Scientific paper

The Use of FT-MIR Spectroscopy and Counter-Propagation Artificial Neural Networks for Tracing the Adulteration of Olive Oil

Neva Grošelj,[1] Marjan Vračko,[1] Juan Antonio Fernández Pierna,[2] Vincent Baeten[1] and Marjana Novič[1]*

[1] *National Institute of Chemistry, Hajdrihova 19, Ljubljana, Slovenia*

[2] *Quality of Agricultural Products department, Walloon Agricultural Research Centre (CRA-W), Gembloux, Belgium*

** Corresponding author: E-mail: marjana.novic@ki.si*
Phone: + 386 1 4760 253; Fax: + 386 1 4760 300

Received: 17-04-2008

Dedicated to the memory of Professor Ljubo Golič

Abstract

The aim of this work is to detect the presence of refined hazelnut oil in refined olive oil, using the Counter-propagation Artificial Neural Networks (CP-ANN) model. The oil samples were analyzed by FT-MIR spectroscopy. They were classified as pure olive oil (Class 1), pure hazelnut oil (Class 2), and two type of adulterated olive oil samples, one with more than (or equal to) 10% of hazelnut oil (Class 3), and the other with less than 10% of hazelnut oil (Class 4). In addition, an external set of blind samples was also analyzed by FT-MIR.

Five CP-ANN models with different number of selected infrared spectral regions were built up and tested for their classification ability. On the basis of leave-one-out cross validation procedure the best models were selected and further used for blind samples prediction. The results obtained show that the models clearly separate different groups and classify correctly the pure olive oil and the hazelnut oil. Moreover a reasonable discrimination between both mixtures and pure oils was achieved.

Keywords: Counter-propagation artificial neural networks, hazelnut oil, MIR spectroscopy, olive oil adulteration.

1. Introduction

Infra red (IR) spectroscopy in combination with different statistical methods has been widely used for analysis of different kinds of commodity adulterations. Sivakesava and Irudayaraj[1] report the detection of inverted beet sugar adulteration of honey with the use of IR spectroscopy in combination with partial least squares (PLS) regression. Picque et al. applied the same methods for monitoring and control the different processes in the alcoholic and lactic fermentation[2]. The IR spectra of sugars in aqueous mixtures[3] and fruit juices[4] were used together with the principal component regression (PCR) and PLS to develop models for sugar content determination. Suchanek et al. used the mid-infrared spectra together with principal component analysis (PCA) and linear discriminant analysis (LDA) for the quantitative analysis of a variety of green coffee.[5] Lai et al. analyzed samples of vegetable oils, where FT-IR spectroscopy was used in conjunction with PCA and linear discriminant analysis to determinate the authenticity of vegetable oils[6]. In the last decades the neural networks have been extensively used as an alternative modeling technique for the statistical analysis of measured data. Zupan et al. report the models for classification of olive oils.[7] Olive oil samples were represented by fatty acids composition and the classes were associated with the land of origin (region) in Italy. Authors found that the Kohonen neural networks were superior to error-back propagation neural networks. Brodnjak-Vončina et al. studied samples of vegetable oils using the chromatographic data and neural networks as a method for statistical analysis.[8] Sayago et al. studied the samples of olive oil adulte-

rated with hazelnut oil with fluorescence spectroscopy.[9] PCA and error-back propagation neural networks in combination with genetic algorithm were used for variable selection and modeling. Oliveira et al. investigated diesel/biodiesel blends adulterated with vegetable oils. FT-NIR spectroscopy and FT-Raman data were analyzed with partial least square, PLS, and ANN methods. The best accuracy for the prediction for an external validation set was obtained with ANN model.[10]

