate ($\sim$500 ng/cm^2). • Spreading of interferents on the support surface as a thin layer/film (<2 μg/cm^2).	PCA	[7]
• RDX • TNT (Trinitrotoluene) • Composition-B	• Arizona road dust • Commercial lubricant (WD-40) • Oil from the surface of the skin in a fingerprint pattern	Aluminum foil	• Depositing of a discrete volume of solutions (powdered explosives in acetone) that then is allowed to evaporate ($\sim$500 ng/cm^2). • Spreading of interferents on the support surface as a thin layer/film (<2 μg/cm^2)	• Linear correlation • PCA • Soft independent method of class analogy (SIMCA) • Partial least squares–discriminant analysis (PLS–DA)	[8]
• EGDN (Ethylene glycol dinitrate) • NG (Nitro glycol) • RDX • TNT • DNT (Dinitrotoluene) • PETN (Pentaerythritol tetranitrate) • HMX (Cyclotetramethylene tetranitramine) • TETRYL (2,4,6-Trinitrophenylmethylnitramine) • TATP (Triacetone triperoxide)	• Diesel oil • Paraffin wax • Grease lubricant • Coal NIST 1632b • Glue LOCTITE • Hand cream	Aluminum discs	• Spreading, over support surface, of small droplets of commercial solutions of explosives (0.1–1.0 mg/ml in methyl or ethyl alcohol, or acetonitrile) that then are allowed to evaporate thermally, leaving unevenly distributed residues. • Spreading of interferents on the support surface as thin layers/films of uncontrolled thickness.	PCA	[9]
• RDX • TNT • Composition-B	• Arizona road dust • Commercial lubricant (WD-40) • Oil from the surface of the skin in a fingerprint pattern	Aluminum foil	• Depositing of a discrete volume (10 μL) of solutions ($\sim$1 mg/mL powdered explosives in acetone) that then is allowed to evaporate. Crushing and spreading of small amounts (<1 mg) of RDX powder with a Teflon block. Direct transferring of RDX powder from the fingertips to the support. • Spreading of interferents on the support surface as a thin film. Printing of grease from skin by handling repeatedly the surface of the support.	PLS–DA	[10]
RDX	• Arizona road dust • Commercial lubricant (WD-40)	• White rubber (drink coaster) • Red silicone (pot-holder) • Blue plastic (floppy disk case)	Spreading of several milligrams of the dry, powdered explosive on the surfaces of the different supports with a Teflon block,	PLS–DA	[11]

Table 1 (*continued*)

Explosives	Potential interferents	Support/s	Sample preparation/form	Modeling approach	Ref.
		• Aluminum foil • Wood (lumber) • Cardboard (box) • Flooring samples • Tan silicone (pot-holder) • Steel	resulting in a thin layer ($\sim$100 μg/cm^2).		
• RDX • TNT • Composition-B	• Arizona road dust • Sand • Diesel fuel • Lubricant oil • Fingerprint oil • Blank (no residue) car panels	Car panels	• Crushing and smearing of dry, powdered explosives ($\sim$1–2 mg) on the support surface (coverage of $\sim$10 μg/cm^2). • Arizona road dust and sand were added in a similar manner as the explosives. Spreading of diesel fuel and lubricant oil on the surface as thin layers. Printing of grease from skin by handling repeatedly the surface of the support.	PLS–DA	[12]
• DNT • TNT • C4 or Composition-C (plastic explosive, 91% RDX) • H15 (plastic explosive, 75% RDX) • $NaClO_3$	–	• Aluminum foil • Polymethylmethacrylate (PMMA) foil • Microscope slide	Pressing the solid explosives on the support surface with a mortar handle ($\sim$500 ng/mm^2).	The used algorithm involves three steps: (1) Application of the method of normalized coordinates (MNC), (2) comparison of selected coordinate values with fixed thresholds and, (3) decision making based on logical conjunction.	[13]
• EGDN • NG • RDX • TNT • DNT • PETN • HMX • TETRYL • TATP	• Diesel oil • Paraffin wax • Hand cream	Aluminum discs	• Spreading, over support surface, of small droplets of commercial solutions of explosives (0.1–1.0 mg/ml in methyl or ethyl alcohol, or acetonitrile) that then are allowed to evaporate thermally, leaving unevenly distributed residues. • Spreading of interferents on the support surface as thin layers/films of uncontrolled thickness.	Simple procedure described in following: (1) Acquisition of a sufficient number of single-shot spectra, (2) Normalization of line intensities on ionic emission in all the individual spectra (3) Linear fitting of the normalized functions and retrieving of the relative slopes and intercepts (4) Verifying if one or more of the calculated slopes and intercepts fall inside the limits for the explosives.	[14]
–	• 2-Cloroethyl ethyl sulfide (CEES) • Dolomitic limestone • Ovalbumin • Chloroform	• Aluminum • Polycarbonate • Steel	–	MLPCA (Maximum likelihood principal components analysis)	[15]

spreading of thin films, leaving all residues evenly distributed throughout the support (in all cases aluminum), is the most common way of preparing the experimental samples. Similarly, linear correlation, that bases on the matching of an unknown spectrum to a previously established library of representative spectra, and the soft independent modeling of class analogy (*SIMCA*), a well known supervised classification method which describes each class separately, have been considered to cope with this issue under the above set of boundary conditions [8].

On the other hand, when the target is heterogeneous and each single LIBS event may consume varying amounts of material present on the same surface, partial least squares–discriminant analysis (*PLS–DA*) [8,10–12] seems the best alternative method used for these classification and discrimination problems. The so-called normalized coordinates method, based on linear algebra, allows to detect the existence of a mixture of several substances in a sample, even when the material is placed behind a barrier [13].

Another procedure [14] to discriminate explosives from interferents through the intercept and slope, respectively, for different, linearly fitted functions of different atomic lines intensities after their normalization to the most sensitive emission for the support, has been described.

