Multicomponent Analyzer of Volatile Compounds Characterization Based on Artificial Neural Networks

Taras Chaikivskyi[1[0000-0002-1166-8749]], Bogdan Sus[2[0000-0002-2566-5530]], Oleksandr Bauzha[2[0000-0002-4920-0631]] and Sergiy Zagorodnyuk[2[0000-0003-3415-7746]]

[1]Lviv Polytechnic National University, Bandera str, 12, Lviv, 79013, Ukraine
taras.v.chaikivskyi@lpnu.ua

[2]Taras Shevchenko National University of Kyiv, Volodymirskaya str., 64, Kyiv 01033, Ukraine
bnsuse@gmail.com, asb@mail.univ.kiev.ua, kola@univ.net.ua

Abstract. An automated multicomponent system for measuring the content of monatomic phenolic compounds in alcohol solutions has been developed. To assure the quality control of widespread alcohol solutions, including medications and strong alcoholic beverages, a new multi-sensor automated quality monitoring system has been proposed. The system is based on a multi-layer artificial neural network for the digital processing of sensor signals. Taking into account the cross-sensitivity of the sensors and the application of selective sensors reduce the error in determining the concentration of light phenolic compounds. The sensor signal processing process is complemented by parallel channels for calculating the ethyl alcohol concentration as well as the illumination and temperature values by linearly converting the output signals. Simulation models using sensors based on the electronic theory of impurity sorption on the surface of semiconductors were acquired.

Keywords: semiconductor sensor, neural network, volatile solutions, microcontroller.

1 Introduction

The application of semiconductor sensors as working elements in multi-sensor systems makes it possible to effectively solve typical problems of modern analytical chemistry - to analyze the compositions of liquids and gases. Such analyzing systems, also known as "electronic tongues" and "electronic noses," work with chemical sensors of different operating principles and biosensors in combination with multi-step mathematical algorithms for data processing. Therefore, the development and implementation of new analytical systems is an urgent technological challenge and is fully in line with global trends. The practical principle of the electronic nose is best explained through its natural counterpart. Volatile compounds reach the olfactory epithelial tissues, where they reach the olfactory receptors. This generates a corresponding neuron, which transmits information about this fact to the brain. In it, the information of neurons is consolidated and structured according to existing patterns that enable the subject to recognize a specific odor [1].

Preferably, there is a need to make extensive application of artificial electronic sensor systems during manufacturing products on an industrial scale. In particular, the control and monitoring of the processes of growing, maturing and spoiling organic raw materials are extremely important and valuable for ensuring the safety and quality of food products [2]. In the production of wine and spirits for quality control and identification of the aroma mandatory state certification for alcoholic beverages is used [3]. Today, the quality of the wine is mainly evaluated by physical and chemical indicators, such as alcohol content and subjective evaluations, as expert tasters. Obviously, this procedure, which is performed over a long period of time and requires expert evaluation, has a fairly high cost. Analysis of light volatile compounds released by alcoholic beverages can help to classify its specific characteristics, to identify in a timely manner the features and complexities of processing and long-term storage. Chemical analytical methods also require long and complex investigations, but in the result of the accumulation and structuring of the results of the data, these methods give higher reliability and accuracy of measurements than the subjective evaluation of a strong drink by a specialist taster [4]. Most of the chemical components that provide the consumable quality of the beverage can alter the signals of the sensors, and therefore provide the beverage categorization according to qualitative and quantitative criteria.

The presence and concentration of light volatile compounds in alcoholic beverages are also monitored by means of multi-component semiconductor gas analyzers. Semiconductor sensor systems provide the ability to control the concentrations of several lightweight volatile compounds. But concentration measurement gives a relatively high error, which limits their scope. This error is related to the cross-sensitivity of the gas analyzers. One method of reducing this error is to design and use highly selective sensor systems. They are an important component of the concept of "lab-on-a-chip" or "micro-total-analysis". This is the concept of making miniature devices that enable the sequencing of chemical processes or chain chemical reactions on a single chip with a small area. In addition, state-of-the-art lab-on-a-chip devices provide real-time chemical or quantitative chemical analysis [5].