The study of olive oil adulteration with hazelnut oils was initiated due to the lack of official analytical methods and due to the difficulty of detection of presence of hazelnut oil at low percentages (below 20%). To cover the odor of hazelnut oil, the blend with lampanate olive oil can be made. Since the lampanate olive oil should be refined before it can be sold, it is important for protection of customers to develop a method to detect refined hazelnut oil in the refined olive oil. This is a difficult task, because the triglycerides and the compounds of the unsaponifiable matter of these refined oils are very similar.[9] Beaten et al. studied the set of 189 samples of olive oils adulterated with refined hazelnut oil and with FT-Raman and FT-MIR spectroscopic techniques.[11] Stepwise discriminant analysis (SLDA) was applied for statistical analysis and discrimination. They analyzed the IR spectra of mixtures of olive oil and hazelnut oil with the stepwise linear discriminant analysis (SLDA). The study shows that the complete discrimination between olive and hazelnut oils is possible and that adulteration can be detected if the presence of hazelnut oil in olive oil is higher than 8% and if the blends are of Turkish olive and hazelnut oils. One of the purposes of this study was to select the regions in spectra, which are highly relevant for discrimination models. They selected ten wavelengths, which have been also used in the presented study.

In the present study we apply the Self Organizing Map (SOM) alias Kohonen neural network and Counter Propagation Artificial Neural Network (CP-ANN) as methods for clustering and classification. The methods are briefly described in section 2. The data and the technical details of models are described in section 3, while in section 4 we present the results and discussion.

2. Modeling Methods

Self-organizing maps (SOM) alias Kohonen neural networks and counter propagation artificial neural networks (CP-ANN) have been applied for cluster analysis and classification of samples.[12,13,14,15] In cluster analysis one investigates whether or not the objects of data set form groups, which are in some manner separated from each other. It is to emphasize that the clustering structure is an intrinsic property of the data set. On the other hand, in classification procedure the objects of the data set are ordered into predefined classes. The rules, which govern the classification, can be defined accordingly to any regulation, or can be defined accordingly to the cluster structure. One of available tools for cluster analysis is the SOM. Basically, the SOM represents a mapping of objects from multi-dimensional representation space into a two-dimensional array of neurons, which are indeed the vectors of weights. The dimension of the vector is equal to the dimension of the representation space that equals to the number of independent (input) variables. The total number of neurons, i.e., the dimension of the SOM, is optimized during the construction procedure.

The learning in SOM runs in two steps. In the first step the algorithm selects the neuron which weights are closest to the object ("winning neuron"). In the second step of learning the weights of the winning neuron are modified towards the object values and in the same time the neighboring neurons are modified to become similar to them. The learning runs over all elements of learning set what represents one learning epoch. The learning epochs are repeated until the weights are stabilized. A final result of the learning is organization of objects in two-dimensional network on the basis of similarity among them. If any cluster structure in the data set exists it is visualized as groups of objects in two-dimensional network. In practice the objects are labeled to be easily detected and recognized in the map.

The CP-ANN is a generalization of SOM. In comparison to SOM it has a two-layer structure. The first layer is input layer and has the same structure as SOM while the second layer is located beneath and is associated to the output variables (properties). The learning in CP-ANN has an additional step. The first step runs in the input layer and is the same as in SOM, i.e., the objects are arranged into the map accordingly to similarity relationship among them. In the second step of learning the positions of objects are projected from the input layer to the output layer and the weights there are modified to become equal to corresponding output values. The reader can find more details about architecture and learning strategy of SOM and CP-ANN in many textbooks and articles.[12,13,14,15]

The SOM and CP-ANN are suitable tools for cluster analysis and classification.[16,17,18] Visual inspection of objects organized in SOM enables one to recognize the clusters. On the other hand, the CP-ANN with multi-dimensional output layer can be used to construct the classification models (see Figure 1).

In such a case the dimension of the output layer is equal to the number of classes and "a property" of an object is described as a vector. For example, in classification system of four classes an object from third class is described with the vector (0, 0, 1, and 0). In the prediction phase, when a new object is presented to the model, the response is also a vector with elements expressed as real numbers between 0 and 1. The numbers represent the affiliation to a particular class. Several cases can occur:

