Article

Application of computational intelligence methods for the automated identification of paper-ink samples based on LIBS

Krzysztof Rzecki 1*, Tomasz Sońwicki 1, Mateusz Baran 1, Michał Niedźwiecki 1, Małgorzata Król 2, Tomasz Łojewski 3, U Rajendra Acharya 4,5,6*, Özal Yıldırım 7, and Paweł Plawiak 1,8

1 Faculty of Physics, Mathematics and Computer Science, Cracow University of Technology, Warszawska 24, 31-155 Krakow, Poland; krz@pk.edu.pl (K.Rz.); tosniwicki@pk.edu.pl (T.S.); plawiak@pk.edu.pl (P.P.); mbaran@pk.edu.pl (M.B.); nkg@pk.edu.pl (M.N.)
2 Laboratory for Forensic Chemistry, Faculty of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland; krolm@chemia.uj.edu.pl (M.K.)
3 Faculty of Materials Science and Ceramics, AGH University of Science and Technology, Mickiewicza 30 Av., Krakow 30-059, Poland; lojewski@agh.edu.pl (T.L.)
4 Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore; Rajendra_Udyavara_ACHARYA@np.edu.sg (R.A.)
5 Department of Biomedical Engineering, School of Science and Technology, Singapore Polytechnic, Singapore; Rajendra_Udyavara_ACHARYA@np.edu.sg (R.A.)
6 School of Medicine, Faculty of Health and Medical Sciences, Taylor’s University, 47500 Subang Jaya, Malaysia; Rajendra_Udyavara_ACHARYA@np.edu.sg (R.A.)
7 Department of Computer Engineering, Munzur University, Tunceli, Turkey; oyildirim@munzur.edu.tr (O.Y.)
* Correspondence: plawiak@pk.edu.pl (P.P.), krz@pk.edu.pl (K.Rz.)

Abstract: Laser-induced breakdown spectroscopy (LIBS) is an important analysis technique with applications in many industrial branches and fields of scientific research. Nowadays, the advantages of LIBS are impaired by the main drawback in the analysis of collected data. This procedure is essentially based on the comparison of lines present in the spectrum with a literature database. This paper proposes the use of various computational intelligence methods to develop a reliable and fast classification of non-destructively acquired LIBS spectra into a set of predefined classes. We focus on a specific problem of classification of paper-ink samples into 30 separate, predefined classes. For each of 30 classes (10 pens of each of 5 ink types combined with 10 sheets of 5 paper types plus empty pages) 100 LIBS spectra are collected. Four variants of preprocessing, seven classifiers (Decision trees, Random forest, k-Nearest Neighbour, Support Vector Machine, Probabilistic Neural Network, Multi-Layer Perceptron, and Generalized Regression Neural Network), 5-fold stratified cross-validation and test on an independent set (for methods evaluation) scenarios are employed. Our developed system yielded an accuracy of 99.08% with average classification time of about 0.12 s is obtained using the random forest classifier. Our results clearly demonstrates that machine learning methods can be used to identify the paper-ink samples based on LIBS reliably at a faster rate.

Keywords: classification; computational intelligence methods; discrimination power; LIBS; machine learning; paper-ink analysis

1. Introduction

During the decade, there have been important developments in laser induced breakdown spectroscopy (LIBS). This atomic emission spectroscopy technique, also known as laser induced plasma spectroscopy (LIPS), is used for qualitative and quantitative chemical analysis of samples in all states of matter [1]. In this technique, high-power, short-duration laser pulse causes an ablation of the analyzed material, which due to its high temperature (10 000K) dissociates into excited ions and atoms. When
plasma cloud cools down this excited species revert to lower energy states and emit optical radiation
which could be recorded and analyzed, revealing information about the elemental composition of
the sample [2]. LIBS spectra are generally very rich in emission lines coming from excited atoms and
ions occuring in a high temperature plasma cloud. Physical and chemical phenomena behind LIBS
are not fully understood yet, as they are very complex in nature. Nevertheless, LIBS applications
are recently rapidly growing due to the number of advantages of this method. The most important
ones are minimally destructive measurement with little or no sample preparation, efficiency and
possibility to analyze in real-time all elements in a single laser shot in all three states [3]. For solids both
mapping (2D) and depth profiling (3D) can be obtained. LIBS can be used for quantitative chemical
analysis, material identification and discrimination. This method can be applied in the laboratories and
industrial plants, and even at stand-off distances of tens of meters [4]. These properties predispose the
LIBS technique for use in many fields: space exploration [5], remote analysis of hazardous materials [6],
on-line quality control in various industries [7], cultural heritage studies [8], forensic chemistry [9,10],
geology [11], weld quality assurance [12], robotics [13] and many others [14–17].