Last but not less effective, another alternative method namely maximum likelihood principal component analysis (*MLPCA*) [15], which is an errors-in-variables modeling method that accounts for measurement errors in the estimation of model parameters, has been successfully applied to modeling detection of explosives residues by LIBS. This approach explicitly model LIBS spectra resulting from the residue and the support to attempt to separate the response from each of the two components. This separation process is performed jointly with classifier design to ensure that developed model is able to detect residues of interest without being confused by variations in the substrates.

Nonetheless, although these conventional chemometric techniques have proven to be able to solve these issues in a more or less efficient way, the problem in the recognition of residues by LIBS remains latent today. Developing models that can correctly characterize the complex emission patterns of any material located on a particular surface is critically important.

In the present paper, authors propose a new alternative for fingerprints recognition by LIBS combined with supervised learning methods [16]. This approach focuses on the learning of the emission optical behavior (spectral features intensities) coming from externally supplied instances (*training* dataset) for searching, designing and training classifiers which argue a distribution of classes (hazardous and harmless fingerprints, in the matter at hand). Once assessed by using a *testing* dataset, the resulting classifier is ready to use on sorting future unknowns [17–20]. Bearing in mind the main demands of any classification task, namely, the correct sorting of as much of instances as possible, the capability of updating itself as many times as necessary and be the most straightforward and easy-to-interpret, and the broad range of products assayed here, authors decided to implement a decision tree classifier [21,22]. This approach is an inductive algorithm more interpretable than other common chemometric strategies because they combine simple issues from data in an understandable way. Furthermore, decision tree approach has substantial advantages for identification problems because of its flexibility and ability to handle non-linear relations between features and classes, thus improving the recognition accuracy to a great extent. To the best of our knowledge, it is the first time that this strategy applies to the matter being discussed, so that this research represents the latest advance in the studies to evaluate the aptitude of a LIBS sensor for classification of fingerprints that are left on objects of courier services.

2. Experimental

2.1. Sensor set-up

The experimental configuration being used is schematized in Fig. 1. A Quantel Brilliant Ultra Q-Switched Nd:YAG laser (20 Hz, @1064 nm, 8 ns pulse width) was used as a excitation source. Laser beam was focused using an optical system integrated by a 1064 nm plane mirror and a plane-convex quartz lens of a focal length of 75 mm. For sampling the targets and generating of their plasma plumes, laser energy of ca. 41 mJ per pulse was focused on a spot size of 470 μm in diameter (irradiance value of 3 GW cm^{-2}).

Plasma light was collected through a collimating lens coupled to the tip of a 600 μm optical fiber (four-furcated cable, 4×600 μm fibers, all legs SMA terminated, total 2 m long, splitting point in the middle) which guided the light to the entrance of a four-channel miniature Czerny-Turner spectrograph (75 mm focal length). Spectrometers were each one fitted with CCD detectors. LIBS spectra that span from 230 nm to 950 nm were achieved. A delay time of 1.28 μs and an integration time of 1.1 ms, were used as timing parameters for spectral data acquisition.

2.2. Samples

To assess the potential of LIBS in fingerprints detection and their discrimination when left on the surfaces of objects for

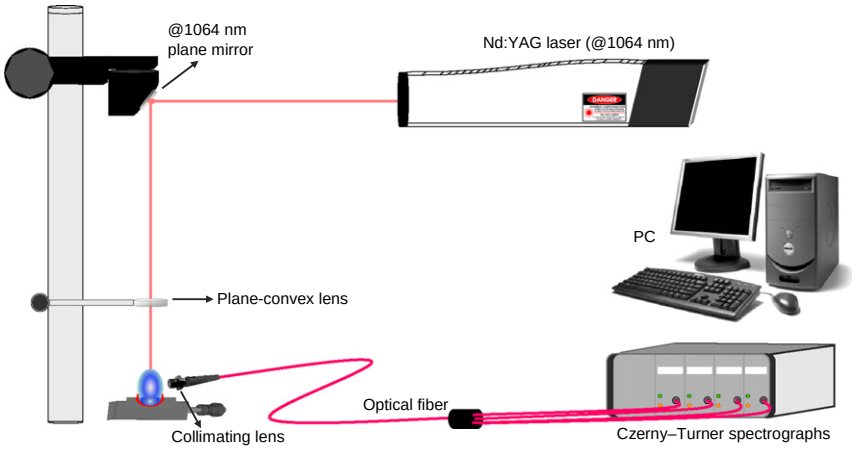

Fig. 1. LIBS experimental set-up.

courier services, a wide range of materials were tested. Four energetic organic compounds, including *DNT* (2,6-dinitrotoluene), *TNT* (2,4,6-dinitrotoluene), *RDX* (trinitramine cyclotrimethylene) and *PETN* (pentaerythritol tetranitrate) (these latter two provided by the Laboratorio Químico Central de Armamento –*LQCA*–, Ministry of Defense, Madrid, Spain), and a pair of powerful *home-made* explosives namely *chloratite* (mainly composed of sulfur –S–, sodium and/or potassium chlorates –$NaClO_3$, $KClO_3$–, and sugar) and *ammonal* (product based on ammonium nitrate –NH_4NO_3– and aluminum –Al– powder), were assayed as positive declarations.

At the same time, seventeen common harmless products that shared some basic elements with those hazardous materials, such as butter, car wax, fuel, fuel oil, hand cream, olive oil, two types of motor oil, soap, sugar, sweetener, KCl, $KClO_3$, NaCl, $NaClO_3$, NH_4NO_3 and sulfur, were taken into account.

In general, objects commonly used for primary packaging are composed of cellulose. Nonetheless, despite that all them are based on wood pulp, each one of these pieces is processed differently. For these reasons, three items, namely a cardboard box (material consisting of several superimposed layers of paper, based on virgin fiber or recycled paper), a padded envelope of brown paper (composed of a *kraft* paper outer envelope with self-adhesive closure and bubble film bag inside) and a conventional envelope of white paper, were chosen as supports for sampling.