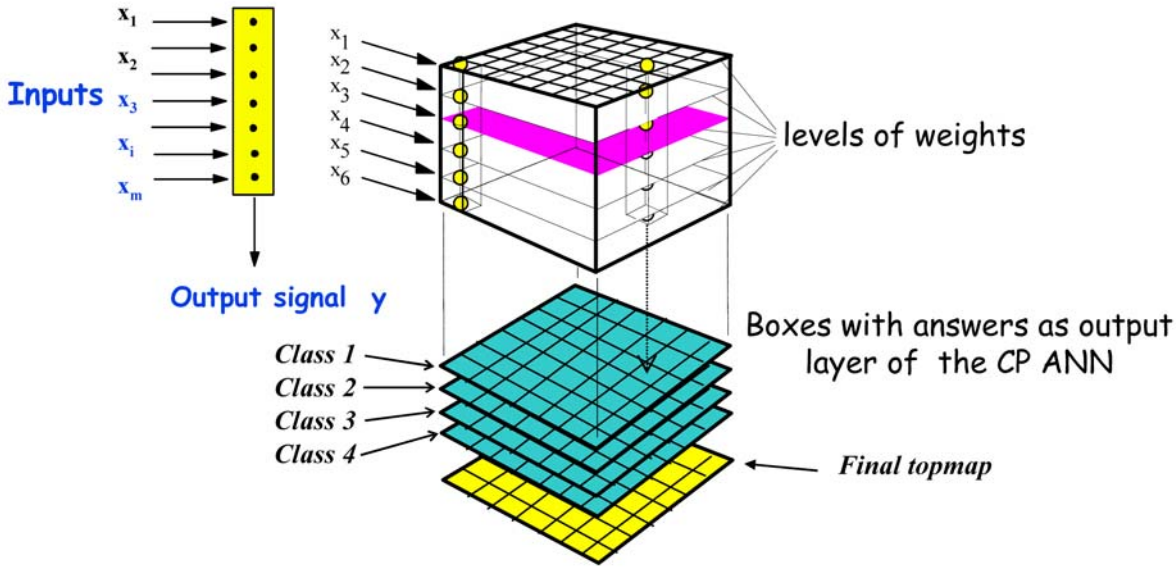

Figure 1: Architecture of the CP-ANN model suitable for classification into four classes.

– One element is essential larger than others, in this case the object is classified to the class corresponding to the position of the largest element.
– Two elements are approximately equal and different from others, in such uncertain classification the object is classified "between the two classes". It is our decision if we accept such a result.
– Three or all elements are approximately equal; in such case the model is not able to classify this object. We know *a priori* that the predicted result is wrong, what is valuable information.

3. Description of Data and Models

Data set consists of 119 samples of oils, for which the FT-MIR spectra were recorded. Each spectrum is represented with 609 variables, i.e., the values of response at 609 wave numbers in the range from 3074 cm^{-1} to 728cm^{-1}. The technical details are reported by Beaten et al.[11] Oils are classified in four classes: pure olive oil (50 samples), pure hazelnut oil (18 samples), and mixture (mixture_1) of olive oil adulterated with more (or equal) than 10% of hazelnut oil (29 samples), and mixture (mixture_2) of olive oil adulterated with less than 10% of hazelnut oil (22 samples). Additionally, we have an external set of 54 samples, for which the classification data were available only for 44 samples. It is to emphasize that this set was never used during the developing of models and is thus suitable for the validation of the developed models.

An important question in the model building is the selection of input variables. We investigated five cases. In the first modeling experiment the intensities at all 609 wave numbers were taken as input variables. In the models 2 to 5 we selected 56, 167, 330, and 393 intensities, respectively. The ranges of selected wave numbers for all models are presented in Table 1. The basis for this selection were ten wave numbers originally selected by Beaten et al.[11]

Table 1: Ranges of selected frequencies. Frequencies of model 1 are from reference Baeten et al.[11]

Model	Range of selected frequencies	Number of frequencies
1	728–3074	609
2	728–748	6
	968–999	9
	1072–1130	16
	1292–1462	13
	1624–1640	5
	1720–1743	7
3	728–783	15
	968–1041	20
	1072–1462	102
	1624–1677	15
	1720–1774	15
4	728–898	45
	948–1199	66
	1357–1462	28
	1647–1824	47
	2001–2553	144
5	728–898	45
	940–1850	237
	2250–2400	39
	2800–3000	72

An in-house program package has been used for modeling[19, 20]. Based on our experience in this field it was developed for chemometric applications and applied in several previous studies[21, 22, 23, 24]. The models were tested on recall ability (RA) and with the leave-one-out cross validation (LOO CV) test. The LOO CV test results were considered to set the technical parameters of the final mo-

del. The dimension of the network was varied from 5x5 to 8x8, the number of epochs from 50 to 500, and the maximal learning rate from 0.1 to 0.9. The following parameters were found as optimal: the dimension of the network: $N_x \times N_y = 6 \times 6$, the number of learning epochs: N_{epoch} 300, and the minimal and maximal learning rates: $\eta_{min} = 0.01$, $\eta_{max} = 0.5$. More details about the training are given in Table 2.