Although currently LIBS is already an established technology, spectrochemical LIBS analysis is
not straightforward. Identification of elemental constituents in the sample is usually based on the
strongest lines present in the LIBS spectrum (so called persistent lines), which are compared with a
literature database collected for all the elements from the periodic table. This analysis is cumbersome
and time-consuming because the emission spectrum is determined also by the properties of the plasma,
not only by the composition of the examined sample [18].

At present, the great advantages of LIBS seem to be impaired to some extent by the main
drawbacks: problems coming from often poor signal reproducibility, the impact of sample composition
on signals recorded for individual components (matrix effects) and the difficulty of performing an
overall reliable data analysis (sometimes over 500 spectral lines need to be interpreted). Thus, a
question immediately raises whether the LIBS method can be supported by the modern achievements
in the field of computational intelligence [19] in an attempt to overcome the limitations of the current
LIBS spectrum analysis methodology. The analogous problem of classification of one dimensional
series of data is well known in computer science. The methods developed for this problem have been
successfully applied in many areas [20–27]. So, application of machine learning methods to LIBS
spectra of samples may certainly give a strong impact for further developments and applications.

Motivated by the aforementioned arguments we decided to apply computational intelligence
methods to the classification of paper-ink samples for forensic purposes, based on LIBS spectra of the
samples. The problem of discrimination of different paper-ink samples has already been addressed
in previous studies based on the LIBS spectrum analysis. Trejos et al. showed [28] that the highest
discrimination power (DP) (96.4%) was obtained when comparisons were done qualitatively by spectral
overlap of the regions of interest (3 different emission lines monitored per element) and quantitatively
followed by pairwise comparisons (1 emission line per element) using ANOVA (analysis of variance).
Kula et al. [29] presented the LIBS method as a useful tool in qualitative elemental differentiation
of ink samples. They have obtained the discrimination power (DP) coefficient of 61%, 82% and 83%
for red, black and blue inks, respectively. Elsherbiny and Nassef [30] studied the dependence of the
obtained spectra of the obtained spectra of various black gel inks on the wavelength of laser excitation
reporting the DP in the range of 88% to 91%. In the next study [31], the elemental analysis with the
use of LIBS was performed on writing and inkjet inks, toners as well as on office paper. The LIBS
results supported by pairwise comparison analysis (ANOVA with Tukey’s post hoc test) provided
discrimination power of 98.4% (3 indistinguishable/190 compared pairs) for the toners and 100% for
the inkjet inks. Moreover, these three undifferentiated toner pairs were discriminated using the Student
t-test at a 95% confidence limit. The authors claimed that LIBS as a suitable tool for the determination
of elemental composition of sample can be a part of a procedure for questioned document examination.

The problem of paper-inks classification, on the other hand, is less common in the literature. To the
best of authors knowledge there are only two articles dealing with LIBS spectra of paper-ink samples
employed in classification/fitting problems by means of computational intelligence techniques. In a paper by Hoehses et al. [32], the benefit of applying several independent chemometric methods to LIBS data was demonstrated. A consecutive methodology of applying soft independent modelling of class analogy (SIMCA) and partial least-squares discriminant analysis (PLS-DA) enabled the step-wise classification of data and separation of inks that were not identified by principal component analysis (PCA). The SVM yielded a correct classification rate of 87% and cross-validation accuracy amounted to 81%. In the second paper [33], multiple methods such as three comparative functions (linear correlation, overlapping integral, and sum of squared deviations) and two advanced statistical methods (multivariate curve resolution alternating least squares (MCR-ALS) with classification tree and discriminant analysis (DA)) were applied to statistically evaluate LIBS spectra. The newly introduced MCR-ALS/DA methodology showed identification of the paper and printer type with an accuracy of 96.3% and 83.3%, respectively.