Experimental targets were prepared by depositing a fingerprint of any finger from two laboratory operators. Five independent fingerprints of each material were left on the surface of 5 independent fragments of each one of the packaging items. It is important to note that no quantification was attempted during experimentation.

2.3. LIBS data

A total amount of 5040 single-shot LIBS spectra to cope with all possible residue/support combinations were collected. The complete LIBS information was comprised by two datasets (*training* and *testing*, respectively) that matched but are independent. Each dataset contained 2520 LIBS spectra. Information within each dataset is broken down as follows: 35 single-shot analysis from 5 independent fingerprints (7 for each) for the total 69 possible combinations residue/support (23 residues on 3 different supports), plus 105 single-shot analysis from 3 clean -*lacking a forced residue*- supports (35 for each). Sampling was performed at discrete, refreshed and well-spaced locations within the particular area of each fingerprint. All data obtained were exported in appropriate format and processed using *PRTools* from Matlab (The Mathworks Inc., South Natick, MA).

2.4. Decision tree classifier

One of the greatest advantages of elemental analysis techniques involves the ability to distinguish between materials exhibiting significant differences in their chemical composition. Regrettably, these techniques show evident difficulties on identification tasks when materials share similar atomic composition. This circumstance requires the use of a systematic approach to design a classification model for a complex input dataset, when a wide diversity of materials is considered. In order to cope with this matter, authors decided the development of a decision tree classifier as learning algorithm.

Decision trees follow a hierarchical structure where at each level a test is applied to one or more attribute values that may have one or two outcomes. Classification of an object starts at the root of the tree. The object is then evaluated at one node and takes the branch appropriate to its outcome. The process may continue hierarchically through several internal nodes, till it encounters a final leaf at which it is asserted that the object belongs to a particular class. In summary, each final leaf is the result of following down at the tree through a set of mutually exclusive decision rules.

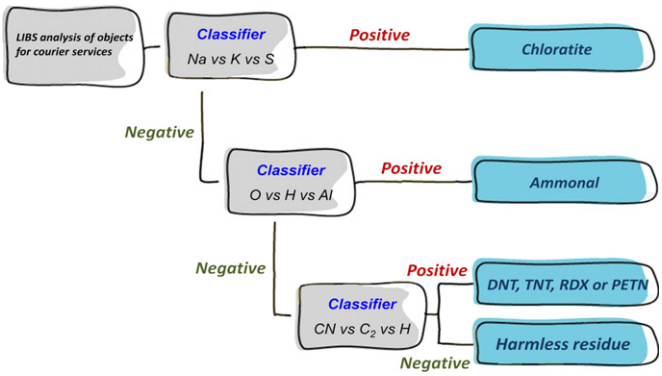

Fig. 2. Scheme of the designed decision tree classifier.

In the present work, a decision tree consisting of three internal nodes has been designed. In this process of multi-stage decision making, each internal node was assigned to a particular group of explosives: (I) *chloratite*, (II) *ammonal* and (III) *organics*. An exclusive machine learning classifier was chosen as a decision strategy to be used at each internal node. In these instances, since LIBS spectra of compounds sharing a similar chemical composition have virtually the same signature, the only useful available spectral information was limited to intensity values of matching features. Consequently, the choice of feature subsets to be used at each internal node (*classifier*) rested on the extraction of such intensity values.

Hence, emission peak intensities extracted were graphed in the best scatter plot at which the classifier that finest fits the relationship between the attributes set and the class of the input data was designed. Variables extracted from spectra of the *training* dataset were firstly used for designing and training exclusive classifiers for the sorting of residues on each single support (also called "*single classifiers*"). Furthermore, to assess whether a unique classifier is able to discriminate between hazardous and harmless fingerprints whatever the packing item assayed, some classifiers were also built by considering the three supports as a whole ("*extended classifiers*"). Finally, to complete the evaluation on labeling of fingerprints for the future sorting of unknowns, all classifiers (either single or extended) were assessed by identifying fingerprints from *testing* dataset. The best classifier on the basis of several performance parameters, namely accuracy (its overall correctness), sensitivity (its ability to correctly identify positive declarations), specificity (its ability to correctly identify negative findings that do not raise an alarm) and precision (accuracy on the discrimination of both classes), was decided for each internal node. Despite that a large variety of different machine learning classifiers were first trained and then evaluated, a neural network classifier trained by the *Levenberg–Marquardt* rule was decided for each one of the three internal nodes [23,24].

Finally, it must be pointed out that the effectiveness of the classification may vary depending on the final configuration of the internal nodes within the decision tree. From the six possible alternatives, the best sequence of actuations was planned and organized on the basis of how the decisions taken at the earlier stages of the sequence and the outcomes from first internal node affect, to a lesser extent, the decisions and outcomes at the later stages of that sequence. After evaluation of their performance parameters, a configuration following the programmed sequence *chloratite*:*ammonal*:*organics* was decided as the most efficient tree structure. Fig. 2 shows the scheme of the designed decision tree.