Table 2: Model with the dimension 6x6 was trained with 300 epochs.

	Answer	Recall ability	Leave-one-out
Model 1	*Correct*	68	40
609 variables	*False*	25	71
	Uncertain	26	8
Model 2	*Correct*	70	43
56 variables	*False*	19	71
	Uncertain	30	5
Model 3	*Correct*	**71**	**46**
167 variables	*False*	**30**	**66**
	Uncertain	**18**	**7**
Model 4	*Correct*	64	42
330 variables	*False*	20	71
	Uncertain	35	6
Model 5	*Correct*	70	37
393 variables	*False*	27	75
	Uncertain	22	7

4. Results and Discussion

It follows from Table 2 that the best model was obtained with 167 input variables (the third model). It gives for the RA test 71 correct answers, 30 false and 18 uncertain ones. For LOO CV test the numbers of correct, false and uncertain answers are 46, 66, and 7, respectively. The detailed predictions are given in Table 3.

One can observe that from 50 samples of pure olive oil 35 samples were predicted correct, 6 were predicted as mixture_1, 2 as mixture_2 and 7 uncertain. None of the samples was predicted as pure hazelnut oil. 18 samples of hazelnut oil were 17 times predicted correctly and once as mixture_2. 29 samples of mixture_1 were predicted 12 times correctly as mixture_1, two times as mixture_2, six times as uncertain, nine times as pure olive oil, and never as pure hazelnut oil.

Similar predictions are for mixture_2 (see Table 3). It is to emphasize that pure olive oil and pure hazelnut oil are completely separated, and also the hazelnut oil is well separated from mixtures. On the other hand, the separation between pure olive oils and mixtures is not completely achieved. Samples arranged into SOM are shown in Figure 2.

The squares (6 × 6) represent the neurons. The population of objects on each neuron is represented by two labels; the first one is related to a class ("oo" for pure olive oil, "ho" for hazelnut oil, "m1" for mixture_1, and "m2" for mixture_2).

The labels in brackets show how many of objects of a particular class are located on the neuron. For example, the neuron (1,1) at the down left corner is populated with three objects of pure olive oil, one object of mixture_1 and two objects of mixture_2. The neurons without labels are empty, i.e., they are not populated. Figure 3, which show the surfaces of four output layers (associated with the four classes), gives more insight into the distribution of objects.