In the present paper, we show that the similar problem can be solved using computational intelligence methods. With our approach much better DP is achieved and the processing time is greatly reduced. Our results show that the machine learning algorithms can be used to analyze the LIBS spectra.

The classification problem discussed in this paper is very difficult due to similarities between LIBS spectra from different paper-ink samples and differences between such spectra from one single class. Additionally, we operate over the long samples with significant noise. To solve this classification problem, we tested many preprocessing ways and computational intelligence methods, however the paper presents only selected and best ones.

2. Materials and Methods

2.1. Materials

Fifty ballpoint pens (10 items of each of 5 models) produced by four different manufacturers from Germany and Poland (more details in Table 1) were purchased in Poland. Fifty sheets (10 sheets of each of 5 types) of five Canadian certified reference papers were used (papers denoted A, D, L, N and O from the set “Fillers in paper” supplied by A.S.O. Design Canada).

Each particular ink from each of 10 pens of 5 types was deposited (as straight lines) on each of 10 sheets of 5 types of papers in the form of straight lines, using standard hand pressure. All 2500 (50 sheets of papers with 50 deposits each) paper-ink samples were placed in plastic bags and stored in darkness at room temperature.

In the experiment data for 30 classes (A, A+B, A+R, A+S, A+SB, A+T, D, D+B, . . . , O+SB, O+T) based on the combination of ink, paper and empty papers were recorded.

2.2. LIBS

The analysis of all paper-ink samples was carried out using a laser induced breakdown spectroscopy system LIBS-6 (Applied Photonics, United Kingdom). It consists of an integrated Q-switched Quantel Ultra Nd-YAG laser (Quantel, France) working at λ = 1064 nm emitting a maximum energy equal to 150 mJ due to one laser pulse (6 ns), and an Avaspec-2048-2-USB2 fibre optic...
Czerny-Turner spectrometer (6-channel) with a CCD detector (Avantes, The Netherlands). The system was also equipped with a camera enabled to observe the analyzed object and a movable sample table. The LIBS-6 system was operated by LIBSoft V6.0.1 software (Applied Photonics, United Kingdom).

Under normal conditions, because of no moving elements inside, a wavelength calibration of the spectrometer was not required. Every measurements were conducted in air under atmospheric pressure. A Q-switch delay time of 165 µs, the integration delay time of 1.27 µs and integration time of 1.2 ms were used. Samples were analysed directly without any special preparation. They were placed at the sample table at the focal point of the focusing lens (at a distance determined by the nozzle about 70 mm from the optical head). The diameter of the ablation spot, ranged from 0.6 mm to 1 mm, was dependent on the analysed material.

The results of LIBS analysis are an emission spectrum – intensity distribution of the radiation energy (expressed in electron volts) emitted by the analysed object depends on the wavelength measured in nanometres. The emission spectra were collected in the UV-vis range (185 nm to 904 nm, spectral resolution 0.1 nm).

In total, 30 classes of experimental data listed above were recorded. For each class 100 LIBS spectra were acquired, resulting in 3 000 samples of the LIBS emission spectrum. A sample spectrum is shown in Figure 1. The spectrum of each sample is available on our website [34].

2.3. Data Analysis

The analysis of the LIBS spectrum consisted of the following steps:

1. Independent preprocessing of LIBS samples.
2. Selection of data samples for the cross validation and testing sets.
3. Data analysis based on computational intelligence methods.
4. Evaluation of the results.

Steps of the experiment are presented in the flowchart in Figure 2.