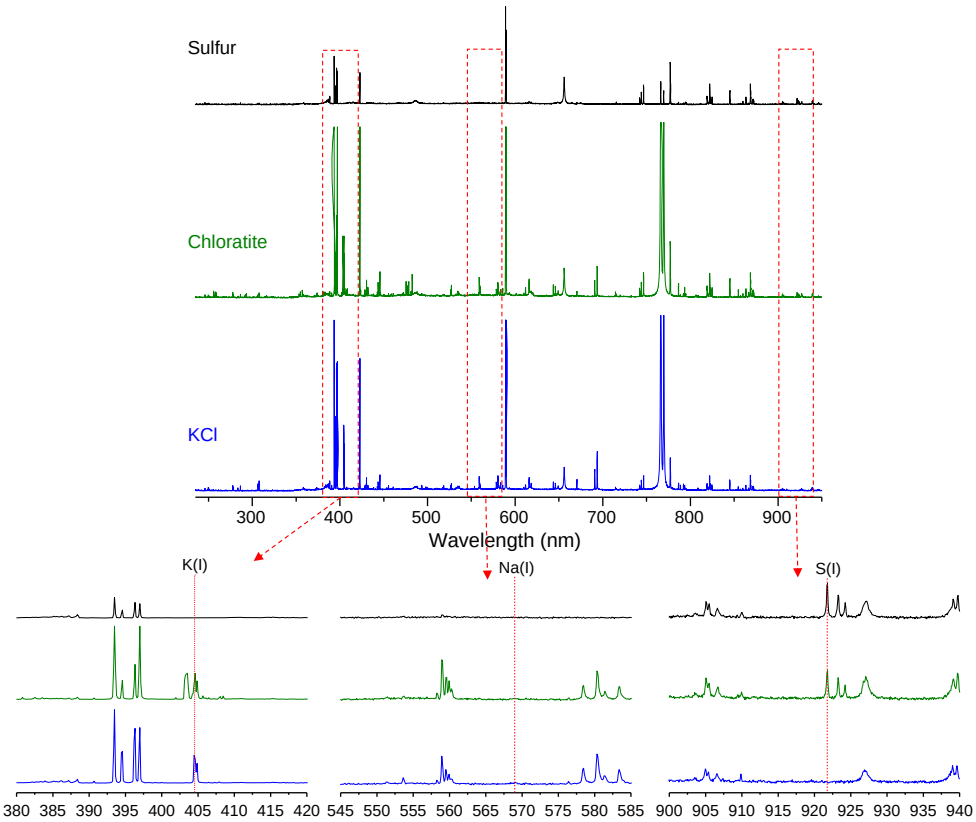

Fig. 3. LIBS spectra gathered from pure sulfur (*black*), *chloratite* (*green*) and KCl (*blue*) fingerprints left on the surface of a padded envelope. The insets show, in detail, the particular spectral windows, where the emissions (K, Na and S) involving discrimination are manifested. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Results and discussion

3.1. LIBS spectra of combinations residue/support

LIBS analysis of experimental targets (residue/support combinations) provides a high-resolution spectrum spanning from 230 nm up to 950 nm. Furthermore, all three supports assayed, when free from any fingerprint, exhibited complex LIBS spectra; highly rich in emission lines. However, no attempts were made to identify all lines observed within LIBS spectra from supports; exception given for those attributes directly related the residues considered, and subsequently, involved in classifying them.

As an example, Fig. 3 shows the LIBS spectra gathered from fingerprints of *chloratite* as well as of their main components – $KClO_3$ and *sulfur*– independently, all of them deposited on the surface of a padded envelope.

As observed, both chloratite and $KClO_3$ spectra show characteristic line of K(I) at 404 nm, whereas this is not manifested in the sulfur spectrum (inset at bottom-left in Fig. 3). In contrast, despite that the most sensitive emissions of Na(I) at ca. 589 nm are detected in all the cases, neither of three situations shows the emission from Na at 569 nm (inset at bottom-center in Fig. 3). This circumstance indicates that, in the matter at hand, the particular presence of the 589 nm lines is attributed either to any eventual contamination bearing no relation with the fingerprints, or to an emission originating from the support. Finally, while both chloratite and sulfur spectra present characteristic emission signal of S(I) at 921 nm, this feature is not evidenced in the $KClO_3$ spectrum (inset at bottom-right in Fig. 3).

Hence, as deduced, the presence of, at least, two of these signals is a mandatory requirement for a positive declaration on the presence of the explosive; the mere occurrence of only one of them is not sufficient. This statement holds also true for *chloratite* made up of $NaClO_3$.

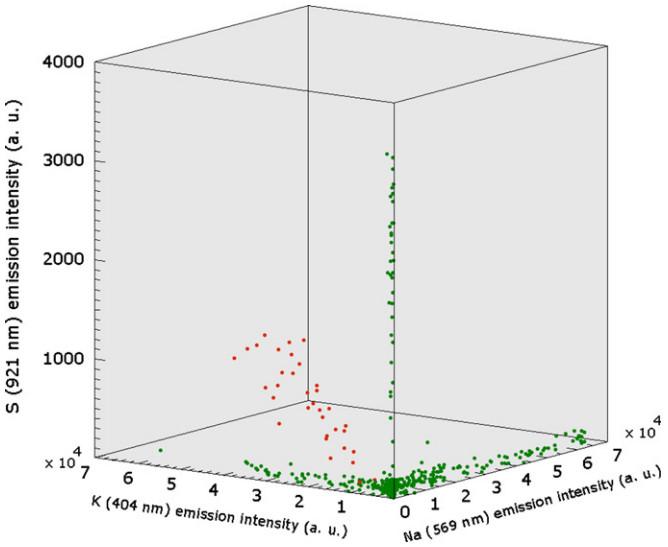

Fig. 4. *3D* scatter plot projected onto the subspace of the K (404 nm), Na (569 nm) and S (921 nm) signals for *chloratite* (*red*) and harmless (*green*) fingerprints left on the surface of on the surface of a cardboard box. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In the case of *ammonal* (product based on NH_4NO_3 and Al powder) only four emission signals, namely Al(I) at 396 nm, H(I) at 656 nm, N(I) at 746 nm, and O(I) centered at 777 nm, are of substantial value for its positive diagnostic due to its direct linking with the material. Regrettably, some minor setbacks may disturb

their usefulness on sorting ammonal. For instance, it is currently not possible to use the most sensitive line of aluminum (396 nm) since this element is present in the three supports assayed (data not shown). This forces to employ the emission at 309 nm that goes unrecognized for clean supports whereas manifests when *ammonal* residue is present. On the other hand, contributions from the surrounding air and the environmental relative humidity may also distort emissions of H, N, and O in an unpredictable manner, depending on the changing conditions of the plasma plume. This last distortion also influences the sorting of fingerprints from organic materials since, for the vast majority of these compounds, H, N and O emissions are the only ones accompanying C, CN and C_2 signals as attributable features to such materials.