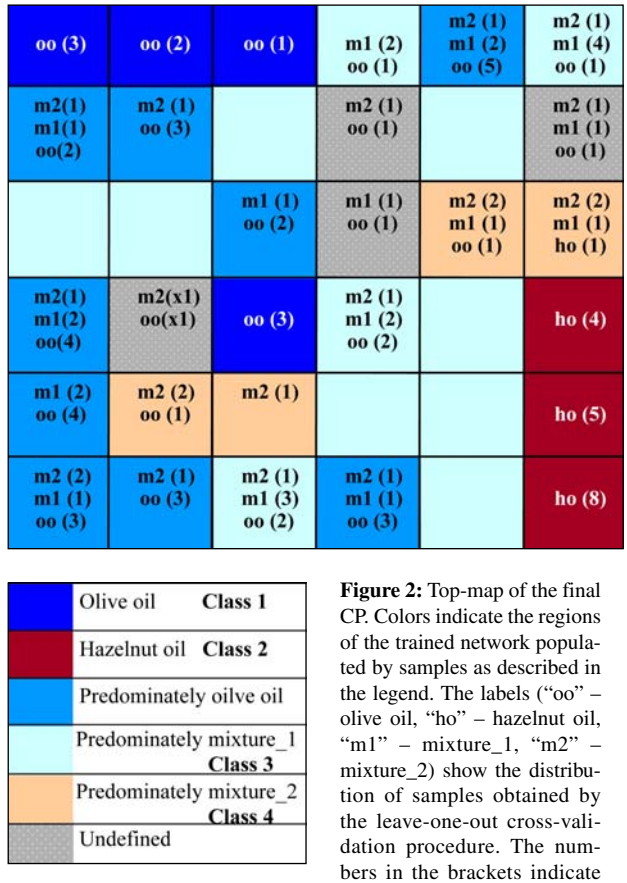

Figure 2: Top-map of the final CP. Colors indicate the regions of the trained network populated by samples as described in the legend. The labels ("oo" – olive oil, "ho" – hazelnut oil, "m1" – mixture_1, "m2" – mixture_2) show the distribution of samples obtained by the leave-one-out cross-validation procedure. The numbers in the brackets indicate the number of samples on a particular neuron. (see online version)

Table 3: Prediction matrix for the Model 3 for the leave-one-out test.

Experimental \ Predicted	Olive oil	Hazelnut oil	Mixture_1	Mixture_2	Uncertain
Olive oil	**35**	0	6	2	7
Hazelnut oil	0	**17**	0	1	0
Mixture_1	9	0	**12**	2	6
Mixture_2	7	0	3	**7**	5

The highest occupation of pure olive oil samples is top left corner, while the pure hazelnut oil samples are collected in down right corner (Figures 3a, b). The mixtures 1 and 2 are distributed in between (Figures 3c, d). One can see that the clusters of pure oils are well separated.