2.3.1. Signal preprocessing

The initial preprocessing of the raw data was performed to reduce the number of data points corresponding to individual samples of the LIBS spectrum. This preprocessing step was applied because the spectral range of LIBS is very broad (11 746 spectral lines in this case) and from the viewpoint of computational intelligence applied to classification tasks it contains irrelevant information. It can be expected that the reduction of the datasets enhances extraction of characteristic features of the distinctive classes (types of material) [35]. Four preprocessing steps were taken under consideration:

- Data reduction by removing the data points from the beginning (first 3 746 data points) and from the end (last 1 000 data points) of the LIBS spectrum, which do not contain relevant information. 7 000 data points are left for further analysis.
5 of 15

- Normalization of energy values to the interval $[0, 1]$.
- Standardization of energy values, so the mean value becomes equal to 0 and standard deviation becomes equal to 1.
- Application of arithmetic averaging over consecutive number $AP$ (called averaging parameter) of data points after data reduction and prior to either normalization or standardization. As a result, we reduce the original length of the data vector by a factor of $AP$. The tested values of $AP$ ranged from 5 to 100.

Four preprocessing ways based on these steps were constructed and evaluated: data reduction, averaging (applied or not), and either standardization or normalization. These steps are depicted at the bottom of Figure 2.

Custom software in Python language was developed for data preprocessing tasks.

2.3.2. Cross-validation

The set of 3 000 LIBS emission spectrum samples was divided into two subsets [36] a training subset containing 90% (2 700) of samples and a test subset containing 10% (300) of samples. Selection of samples for both subsets was performed in a stratified way. The cross-validation subset was used for parameter optimization and then the test subset was used for final evaluation of different types of classifiers with optimized parameters.

The 5-fold stratified cross-validation method was applied to build the training and validation data sets. In each of the five analyzed combinations a training set of 72 LIBS spectra from each of 30 classes (2 160 in total) and a validation set of 18 LIBS spectra from each of 30 classes (540 in total) were used.

Evaluation of the classifiers with optimized parameters was performed separately for each training subset using the test subset. Results of evaluation based on each training subset were obtained by averaging the folds.

2.3.3. Computational Intelligence Methods

The samples of LIBS spectra, after preprocessing and cross-validation were fed to the classifiers: Generalized Regression Neural Network (GRNN) [37], Probabilistic Neural Network (PNN) [38], Multi-Layer Perceptron (MLP) [39], Support vector machine (SVM) [40], Decision trees (DT) [41], $k$-nearest neighbour (kNN) classifier [42] and Random forest (RF) [41].

Each method is potentially dependent on a set of parameters that are either quantitative or qualitative. They have an influence on the overall performance of the method. These parameters were
separately optimized for each machine learning algorithm to receive the lowest number of erroneous classes. The basic categorical parameters for each method were set to constants listed in Table 2.

Table 2. Computational intelligence methods and their basic parameters used for LIBS spectra identification.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Decision trees</td>
<td>Criterion: gini, Splitter type: best,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximum depth: none</td>
</tr>
<tr>
<td>2</td>
<td>Random forest</td>
<td>Criterion: gini, Maximum depth: none</td>
</tr>
<tr>
<td>3</td>
<td>kNN</td>
<td>Distance metric: Minkowski</td>
</tr>
<tr>
<td>4</td>
<td>SVM</td>
<td>Type: nuSVC, Type of kernel function: radial basis function</td>
</tr>
<tr>
<td>5</td>
<td>Neural Network</td>
<td>Type: PNN</td>
</tr>
<tr>
<td>6</td>
<td>Neural Network</td>
<td>Type: GRNN</td>
</tr>
<tr>
<td>7</td>
<td>Neural Network</td>
<td>Type: MLP</td>
</tr>
</tbody>
</table>

Methods based on adaptive neuro-fuzzy inference system (ANFIS) [43] and Gaussian process [44] were also tested, but they are computationally intensive and take long time to receive results even with reduced data set.

The custom software developed for this study uses the scikit-learn Python library [45] as a source of implementations of employed computational intelligence algorithms.

2.3.4. Evaluation criteria

Evaluation of classification process was based on methodology described in [46,47]. The 5-fold cross-validation strategy was adopted in our experiment. Five performance parameters namely accuracy (ACC), sensitivity (SEN), specificity (SPE), mean values and Cohen’s kappa (κ) were calculated. The mean values were calculated to estimate the overall performance of the computational intelligence methods used in this study separately for the task of recognition of each of the different classes of LIBS spectra.