Semiconductor gas-sensitive sensors have a low response time, but they are also sensitive to other physical parameters, including light, temperature, humidity, and background concentrations in the air of other organic volatile compounds. The effect of these factors has a significant effect on the measuring signal. Currently, several technological approaches are being used to minimize the impact of these factors. These approaches include the method of temperature compensation for gas-sensitive sensors and the pulse mode of sensor heating. By applying these methods, the effect of evaporation of water and ethanol on the sensitivity of the sensors was significantly reduced [6]. In this work, the measuring cell was purged for 10 additional minutes with pure helium to completely evaporate and dry the sensors before taking the measurements. It should be noted that such a procedure is technologically complicated and costly. Such solutions do not always provide the required level of error and may have a low reaction time for gas analyzers. Some studies have shown that under conditions of light irradiation, the photocurrent of a semiconductor strongly depends on the recombination characteristics and the charge state of the surface. The approach of using solar cells as sensory structures is interesting. Induced light flux (LBIC) measurements reflect the spatial distribution of the photocurrent of a solar cell. Depending on the diameter of the excitation beam and the distance, a lateral resolution of several 10-4 m to 10-6 m can be achieved. The concept of the LBIC sensor system is described in [7]. The conductivity of the sensitive layer of a semiconductor sensor depends on the concentration of free electrons, which is proportional to the fraction of the surface directly covered by the molecules of the adsorbed light compound.

However, even such sensors also have low selectivity and significant measurement error. The responses of each low-selective sensor are slightly dependent on the type of adsorbent. One option to increase selectivity is to increase the number of sensors and create their array. The set of multiple responses to the sensory system forms a unique imprint for each substance composition. The artificial neural network apparatus can be used to detect the composition of volatile compounds in sensor systems. A computer program can evaluate the signal pattern and can compare the flavors of different samples. The measurement results are comparative, not quantitative, and therefore presented as an "imprint".

Currently, artificial neural networks are used in many fields of science and technology. They are powerful tools for solving complex problems whose answers are not obvious and cannot be determined by classical mathematical algorithms. Neural networks have been increasingly used since the introduction of the backpropagation algorithm [8]. The study of fingerprint recognition using multisensory data is an area where rapid technological development of the data collection system is required, which are quantitative indicators of human quality of life. In particular, multiple sensor data collected from a smart home using several deep neural networks were investigated in [9] Similarly, the application of machine learning methods, including linear regression, neural network, and vector machine support, to determine the dependence of wine quality on the aromatic content and to predict wine quality were explored in [10]. The paper demonstrates that the value of a dependent variable can be more accurately predicted if only the important features of the wine are taken into account in the forecasting but not all the features are taken into account. In this case, a linear regression was implemented to determine the dependence of wine quality on the various 11 physicochemical characteristics. Wine quality assessment is one of the key elements in this work and this assessment can be used for product certification. This type of quality certification helps to ensure long-lasting quality of wine and its competitiveness. The wine has different characteristics, including density, pH, strength, acidity, etc. Wine quality can be assessed by two types of tests; the first is a physicochemical test and the second is a sensory test [11].

Low-accuracy electronic sensor systems were used for pattern recognition and selectivity. Artificial neural networks can also be effectively used to increase the selectivity of semiconductor sensor systems. Therefore, the creation of multi-component gas analyzers based on neural networks that provide simultaneous quantitative analysis to determine the concentrations of components, increase selectivity and reduce sensitivity to environmental factors are and remain relevant technical problems of analytical chemistry.

2 Hardware Setup

A block diagram of a multicomponent analyzer measuring complex for volatile compounds is shown in Fig. 1.