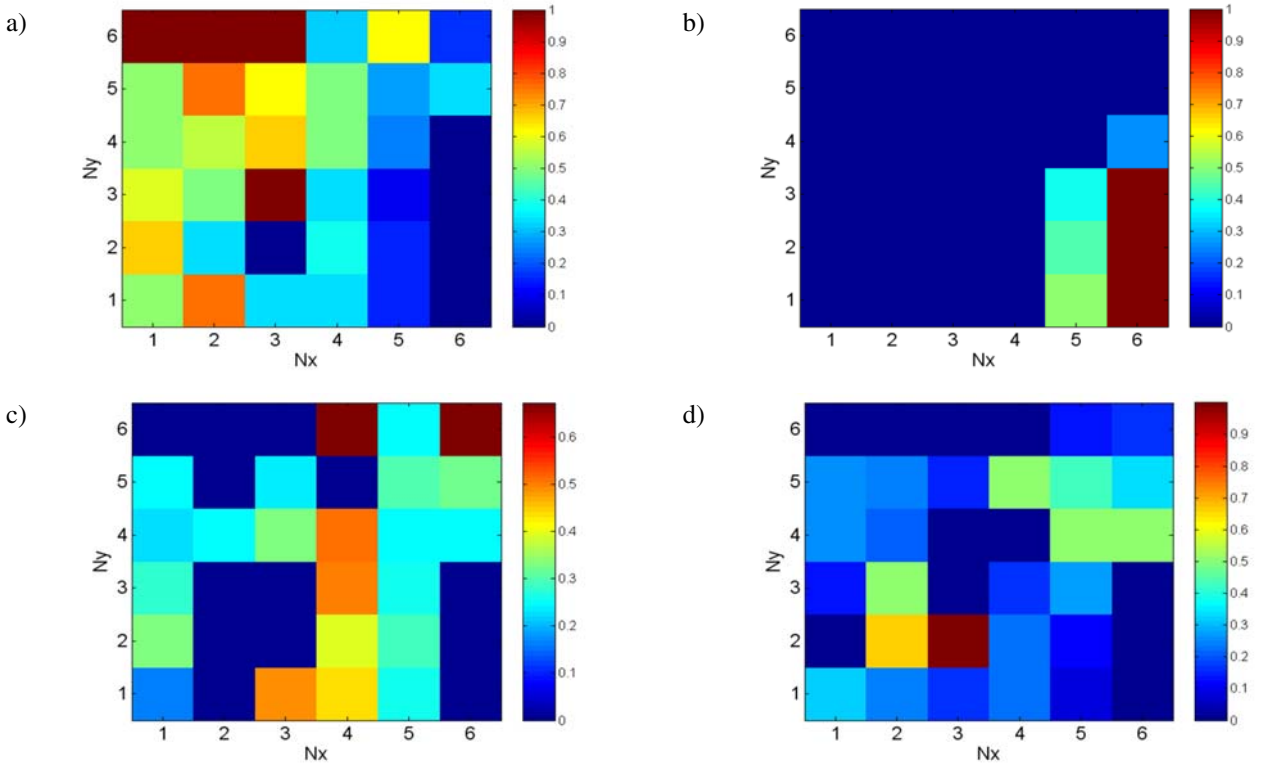

Figure 3: Four layers of the CP-ANN output layers. The numbers, which show the affiliation to a particular class, are represented by a color scale. (see online version)

The optimal model, which is described above, was used to classify the 54 samples of the external set. In the first step we analyzed model responses vectors (affiliations to classes). As an example Table 4 shows the classification results for seven samples. The boundaries between the classes can be determined for three different cases, (i) one element is essentially larger than other, (ii) two elements are of similar size, and (iii) three elements are of equal size. The sum of all four elements of the classification vector for prediction of four classes is always equal to 1.0. A trivial boundary is 0.5. If one of the four elements in the classification vector is over 0.5, the predicted class is unambiguous, i.e. the position of the largest component in the classification vector. We accept as a positive prediction also the situation, when the largest component is below 0.5, but it must be for more that 50% larger that any of the remaining components. The samples 143 and 167 are unambiguously classified as pure olive oil and mixture_2, respectively. Samples 102 and 165 are also clearly classified as mixture_1 and pure olive oil, respectively. The sample 156 is classified as olive oil, or mixtures 1 and 2. The samples 135 and 142 are classified between three classes, i.e. pure hazelnut oil, mixture 1, and 2. Further results of classification are shown in Table 4.

In the second step we analyzed the positions of samples in the SOM. The map is shown in Figure 4. The colors of digits, which label the samples, indicate the classification as described above. Five samples, which are classified as pure hazelnut oil, are located close to region of hazelnut oil (right down corner).

Table 4: Prediction of classes for seven samples of external data set

Id of the sample	Pure olive oil	Pure hazelnut oil	Mixture_1	Mixture_2
102	0.3270	0.0000	0.6730	0.0000
135	0.1407	0.4485	0.2885	0.1223
142	0.1001	0.3809	0.2531	0.2658
143	1.0000	0.0000	0.0000	0.0000
156	0.3801	0.0000	0.3981	0.2218
165	0.7570	0.0000	0.0000	0.2430
167	0.0003	0.0000	0.0002	0.9995

Around are samples of both mixtures (the first line of the map, the center of the map, and the last column of the map). In the upper left part are located samples classified as pure olive oils. We can compare the data with same experimental information, which is available for 44 samples. They are olive oils, which have been adulterated with 8% to 20% of hazelnut oil (see Table 5).

Table 5: The numbers of samples from the external data set adulterated with different concentrations (given in %) of hazelnut oils.

	Mixture_1						Mixture_2		Olive oil
Adulteration (%)	20	15	13	12	11	10	9	8	0
Number of blind samples	4	5	4	4	4	5	6	2	10

Table 6: Prediction matrix for the samples of external set.

Experimental \ Predicted	Olive oil	Hazelnut oil	Mixture_1	Mixture_2
Olive oil	**4**	–	3	3
Hazelnut oil	–	**0**	–	–
Mixture_1	4	–	**18**	4
Mixture_2	1	–	2	**5**

Twenty six of them belong to class mixture_1, eight to mixture_2, and ten of them are pure olive oils. Table 6 shows the prediction matrix for 44 samples. The model correctly excludes the class of pure hazelnut oil, i.e. none of the samples is classified as pure hazelnut oil. Majority of mixtures are also correctly predicted, while the largest error occurs in pure olive oil samples. Four out of ten olive oils are classified correctly, others as mixtures. However, it is obvious from Table 5 that the concentration of hazelnut oil increases gradually and therefore, the classification of a sample as mixture_1 or mixture_2 must be considered with caution. From the comparison of these results with the previous study[11] we can conclude, that the separation of pure hazelnut oil from other samples in the training set is reasonably successful. The detection of mixtures of blind samples was only partially successful in both studies, while the blind samples of pure olive oil were more often misclassified in this study than with the stepwise linear discriminant analysis (SLDA). Obviously, the applied neural network method, which is nonlinear and very sensitive, detected differences between the two groups of pure olive oil samples (see Fig. 3a, two regions with brown squares labelling the olive oil samples). The splitting of pure olive oil samples into two clusters enabled the mixture samples to be placed between these two regions, which enlarged the classification error. Our intention is to improve the results by selecting additional spectral regions (representing the variables for our model) with an automated variable selecion method.