To test whether there are classes which are classified with higher accuracy, specificity or sensitivity, we calculated for each class \( S \) the number of true positives \( TP(S) \), false positives \( FP(S) \), true negatives \( TN(S) \) and false negatives \( FN(S) \) separately. Then, the accuracy \( ACC(S) \), sensitivity \( SEN(S) \), and specificity \( SPE(S) \) with respect to class \( S \) are defined as averages over five folds of cross-validation \((K = 5)\) is the number of folds):

\[
ACC(S) = \frac{1}{K} \sum_{i=1}^{K} \frac{TP_i(S) + TN_i(S)}{TP_i(S) + FP_i(S) + TN_i(S) + FN_i(S)}
\]

\[
SEN(S) = \frac{1}{K} \sum_{i=1}^{K} \frac{TP_i(S)}{TP_i(S) + FN_i(S)}
\]

\[
SPE(S) = \frac{1}{K} \sum_{i=1}^{K} \frac{TN_i(S)}{TN_i(S) + FP_i(S)}
\]

where \( TP_i(S), FP_i(S), TN_i(S), FN_i(S) \) are, respectively, the number of true positives, false positives, true negatives and false negatives for the \( i \)th fold of cross-validation with respect to class \( S, i = 1, 2, \ldots, K \).
Figure 3. Visualization of data preprocessing for single sample from the class A. Figure (a) depicts an example of a raw spectrogram, (b) shows data after reduction, (c) represents averaged data for AP equal to 20, which is further normalized as shown in (d).

The overall values of ACC, SEN and SPE for the classification system are the arithmetic means of ACC(S), SEN(S) and SPE(S) over all classes.

To evaluate the degree of discrimination between two different samples discrimination power coefficient (denoted DP) was proposed [48]. DP is the ratio of the number of correctly identified pairs of test samples (identified as from different classes) to the number of all possible pairs of test samples. It can be calculated by the following equation:

$$DP = \frac{2D}{T(T-1)} = 1 - \frac{2N}{T(T-1)},$$  \hspace{1cm} (4)

where:

- $D$ – the number of differentiated pairs, that is the number of pairs correctly identified as belonging to the same class or correctly identified as belonging to different classes,
- $N$ – the number of non-differentiated pairs, $N = T(T-1)/2 - D$,
- $T$ – the total number of analysed samples (the total number of possible pairs of samples is equal to $T(T-1)/2$).

3. Results

An example of a raw spectrum and the output from successive stages of signal preprocessing (data reduction, averaging, normalization) are shown in Figure 3. The averaging stage is optional and normalization can be replaced by standardization.
The averaging within preprocessing is performed mainly to increase the speed of training and testing, however it can also be described as a feature extraction procedure. Obviously the higher AP is, the faster training and validation are for each of seven tested classification methods. The computation time and classification accuracy were investigated for AP equal to 5, 10, 20, 50, 100, 150, 200 and 250. The best trade-off between the number of erroneous classifications and the reduction in the processing without averaging (values of speed-up were dependent on the classification method). Thus, for AP equal 20 the number of final spectral lines used decreased from 11,746 to 350.

Parameter selection is a key part in reaching the optimal overall performance of a classification system. There are many possible options available, so it is important to analyze the outcomes of experiments for different values of parameters to demonstrate the possibilities of machine learning methods. Particular classification methods depend on the various basic parameters set as listed in Table 2 and some other parameters that were optimized. Optimization of these parameters was performed in two steps. The first step was to find out the general range of values for each parameter for the fine-tuning procedure. Then the detailed grid search of the selected ranges of parameters was performed.

The classification algorithms, the tuning parameters and range of these parameter values are described below.