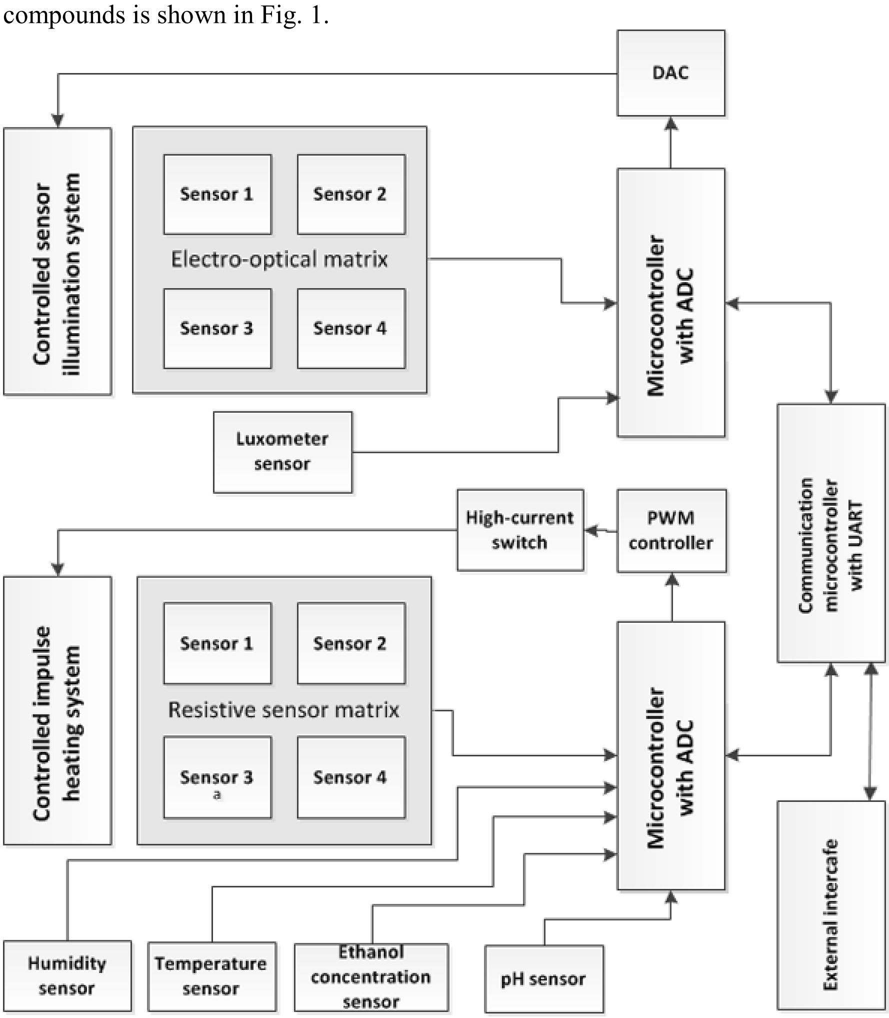

Fig. 1. Structural diagram of the measuring complex.

The measuring chamber is shown in fig. 2.

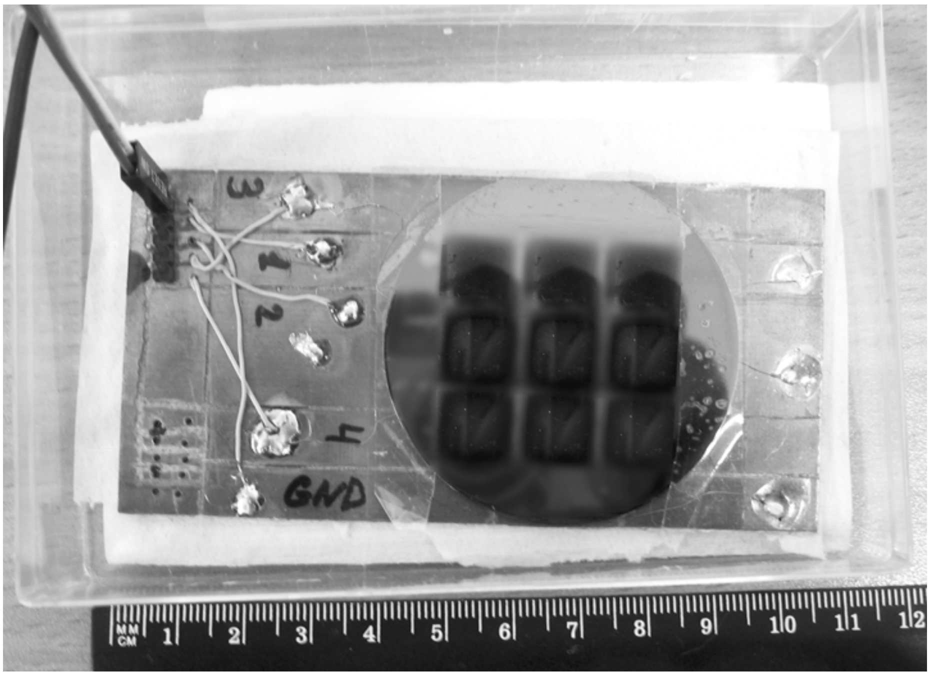

Fig. 2. Measuring chamber with Optical Conversion Sensor.