5. Conclusions

The objective of this work was to build the model for classification of olive oil samples with respect to adulteration with hazelnut oil. The samples were represented by FT-MIR spectra and the CP-ANN method was taken as a modeling technique. Classification of samples was achieved using the proper architecture of the CP-ANN. In addition, we analyzed the organization of samples into clusters looking at the top-map of CP-ANN. This inspection can support the predicted classification and give an insight into similarity relationship between the samples. We achieved a firm separation between olive oil and hazelnut oil. On the other hand, the separation between mixtures and olive oil is only partially achieved due to the composition of mixtures and nature of measured data.

6. Acknowledgement

The authors acknowledge Slovenian Ministry of Higher Education, Science and Technology for the financial support (grant P1-017).

7. References

1. S. Sivakesava, J. Irudayaraj, Detection of Inverted Beet Sugar Adulteration of Honey by FTIR Spectroscopy, *J. Sci. Food Agric.* **2001**, *81*, 683–690.
2. D. Picque, D. Lefier, R. Grappin, G. Corrieu, Monitoring Of Fermentation By Infrared Spectrometry, *Analytica Chimica Acta* **1993**, *279*, 67–72.
3. S. Sivakesava, J. Irudayaraj, Determination of Sugars in Aqueous Mixtures Using Mid-Infrared Spectroscopy, *Applied Engin. Agric*. **2000**, *16*, 543–550.
4. N. Duputy, M. Meurens, B. Sombret, P. Legrand, J.P. Huvenne, Determination of Sugars and Organic Acids in Fruit Juices by FT Mid-IR Investigation of Dry Extract, *Applied spectroscopy* **1992**, *46*, 860–863.
5. M. Suchanek, H. Filipova, K. Volka, I. Delgadillo, A.N. Davies, Qualitative Analysis of Green Coffee by IR Spectrometry, *Fresenius J. Anal Chem* **1996**, *354*, 327–332.
6. Y. W. Lai, E. K. Kemsley, R. H. Wilson, Potential of Fourier Transform Infrared Spectroscopy for the Authentication of Vegetable Oils, *J Agric. Food Chem.* **1994**, *42*, 1154–1159.