- Decision trees – the number of features to consider when looking for the best split was optimized in two different ranges: from 100 to 7,000 with step equal to 100 when no averaging in preprocessing was done, and from 10 to 350 with step equal to 10 when averaging in preprocessing was performed.
- Random forest – two parameters dependent on averaging were optimized during preprocessing. If averaging was not applied, the number of trees in the forest was optimized in range from 10 to 200 with step 10 and from 200 to 1,000 with step equal to 50 and the number of features to consider when looking for the best split was optimized in the same range. If averaging was performed, both parameters were optimized in range from 10 to 350 with step 20.
- kNN – the number of neighbours was optimized in the range from 1 to 4 and exponent used to calculate the Minkowski distance was optimized from 1 to 10 with step 1.
- SVM – the gamma parameter of the RBF kernel function was optimized in range from 0.01 to 1.00 with step 0.01 and the nu parameter of the nu-SVC algorithm, related to the error tolerance of the SVM classification, was optimized in range from 0.01 to 1.00 with step 0.01.
- PNN – the radius (the spread) of the kernel function of the network (standard deviation for the probability density function of the normal distribution). This parameter was optimized in range from 0.01 to 0.20 with step equal to 0.01 when normalization in preprocessing was used and from 0.1 to 1.0 with step 0.1 when standardization was used.
- GRNN – the spread parameter with identical meaning and range as in PNN was optimized;
- MLP – the number of neurons was optimized in a range from 10 to 200 with step equal to 10. The activation function was selected from ‘identity’, ‘logistic’, ‘tanh’, ‘relu’. The solver for weight optimization was chosen from ‘lbfgs’ (an optimizer in the family of quasi-Newton methods), ‘sgd’ (a stochastic gradient descent) or ‘adam’ (a stochastic gradient-based optimizer proposed in [49]).

These optimal parameters were selected in a series of experiments to maximize the mean ACC, SEN and SPE values described in the section 2.3.4.

The best results of the experiments with different preprocessing methods, machine learning algorithms, based on 5-fold cross-validation (3,000 samples were divided into training set containing 2,160 samples, validation set with 540 samples and test sets containing 300) are presented in Table 3. The preprocessing involves data reduction, with or without averaging (as in table noted) and standardization. Generally the classification results were worse, when the preprocessing included normalization instead of standardization. In the table, we show the values of ACC, SEN, SPE (and
Table 3. Performance results for optimized parameters.

<table>
<thead>
<tr>
<th>Averaging Method</th>
<th>Parameters</th>
<th>ACC (%)</th>
<th>SEN (%)</th>
<th>SPE (%)</th>
<th>MEAN (%)</th>
<th>κ (%)</th>
<th>DP (%)</th>
<th>Training time [s]</th>
<th>Testing time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No DecisionMax features trees = 6 100</td>
<td></td>
<td>98.08</td>
<td>71.13</td>
<td>99.00</td>
<td>89.40</td>
<td>70.14</td>
<td>98.40</td>
<td>24.13</td>
<td>0.0002</td>
</tr>
<tr>
<td>No Random forest N estimators = 700 Max features = 950</td>
<td></td>
<td>99.08</td>
<td>86.27</td>
<td>99.53</td>
<td>94.96</td>
<td>85.79</td>
<td>99.08</td>
<td>1667.57</td>
<td>0.1198</td>
</tr>
<tr>
<td>Yes kNN N = 1 Exponent = 1.00</td>
<td></td>
<td>97.14</td>
<td>57.07</td>
<td>98.52</td>
<td>84.24</td>
<td>55.59</td>
<td>97.94</td>
<td>0.0505</td>
<td>0.0013</td>
</tr>
<tr>
<td>Yes SVM Nu = 0.13 Gamma = 0.02</td>
<td></td>
<td>98.92</td>
<td>83.80</td>
<td>99.44</td>
<td>94.05</td>
<td>83.24</td>
<td>99.00</td>
<td>3.60</td>
<td>0.0012</td>
</tr>
<tr>
<td>Yes PNN Spread = 0.9</td>
<td></td>
<td>97.25</td>
<td>58.73</td>
<td>98.58</td>
<td>84.85</td>
<td>57.31</td>
<td>97.98</td>
<td>0.0115</td>
<td>0.0194</td>
</tr>
<tr>
<td>Yes GRNN Spread = 0.1</td>
<td></td>
<td>96.94</td>
<td>54.07</td>
<td>98.42</td>
<td>83.14</td>
<td>52.48</td>
<td>97.56</td>
<td>0.0085</td>
<td>0.0070</td>
</tr>
<tr>
<td>No MLP N. of neurons = 120</td>
<td></td>
<td>98.22</td>
<td>73.27</td>
<td>99.08</td>
<td>90.19</td>
<td>72.34</td>
<td>98.36</td>
<td>221.76</td>
<td>0.0011</td>
</tr>
</tbody>
</table>

their mean value denoted MEAN, κ and DP, and average time of the training for all 5-fold training and testing samples.