Nowadays, the most promising and convenient are sensor systems based on compact digital devices - microcontrollers equipped with built-in ADCs. The complex consists of sensors, sensitive to light volatile compounds and sensors of environmental parameters, microcontrollers of control and preprocessing of signals, and also the communication microcontroller of communication. The microcontroller (1) checks the illumination brightness of the sensor plate in an assigned intensity range. Brightness values are obtained from the light sensor. Brightness ratios and corresponding photoresponses of different samples are stored by the neural network. ADC microcontrollers convert analog sensor signals into digital signals, perform signal correction and normalization, control sensor backlight modes with optical re-creation and change of temperature measurement modes. The communication microcontroller receives data from ADC microcontrollers and processes the data according to algorithms implemented by the artificial neural network. Additionally, software and hardware encryption of the accumulated data may be implemented [12]. The developed setup can also operate with wireless data modules and be practiced for laboratory work during the learning process [13-15].

When using a 4-channel barrier sensor, a bias voltage should be applied to the p-n junctions. The photocurrent is determined by the rate of superficial recombination of charge carriers. Each channel has a distinct sensitivity to analytes. The equivalent circuit of the sensor is shown in fig. 3.

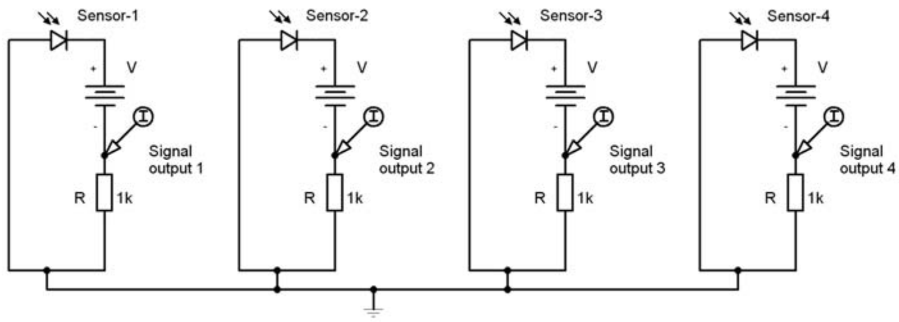

Fig. 3. Equivalent circuit of 4-channel sensor with p-n junction

Since the sensor measurement data also depend on the illumination brightness of the p-n junctions, the illumination intensity varies linearly with the DAC converter in the specified range during the measurements. There may be several photocurrent extremes in the range of illumination intensity, which could be used as additional informative parameters.

3 Experiment

Experiments with a set of substances of light alcohol compounds were carried out. The evaporation of the substances was above the sensors of the electronic nose. The sensors were illuminated by the incident light of a given intensity. The results of photocurrent measurements from 4 electronic nose sensors are shown in Fig. 4.

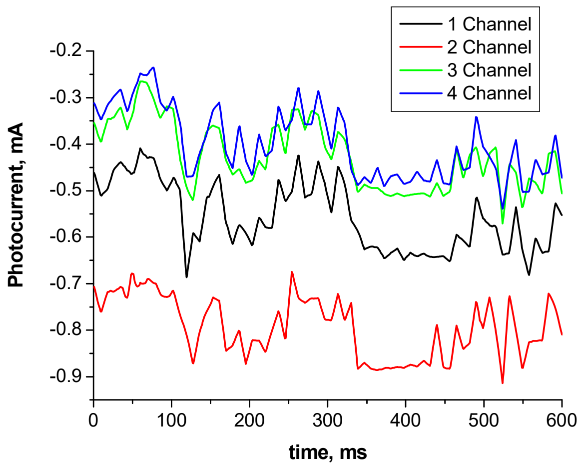

Fig. 4. Results of photocurrent measurements (sample # 1).