In short, all this significantly limits the classification potential of LIBS; the analytical challenge is heightened.

3.2. Identifying fingerprints from their LIBS spectra

As mentioned, particular classifiers were chosen as a decision strategy to be used at each internal node of the decision tree. Thus, for each one of groups of explosives considered, sorting was optimized in terms of both emission signals and machine learning classifiers. Best results achieved on differentiating fingerprints of *chloratite*, *ammonal*, and organic explosives from those of some harmless common products are summarized in the following sections.

3.3. Chloratite fingerprint identification

As above mentioned, *chloratite* is a mixture of sodium and/or potassium chlorates and pure sulfur. Hence, the best spectral features extracted for classification purposes were emission lines associated to atomic Na (569 nm), K (404 nm) and S (921 nm). As discussed above, less sensitive emissions for Na and K were chosen in order to distinguish between the fingerprint and a mere contamination of the support. Fig. 4 illustrates the *3D* scatter plot showing the LIBS intensities for *chloratite* fingerprints (red dots) compared to those related to all common harmless products (green dots), when left on the surface of a cardboard box. From this diagram, it is clear that a simple decision boundary for the neural network classifier trained by the *Levenberg–Marquardt* rule permits a straightforward distinction of chloratite traces from those harmless fingerprints; even fingerprints of compounds that partially share the elemental composition with chloratite, namely chlorides, chlorates and pure sulfur, are separated right away.

Table 2 summarizes the performance parameters of this particular classifier when used to identify fingerprints from both *training* and *testing* datasets. From the reported results, a comparison between the performance on the use of either a simple classifier for each support or an extended one for all of them together, may also be made.

As seen, all developed classifiers were highly accurate and precise, as evidenced by their values close to 100% for both parameters. Furthermore, in all circumstances, the specificity of the model was found to be 100%, indicating that none of the harmless fingerprints was misidentified as *chloratite* by classifiers. The same is true for the sensitivity since only one of five fingerprints of *chloratite* placed on the cardboard box was missed by the model. This means that, for the scanning of the cited fingerprint, none of the 7 single-shot analyses identified the residue as *chloratite*.

Low surface concentrations of *chloratite* in the support as well as low sulfur content in *chloratite* formulations both may increase the mislabeling rates of this explosive by the classifier. Clearly, this error cannot be attributed to a malfunction of the model but to the inherent detection power of LIBS. Also, a number of additional factors beyond the control of the operator may influence the detection of a fingerprint. These factors include the force applied during deposition, duration of surface contact, the angle of deposition, and the level of adhesion of the fingerprint to the support (directly related to their physicochemical properties) [25,26].

The data in Table 2 also demonstrate that discrimination between *chloratite* and harmless fingerprints is successful using a unique classifier whatever the packing object assayed. As observed, no difference was detected in the performance of the extended classifier when tested on the residues placed on any support with respect to the performance of each simple classifier for sorting their particular fingerprints.

3.4. Ammonal fingerprint identification

In accordance with the above approach, a particular classifier, also trained by the *Levenberg–Marquardt* rule, was built for identifying *ammonal* fingerprints. In this case, as this explosive is mainly made up of NH_4NO_3 and Al powder, the emission lines associated to Al (309 nm), H (656 nm) and O (777 nm) were extracted as the most appropriate for the *3D* scatter plot. Fig. 5 shows the corresponding diagram. As shown, all the fingerprints from harmless products (green dots) are entirely differentiated from *ammonal* (red dots), when left on the surface of a conventional envelope. Again, as discussed above, the use of a less sensitive line from Al enabled us to determine whether the source of this metal was an *ammonal* fingerprint or, on the contrary, its origin was closely related to the support.

Table 2
Performance parameters (expressed in %) obtained in the recognition of fingerprints of *chloratite* deposited on the surface of different objects for courier services using a neural network classifier trained on the basis of the *Levenberg–Marquardt* rule.

Classifier[a]	*Support*	*Data set*	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*	*False positive rate*	*False negative rate*[b]
Simple	***Cardboard box***	*Training*	100	100	100	100	0	0
		Testing	99	80	100	100	0	20 (1)
Simple	***Padded envelope***	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
Simple	***Conventional envelope***	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
Extended	***Cardboard box***	*Training*	100	100	100	100	0	0
		Testing	99	80	100	100	0	20 (1)
	Padded envelope	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
	Conventional envelope	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0

[a] Simple classifiers were built by independently using information collected from the analysis of fingerprints on the surface of each particular support. Alternatively, the extended classifier was built by considering data gathered from the analysis of fingerprints on the surfaces of all the supports considered as a whole.
[b] Values in parentheses indicate the fingerprints of *chloratite* that the model has incorrectly classified. More details in the body of the text.

Table 3 lists the performance parameters of the three simple classifiers designed to identify ammonal fingerprints on the surfaces of each courier object together with those for an extended classifier that used to detect ammonal traces regardless the packing item. As the reported results clearly prove, the effectiveness of each classifier is more than satisfactory. All built classifiers were highly accurate and precise (100% for both parameters). Furthermore, while none of the positive declarations passed unnoticed to the model (note that false negative rates are 0% in all cases), at the same time, no fingerprints from harmless materials were mislabeled as an alert (false positive rates of 0% in all cases). In short, all designed classifiers are more than suitable for addressing with the matter at hand. At time, results also evidence that a simple classifier for sorting *ammonal* and harmless fingerprints on each particular courier article is not needed. Notwithstanding this proof of effectiveness, just as for the *chloratite* case, low surface concentrations of ammonal in the support as well as low Al content in their formulations both are factors that may also deteriorate the identification of this explosive.