Table 4 shows the values of ACC(S), SEN(S), and SPE(S) for all S classes and all classifiers with appropriate preprocessing methods. It can be seen from the table that the most easily distinguishable (highest ACC(S), SPE(S) and SEN(S)) samples are in 6 classes: A, A+T, D, N, O and O+T (ACC(S), SPE(S) and SEN(S) reached the value of 100%) using random forest method. Additionally, samples from 2 classes: L and N+T are well recognized by SVM and MLP methods as well. The least distinguishable classes are L+S and O+B with value of SEN(S) less than 50% by random forest method, however SPE(S) value was over 99%. It means that the samples from these classes are often assigned to other classes.

The computations were run on a virtual machine with Intel(R) Xeon(R) CPU L5640 @ 2.27 GHz (12 cores without HyperThreading were used) and 20 GB RAM DDR3 1333 MHz. Independent computations were performed in parallel. The total processing time of each step of computations depends on selected method and its parameters. The training time varied between less than a second for GRNN, PNN and kNN, and almost half an hour for random forest classifiers. The testing time for one sample varied between about 1-2 ms for decision tree and MLP, and more than 0.1 s for random forest, PNN and GRNN classifiers. The time necessary to complete the training was longer but it is less important than the classification speed that is critical at the testing stage. The time required for the identification of a particular sample was less than 1 second (however the exact value depends on the method) and is shown in Table 3.

4. Discussion

The results of experiments confirm that the computational intelligence methods can be used to analyze LIBS data and obtain accurate classification of paper-ink samples (please see Table 3). We have obtained an accuracy of 99.08% for the best classifier. The results of the present study can not be
Table 4. Results of classification per class.

<table>
<thead>
<tr>
<th>Class</th>
<th>Average SEN(%)</th>
<th>Standard deviation SEN(%)</th>
<th>Average ACC(%)</th>
<th>Standard deviation ACC(%)</th>
<th>Average SPE(%)</th>
<th>Standard deviation SPE(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>98.27</td>
<td>99.13</td>
<td>99.66</td>
<td>99.59</td>
<td>99.45</td>
<td>99.59</td>
</tr>
<tr>
<td>S</td>
<td>98.27</td>
<td>99.13</td>
<td>99.66</td>
<td>99.59</td>
<td>99.45</td>
<td>99.59</td>
</tr>
<tr>
<td>N + S</td>
<td>99.59</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>A</td>
<td>98.27</td>
<td>99.13</td>
<td>99.66</td>
<td>99.59</td>
<td>99.45</td>
<td>99.59</td>
</tr>
<tr>
<td>A + B</td>
<td>99.59</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>A + C</td>
<td>99.59</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>A + C</td>
<td>99.59</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Directly compared with those obtained in [29,30] focusing on the problem of discrimination between two samples which were not assigned to a priori classes. However, the methods used in [29,30] used Plasus SpecLine 2.13 software for the identification of spectral lines and National Institute of Standards and Technology (NIST) spectral database, gave values of the DP in the range from 85% to 92%. The same standard methods with visual spectra was applied to the data collected in our study gave DP equal to 90.6%. After adopting our algorithms to the problem of discrimination between two different samples, we achieved the DP of 99.0% with SVM and more than 98% with decision trees, random forest, PNN and MLP classifiers. The limitation of applying our solution to the problem of discrimination is that the classes must be predefined.

Additionally, application of an automated method based on machine learning greatly reduced the time required for spectrum recognition. The recognition time for a single spectrum was less than 1 second.