The dependencies of the sensor data from different samples are shown in Fig. 4.. The normalized values of the photocurrent are shown on the X-axis.

The currents density passed into the measurement sampling is shown on the Y-axis, Fig.5. Photocurrents from different channels are shown in different colors. As can be seen in the Fig. 5., the values on the sensors from different channels overlap.

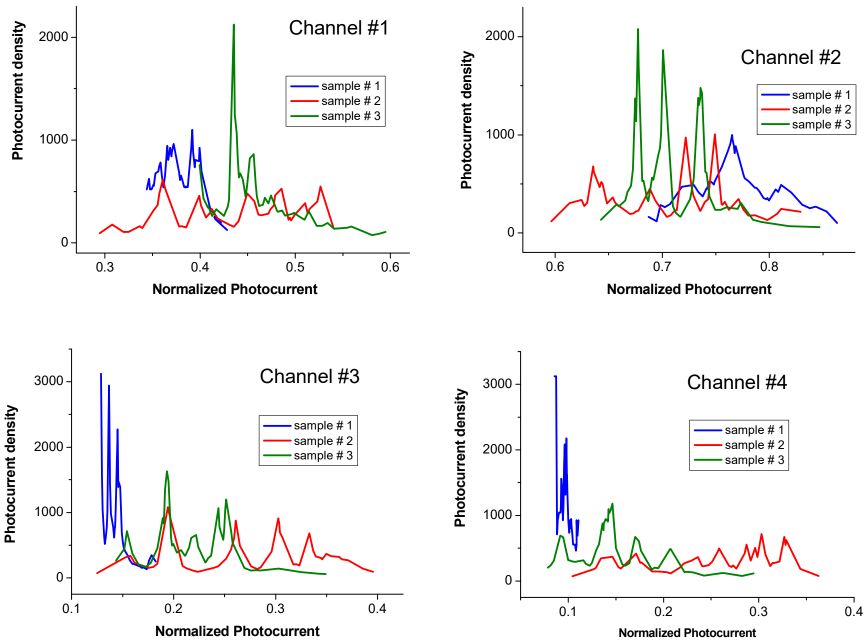

Fig. 5. Sampling of the sensor density measurement results.

4 Software implementation

During the measurements of different samples of alcoholic beverages, it turned out that the results of the sensors for the related samples have similar values. The expected correlation of the sensors with the specimens of the test substances (cognacs) was not observed. Obviously, under such conditions, it is advisable to use an artificial neural network to classify substances. The multilayered, fully connected feedforward neural network architecture has been selected. The artificial neural network receives sets of values of variables, specifically the values of equivalent concentrations of the mixture components of volatile compounds from optical and resistive sensor matrices. The normalized values of the concentrations of the components of the volatile compounds are calculated for each set of input variables. Values depend on the normalized output signals of the sensors. The input layer of the neural network consists of 8 input neurons. Input signals were normalized to 1 and reflected 4 channels of photocurrents from electronic nose sensors. The neural network contains three hidden layers, each of which has 10 neurons. The output layer has 6 neurons. The structure of the neutron network is presented in Fig. 6. This artificial network allows increasing the required number of neurons in layers.

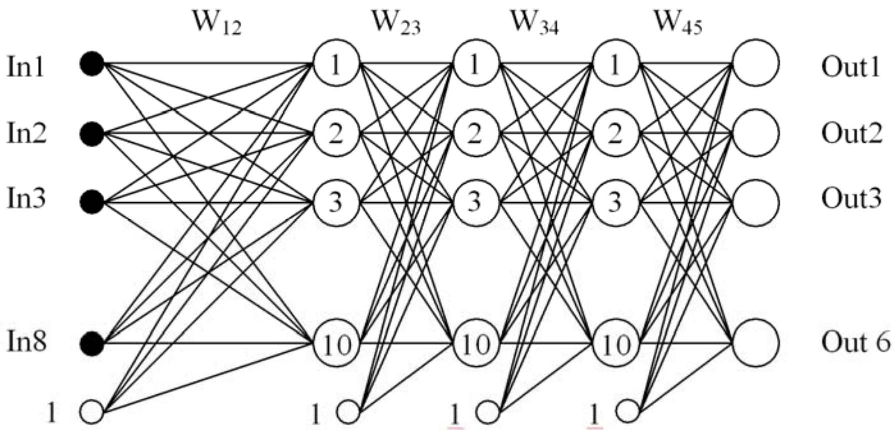

Fig. 6. Multilayer feedforward artificial neural network.