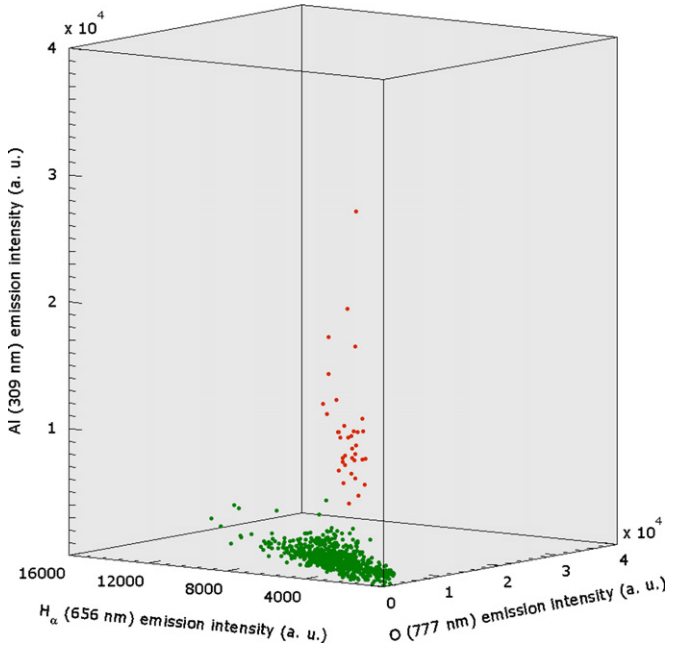

Fig. 5. *3D* scatter plot projected onto the subspace of the H (656 nm), O (777 nm) and Al (309 nm) signals for *ammonal* (*red*) and harmless (*green*) fingerprints left on the surface of on the surface of a conventional envelope. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.5. Organic explosives fingerprint identification

Finally, differentiation between organic fingerprints, either hazardous or harmless, was addressed. A total amount of 60 fingerprints of *organic explosives* (20 per each support) as well as 270 fingerprints of harmless traces (90 per each support) were considered for each dataset, *training* and *testing*, respectively. In this case, the challenge is much more than evident since explosive fingerprints share similar elemental composition with both harmless substances and supports considered in the present study. All of this, together with alterations caused on intensity values of H, N and O emissions by both relative humidity and surrounding air, tends to complicate the task. Notwithstanding these difficulties, atomic signal from H (656 nm) together with the molecular emissions of CN (388 nm) and C_2 (516 nm) were the best spectral features extracted to cope with this pursued purpose.

Fig. 6 shows the resulting *3D* scatter plot of these LIBS features for *DNT*, *TNT*, *RDX* and *PETN* fingerprints (red dots) compared to the innocuous products (green dots), when left on the surface of a cardboard box. As shown, the two classes of fingerprints are grouped into a pair of clouds along the axis of H intensities. Nevertheless, the trend observed for explosive fingerprints shows larger CN intensities together with lower C_2 values, whereas harmless traces exhibit the opposite tendency. Despite this differentiable behavior between classes, an overlapping region between hazardous and innocuous fingerprints also exists, thus deteriorating partially the differentiation. These graphical evidences are perfectly reflected in the performance results of the classifiers designed for this particular scenario, as summarized in Table 4.

In general the accuracy of the classifiers was always close to 90%. The same holds for the specificity. This means that just for a few harmless fingerprints (between 3 and 22, depending on the support and the dataset being considered), at least one of the 7 shots hitting within the involved fingerprint was declared as a positive finding by the classifier.

On the other hand, precision of the classifiers ranged between 44% and 81%, also according to the support or the dataset assessed. In other words, from the 20 explosives targets (5 each of *DNT*, *TNT*, *RDX* and *PETN*) on each support, 9, 3 or 6 fingerprints were incorrectly classified; that is, none of the 7 shots, hitting each one of such fingerprints, identified the trace as a threat.

Table 3
Performance parameters (expressed in %) obtained in the process of recognition of fingerprints of *ammonal* explosive encountered on the surfaces of different objects for courier services through the use of a neural network classifier trained on the basis of the *Levenberg–Marquardt* rule.

Classifier[a]	*Support*	*Data set*	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*	*False negative rate*	*False positive rate*
Simple	***Cardboard box***	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
Simple	***Padded envelope***	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
Simple	***Conventional envelope***	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
Extended	***Cardboard box***	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
	Padded envelope	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0
	Conventional envelope	*Training*	100	100	100	100	0	0
		Testing	100	100	100	100	0	0

[a] Simple classifiers were built by independently using information collected from the analysis of fingerprints on the surface of each particular support. Alternatively, the extended classifier was built by considering data gathered from the analysis of fingerprints on the surfaces of all the supports considered as a whole. More details in the body of the text.

The cause for which these explosive fingerprints may go unnoticed is a complex function of the amount of material in the fingerprint, as discussed above, and the similarity in optical emissions of some harmless fingerprints (sugar and sweetener) and the explosives considered (mainly *PETN*).

These circumstances also seem to be responsible for the substantial differences observed in the performance of the three single classifiers when compare to the extended one. While for the padded envelope the sorting of fingerprints benefits from the use of the extended model, for the cardboard box and the conventional envelope, the simple models performed better.

3.6. Performance of the decision tree

So far we have shown the effectiveness of each particular classifier built to identify each group of explosives versus their most common potential confusants. Nonetheless, as the fingerprint recognition scheme is a hierarchical set of three classifiers working in cascade, the overall classification process must be entirely assessed. In other words, the organized sequence of actuation must be tested in order to evaluate how the decisions made at earlier stages, and their corresponding outcomes, affect the decisions at later stages of the decision tree.