The detailed analysis of the Table 3 shows that the best method for the class of problems addressed in the present study is the random forest classifier. Parameters for this classifier were optimized using the 5-fold cross-validation using averaged and standardized data. The classifier during test over independent test set was able to correctly recognize 1,294 cases of LIBS spectra out of 1,500 cases and
had no erroneous recognitions in training sets (10 800 cases). This classifier achieved an accuracy of 99.08%, sensitivity of 86.27%, specificity of 99.53%, kappa of 85.79% and DP of 99.08% for the test set.

Taking into account the number of erroneous classifications under different settings, the worst classifier was GRNN, however it was the fastest classifier during training.

Four preprocessing methods were applied on the raw data prior to the classification. Hence, we have shown that the success of classification depends on careful selection of preprocessing procedures. However, although in most cases the actual selection of the pair preprocessing-classifier determines the overall performance of the classification method, for some specific selections of classifier the final classification performance depends only marginally on the preprocessing. For this reason, it is not possible to select a single best preprocessing method.

The estimation of an optimal AP for decision trees, kNN, SVM and GRNN methods will be the subject of the future research. It should be pointed out that the problem of optimal data reduction is important. The length of data series influences the time needed to train the classifiers although it has only marginal effect on the time needed for classification of a test sample.

Looking at the spectral data processed in the experiment, the averaging does not seem to be an appropriate preprocessing method for data reduction because the differences between different kinds of samples are inconspicuous. Authors took two other methods into consideration: resampling and wavelet decomposition. The first method resamples the input data at given wavelengths of the original sequence. To obtain 350 samples, resampling with \( AP = 20 \) was performed. This method, yielded worse results during identification when averaging step was replaced. The second method performs one dimensional Haar wavelet decomposition [50]. The initial results are promising, however this method requires more detailed investigation in the future. This technique needs further research to determine the type of wavelet decomposition structure and its parameters which will lead to the best results.

The results of the present study show that modern computational intelligence methods can be used for the classification of LIBS spectra into predefined classes, which solves a broad class of problems related to LIBS applications. The main limitation of this study is that with the chosen settings another broad class of LIBS applications, that is analysis of elemental composition, cannot be solved. The problem of classification addressed in the present study is certainly simpler than the problem of detailed analysis of the composition of materials. However, successful solution of this problem encourages further research in this area. We suppose, because the problem of analysis of composition of materials based on a LIBS spectrum may occur to be equivalent to the recognition of the presence of specific spectral lines, this problem can be brought to a series of classification subproblems.

5. Conclusions

The aim of this study is to solve the difficult problem of distinguishing paper-ink samples. Hence, we have designed a computational intelligence system to solve this problem using LIBS spectra. Difficulty of this problem is caused by the high variability of spectra within a single class of paper-ink samples and strong similarity of samples from different classes.

A machine learning system is developed and based on results presented in Table 3 and Table 4, we can conclude that the system performed its task for majority of LIBS spectra in a short time. The described method has sensitivity of 86% and it needs to be improved further. The main limitation of this method is that we do not list particular elements of the tested material, however it also should be taken under consideration in future work. Additional feature extraction and classification methods can be applied in the future.

Author Contributions: The work presented in this paper was a collaboration of all authors. The manuscript was revised by all authors. Conceptualization, M.K., K.Rz., T.S., and P.P.; Methodology, K.Rz., M.K., and P.P.; Software, T.S.; Validation, K.Rz., and P.P.; Investigation, M.K., K.Rz., T.S., and P.P.; Resources, M.K., M.N., and T.L.; Data Curation, M.K., T.S., and P.P.; Writing – Original Draft Preparation, M.K., K.Rz., and P.P.; Writing – Review & Editing, M.B., K.Rz., O.Y., P.P., and R.A.; Visualization, T.S., and P.P.; Supervision, K.Rz., P.P., and R.A.;
Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Acknowledgments: The authors would like to thank...


34. Team of Science and Industrial Intelligent Applications; subsection “Research” in.


**Sample Availability:** LIBS spectra used in this study are available on the webpage: [http://libs.iti.pk.edu.pl/](http://libs.iti.pk.edu.pl/).