The sigmoid curve was chosen as the activation function.

$$f(x) = \frac{1}{1+e^{-x}}$$

Artificial neural network training was carried out by the method of training with the teacher. To determine the minimum of the error function, the backpropagation method in stochastic gradient descent algorithm was used. The algorithm of artificial neural network training is shown in Fig. 7. During training, the signals In, Out - are the input and output vectors of the neural network, respectively. The W array is a matrix of neural network weights.

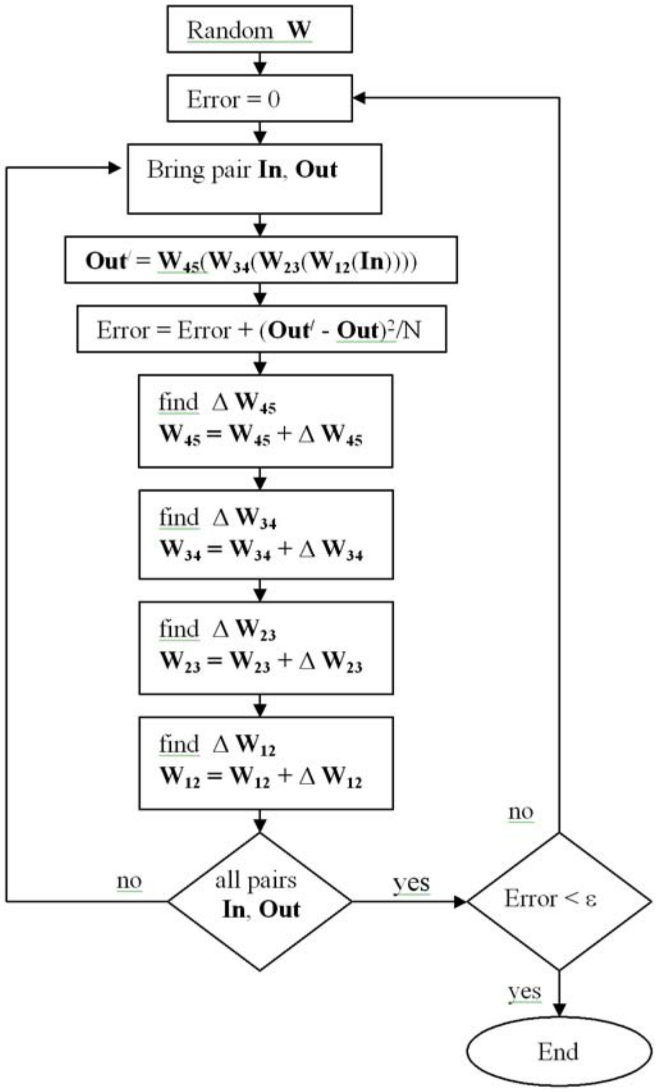

Fig. 7. Backpropagation algorithm.

The user interface of the sensor software with the neural network implementation is shown in Fig. 8.