7. J. Zupan, M. Novič, Liu, J. Gasteiger, Classification of multicomponent analytical data of olive oils using different neural networks *Anal. Chim. Acta* **1994**, *292*, 219–234.
8. D. Brodnjak-Vončina, Z. Cencič-Kodba, M. Novič, Multivariate data analysis in classification of vegetable oils characterized by the content of fatty acids, *Chemometr. intell. lab. syst.* **2005**, *75*, 31–43.
9. A. Sayago, D. L. Garcia-Gonzales, M. T. Morales, R. Aparicio, Detection of the presence of refined hazelnut oil in refined olive oil by fluorescence spectroscopy, *J. Agric. Food Chem*. **2007**, *55*, 2066–2071.
10. F. C. Oliveira, C. R. Brandao, H. F. Ramalho, L. A. da Costa, P. A. Suarez, J. C. Rubim, Adulteration of diesel/biodiesel blends by vegetable oil as determinated by Fourier transform (FT) near infrared spectroscopy and FT-Raman spectroscopy, *Anal. Chim. Acta*. **2007**, *587*, 194–199.
11. V. Baeten, J. A. F. Fernandez Pierna, P. Dardenne, M. Meurens, D. L. Garcia-Gonzalez and R. Aparicio-Ruiz, Detection of the Presence of Hazelnut Oil in Olive Oil by FT-Raman and FT-MIR Spectroscopy, *J. Agric. Food Chem*. **2005**, *53*, 6201–6206.
12. T. Kohonen T., *Self-Organization and Associative Memory Third Edition*, SpringerVerlag, Berlin, FRG, 1989.
13. R. Hecht-Nielsen, Counterpropagation Networks, *Appl. Optics.* **1987**, *26*, 4979–4984.
14. J. Zupan, J. Gasteiger, *Neural Networks in Chemistry and Drug Design*, Wiley-VCH, Weinheim, 1999.
15. J. Zupan, *Basics of artificial neural network, in Nature-inspired methods in chemometrics: Genetic algorithms and artificial neural networks*, Editor: Riccardo Leardi, *Elsevier*, Amsterdam, **2003**, pp 199–229.
16. J. Zupan, M. Novič, I. Ruisanchez, Kohonen and counterpropagation ANNs in analytical chemistry, *Chemom. Intell. Lab. Syst* **1997**, *38*, 1–23.
17. M. Vračko, Kohonen Artificial Neural Network and Counter Propagation Neural Network in Molecular Structure-Toxicity Studies, *Current Computer-Aided Drug Design* **2005**, *1*, 73–78.
18. M. Novič, M. Vračko, *Artificial neural networks in molecular structure-property studies, in Nature-inspired methods in chemometrics: Genetic algorithms and artificial neural networks,* Editor: Riccardo Leardi, *Elsevier*, Amsterdam, **2003**, pp 199–231–256.
19. ANN-train program, ANN-test program, ANN-LOO program, Laboratory of Chemometrics, National Institute of Chemistry, Ljubljana, Slovenia, http://www.ki.si/index.php?id=185&no_cache=1&L=1
20. ANN-visual program, Laboratory of Chemometrics, National Institute of Chemistry, Ljubljana, Slovenia, http://www.ki.si/index.php?id=185&no_cache=1&L=1
21. M. Novič, J. Zupan, 2-D mapping of infrared spectra using Kohonen neural network. *Vestn. Slov. kem. druš*., (Documenta chemica Yugoslavica) **1992**, *39 (2),* 195–212.
22. J. Zupan, M. Novič, I. Ruisánchez, Kohonen and counterpropagation artificial neural networks in analytical chemistry : tutorial. *Chemometr. intell. lab. syst*. [Print ed.], **1997**, *38*, 1–23.
23. E. Maran, M. Novič, P. Barbieri, J. Zupan, Application of counterpropagation artificial neural network for modelling properties of fish antibiotics. SAR QSAR, *Environ. Res.*, **2004**, *15(5/6),* 469–480.
24. E. Vončina, D. Brodnjak-Vončina, N. Sovič, M. Novič, Chemometric characterisation of the quality of ground waters from different wells in Slovenia. *Acta chim. slov.*, **2007**, *54(1)*, 119–125.

Povzetek

Model protitočnih nevronskih mrež (CP-ANN) smo razvili z namenom, da bi določili prisotnost rafiniranega lešnikovega olja v rafiniranem oljčnem olju. Vzorci so analizirani z FT-MIR spektroskopijo. Spektri vzorcev so razvrščeni v štiri razrede, in sicer kot vzorci čistega oljčnega olja (razred 1), vzorci čistega lešnikovega olja (razred 2), vzorci oljčnega olja z dodatkom deset ali več odstotkov lešnikovega olja (razred 3) in vzorci oljčnega olja z dodatkom manj kot deset odstotkov lešnikovega olja (razred 4). Določili smo tudi niz neznanih vzorcev, ki smo jih analizirali po isti poti.
V nadaljnjem postopku smo pripravili pet CP-ANN modelov, tako da smo upoštevali različna spektralna območja z različnim številom vrhov. Modele smo najprej ocenili glede na njihovo zmožnost razvrščanja vzorcev v razrede s testom razpoznavanja, potem pa še z navzkrižnim testom z izpuščanjem po en vzorec, tako imenovanim "leave-one-out" testom. Na ta način smo izbrali najboljši model, ki smo ga uporabili za napoved neznanih vzorcev. Dobljeni rezultati kažejo, da model popolnoma loči vzorce čistega oljčnega in lešnikovega olja, prav tako smo dobili zadovoljivo razvrstitev obeh mešanic v ustrezni razred.