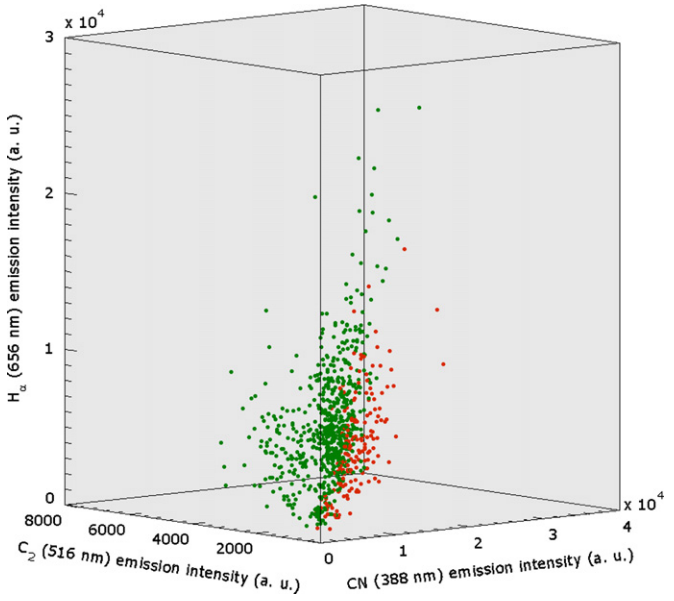

Fig. 6. *3D* scatter plot projected onto the subspace of the CN (388 nm), C_2 (516 nm) and H (565 nm) signals for *organic explosives* (*red*) and harmless (*green*) fingerprints left on the surface of on the surface of a cardboard box. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

A total of 4 decision trees were built and assessed: 3 decision trees (1 exclusive for each support), each one of them composed by their corresponding simple classifiers; and a unique decision tree, which involves the 3 extended classifiers, for evaluating residues on whatever the support. The programmed sequence *chloratite*:*ammonal*:*organics* was the configuration decided in all of them. Hence, only 115 fingerprints (5 for each of the 23 residues on each support) were hierarchically tested through each exclusive decision tree, whereas a total amount of 345 fingerprints were evaluated using the all-encompassing decision tree.

Table 5 summarizes error rates committed by each decision tree based on identification of each group of experimental fingerprints from the testing dataset used as input. Furthermore, number of misidentified fingerprints of each group that have not encountered their final correct leaf is listed in parentheses.

Through exploiting first mover advantage, there are no differences between the particular classifier for sorting *chloratite* fingerprints when embedded in the complete sequence (Table 5) as compared when it acts as a single decision step (Table 2). The classifier intended to be used for identifying *ammonal*, despite actuating in a second place, also works in a wholly effective way. Again, no differences in errors committed were observed between decisions taken when its integration within the decision tree (Table 5) or those derived from its isolated acting (Table 3). These results suggest that any *chloratite* outcome that may pass unnoticed by its identifier does not affect the performance of the following step (the next internal node) in the sequence of action.

Finally, for the classifier intended to verify the identity of *organic explosives*, only minor differences were recognized.

A total of 8, 2, and 4 explosive fingerprints (depending on the support considered) passed unnoticed during the operation of the full decision tree (Table 5), whereas 9, 3, and 6 false negative cases were missed when the classifier acts as a decision element on its own (Table 4). Similar results were reached when identification of harmless compounds was attempted; 6, 21, and 15 fingerprints in contrast to 3, 22 and 11 traces, on cardboard or both padded and conventional envelopes, respectively.

As a final remark, it should added that decision tree encompassing all supports provides slightly better results as compared

Table 4
Performance parameters (expressed in %) obtained in the recognition of fingerprints of organic explosives deposited on the surface of different objects for courier services using a neural network classifier trained on the basis of the *Levenberg–Marquardt* rule.

Classifier[a]	*Support*	*Data set*	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*	*False positive rate*[b]	*False negative rate*[c]
Simple	***Cardboard box***	*Training*	93	100	91	71	9 (8)	0
		Testing	89	55	97	79	3 (3)	45 (9)
Simple	***Padded envelope***	*Training*	89	100	87	63	13 (12)	0
		Testing	77	85	76	44	24 (22)	15 (3)
Simple	***Conventional envelope***	*Training*	95	100	93	77	7 (6)	0
		Testing	84	70	87	54	13 (12)	30 (6)
Extended	***Cardboard box***	*Training*	89	100	87	63	13 (12)	0
		Testing	91	65	97	81	3 (3)	35 (7)
	Padded envelope	*Training*	93	100	91	71	9 (8)	0
		Testing	87	70	91	64	9 (8)	30 (6)
	Conventional envelope	*Training*	92	75	96	79	4 (4)	25 (5)
		Testing	85	50	92	59	8 (8)	50 (10)

[a] Simple classifiers were built by independently using information collected from the analysis of fingerprints on the surface of each particular support. Alternatively, the extended classifier was built by considering data gathered from the analysis of fingerprints on the surfaces of all the supports considered as a whole.
[b] Values in parentheses indicate the harmless fingerprints that the model has classified as a positive threat.
[c] Values in parentheses indicate the fingerprints of *organic explosives* that the model has passed unnoticed. More details in the body of the text.

Table 5
Results on identifying fingerprints of the *testing* dataset when the overall actuation sequence of the decision tree is applied.

Decision tree	*Fingerprints group*	*Error rate* (%)		
		Cardboard box	***Padded envelope***	***Conventional envelope***
Exclusive (***simple classifiers***)				
	Chloratite	0.8 (1)	0.0 (0)	0.0 (0)
	Ammonal	0.0 (0)	0.0 (0)	0.0 (0)
	Organic explosives	6.7 (8)	1.7 (2)	3.3 (4)
	Harmless	5.0 (6)	17.5 (21)	12.5 (15)
All-encompassing (***extended classifiers***)				
	Chloratite	0.8 (1)	0.0 (0)	0.0 (0)
	Ammonal	0.0 (0)	0.0 (0)	0.0 (0)
	Organic explosives	4.2 (5)	4.2 (5)	9.2 (11)
	Harmless	4.2 (5)	7.5 (9)	6.7 (8)

Note: Values in parentheses indicate the number of misclassified fingerprints of each group after being hierarchically tested through the complete decision tree. More details in the body of the text.

to the exclusive decision trees for each particular support. In other words, discrimination between hazardous and harmless compounds may take place using a unique decision tree regardless the packaging objects inspected.

4. Conclusions

In this paper, a cutting edge method for classification of explosive fingerprints using laser-induced breakdown spectroscopy (LIBS) is presented. The approach involves a combination of spectral signatures and a data processing routine based on machine learning methods. A decision tree involving 3 neural network classifiers trained by the *Levenberg–Marquardt* rule in a hierarchical structure, capable for discriminating 345 experimental combinations (5 independent fingerprints from 23 materials left on the surfaces of 3 supports) has been successfully developed. The envisaged approach outperforms satisfactorily a complex challenge, discrimination between compounds sharing a similar chemical composition, in the use of LIBS for the matter at hand. As demonstrated, extraction of optimal features from LIBS spectra reveals hidden patterns that can be used for building a concise classifier for the distribution of both explosive and harmless classes. Thus, after embedding the exclusive features from an unknown that relate with each particular classifier, our custom algorithm (*decision tree*) makes an accurate diagnosis on the identity of the finding. Furthermore, all variables used by this approach are directly linked to tangible emission features, thus breaking away with other complex and abstract variables of other conventional chemometric tools previously used (scores from PCA, for instance). In any case, it is clear that building a universal strategy that addresses all situations is extremely complex, but the design of a model based on a multistage architecture allows to increasing the accuracy, precision, sensitivity and specificity of the sorting tasks. Indeed, the reported results, with error rates not much beyond 10%, demonstrate the suitability of the algorithm for recognizing fingerprints encountered on the surface of some packing objects commonly used.

Currently, the model is being extended to a broader database of harmless materials and daily life objects. Furthermore, studies on the effects of ambient conditions on the performance of the classifiers are in progress.

Acknowledgments

Research was supported by the Excellence Project P07-FQM-03308 of the Secretaría General de Universidades, Investigación y Tecnología, Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía. Part of the research made has been also supported with funds of the contract SEDUCE between Indra Sistemas, S. A. and the University of Malaga.

References

[1] C. López-Moreno, S. Palanco, J.J. Laserna, F.C. De Lucia Jr., A.W. Miziolek, R.A. Walters, A.L. Whitehouse, J. Anal. At. Spectrom. 21 (2006) 55–60.
[2] J.L. Gottfried, F.C. De Lucia Jr., C.A. Munson, A.W. Miziolek, Anal. Bioanal. Chem. 395 (2009) 283–300.
[3] F.C. De Lucia Jr., J.L. Gottfried, A.W. Miziolek, Opt. Express 17 (2009) 419–425.
[4] R. Noll, C. Fricke-Begemann, in: H. Schubert, A. Rimski-Korsakov (Eds.), Stand-off Detection of Suicide Bombers and Mobile Subjects (NATO Security through Science Series B: Physics and Biophysics), Springer, The Netherlands, 2006, pp. 89–99.
[5] C.A. Munson, J.L. Gottfried, F.C. De Lucia Jr., K.L. McNesby, A.W. Miziolek, in: J. Yinon (Ed.), Counterterrorist Detection Techniques of Explosives, Elsevier, The Netherlands, 2007, pp. 279–321.
[6] F.C. De Lucia Jr., J.L. Gottfried, C.A. Munson, A.W. Miziolek, Spectrochim. Acta B 62 (2007) 1399–1404.
[7] J.L. Gottfried, F.C. De Lucia Jr., C.A. Munson, A.W. Miziolek, Spectrochim. Acta B 62 (2007) 1405–1411.
[8] J.L. Gottfried, F.C. De Lucia Jr., C.A. Munson, A.W. Miziolek, J. Anal. At. Spectrom. 23 (2008) 205–216.
[9] V. Lazic, A. Palucci, S. Jovicevic, C. Poggi, E. Buono, Spectrochim. Acta B 64 (2009) 1028–1039.
[10] F.C. De Lucia Jr., J.L. Gottfried, C.A. Munson, A.W. Miziolek, Appl. Opt. 47 (2008) G112–G121.
[11] J.L. Gottfried, F.C. De Lucia Jr., A.W. Miziolek, J. Anal. At. Spectrom. 24 (2009) 288–296.
[12] F.C. De Lucia, J.L. Gottfried, Appl. Opt. 51 (2012) B83–B92.
[13] R. González, P. Lucena, L.M. Tobaria, J.J. Laserna, J. Anal. At. Spectrom. 24 (2009) 1123–1126.
[14] V. Lazic, A. Palucci, S. Jovicevic, M. Carpanese, Spectrochim. Acta B 66 (2011) 644–655.
[15] K.D. Morton Jr, P.A. Torrione, L. Collins, Proc.—Soc. Photo-Opt. Instrum 8018 (2011) 801817.
[16] B. Clarke, E. Fokoue, H.H. Zhang, Principles and Theory for Data Mining and Machine Learning, Springer, New York, 2009.
[17] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, San Diego, 1990.
[18] A.J. Smola, P.L. Bartlett, B. Schölkopf, D. Schuurmans, Advances in Large Margin Classifiers, The MIT Press, Cambridge, 2000.
[19] A.R. Webb, Statistical Pattern Recognition, second ed., John Wiley & Sons, Inc., West Sussex, 2002.
[20] S.B. Kotsiantis, Informatica 31 (2007) 249–268.
[21] L.I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, John Wiley & Sons, Inc., New Jersey, 2004.
[22] B. Tso, P.M. Mather, Classification Methods for Remotely Sensed Data, second ed., Taylor & Francis, London, 2001.
[23] K. Levenberg, Q. Appl. Math. 2 (1944) 164–168.
[24] D. Marquardt, J. Soc., Ind. Appl. Math. 11 (1963) 431–441.
[25] O.P. Jasuja, M.A. Toofany, G. Singh, G.S. Sodhi, Sci. Justice 49 (2009) 8–11.
[26] B.J. Jones, R. Downham, V.G. Sears, Surf. Interface Anal. 42 (2010) 438–442.