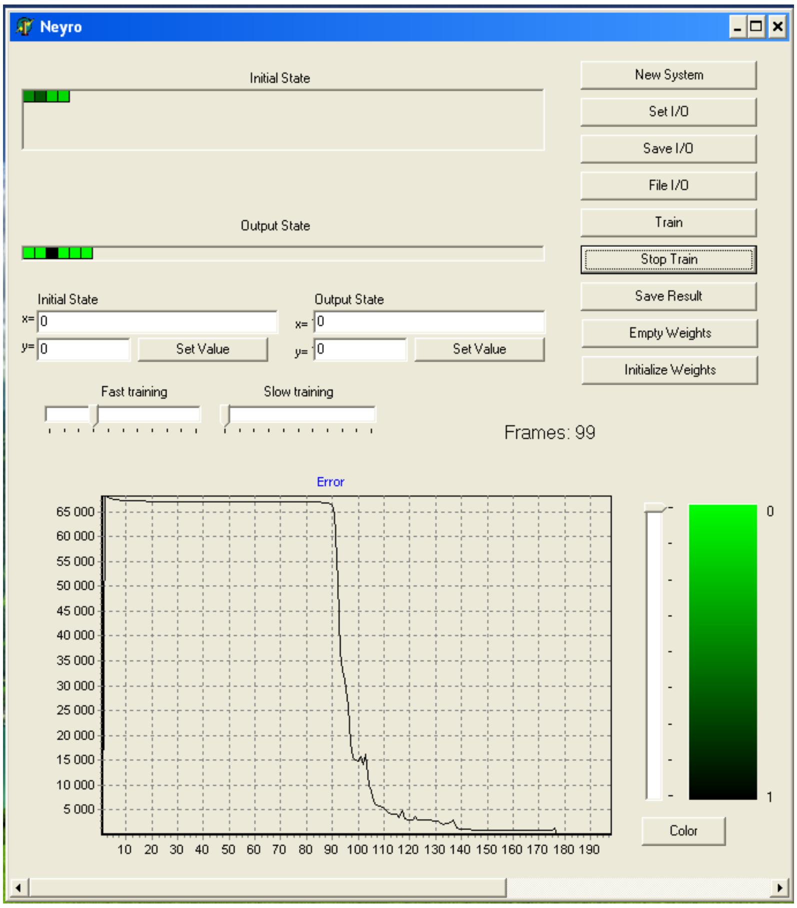

Fig. 8. Program interface

The "New System" button allows the user to configure the neural network. During configuration, it is possible to specify the number of input, output neurons and the number of neurons in each of the 3 intermediate layers. The "Set I / O", "Save I / O", "Set Value", and "File I / O" buttons define the input vectors of the input and output neurons (frames). The process of training the neural network begins with the "Train" button. The Empty Weights button sets the weight of the W_{ij} training to its initial, random position. The "Initialize Weights" button changes the learning scales and exits the system from the point of local minimum error.

The values of the input and output neurons in the frame are represented at the top left in the form of a matrix of rectangles. The color of the rectangles is determined by the value of the corresponding neurons. The scale for matching (and changing) the color and meaning of the neurons are shown at the bottom right.

In Fig. 8. it is shown that the neural network has finished the training and the schedule of the error change can be observed below.

5 Experimental data processing

The experiment was carried on three sample substances. For each substance, 51 measurements of photocurrent were executed, and therefore 153 measurements of photocurrent from three samples were executed, in the future, these measurements are called frames. The influence of time and the influence of the measurement history from the sensors on the current measurement result were not taken into account. A sample of 99 frames was randomly selected from 153 frames, 33 frames from each sample for neural network training. The remaining 54 frames, 16 frames from each sample are left as controls to test the operation of the already trained neural network. The neural network was trained for about 200 epochs, after which the total error was less than $\varepsilon = 0.01$, which was admitted a positive result. After training, the system was tested on control frames. The test gave 98% correct results. In 2% of cases, the non-authentic odor (with indicators of 50-92%) was the main one, while the authentic odor (with indicators of 8-50%) took second place. The result demonstrates the possibility of a system error (due to a lack of training frames) and demonstrates the need to consider not only the winners (first place) but also those source neurons that took second place.

In turn, among 98% of the correct results, 2% of cases were the authentic odor (82-93%), while the non-authentic odor (7-18%) ranked second. In the rest (96%) of control shots, the authentic smell was the undisputed leader (more than 99% probability).

6 Conclusions

During creating automated multicomponent systems for measuring the concentration of mixtures of volatile compounds, the processes occurring on the surface of the semiconductor sensitive layer as the temperature, light, and humidity of the environment are taken into account.

The use of multicomponent low-selection sensors gives the possibility to significantly reduce measurement errors in determining the concentrations of light mixtures, including when exposed to external disturbances.

For the training of neural networks, it is necessary to create and develop specialized mathematical models that describe the transformation processes in the used sensors. With the help of a neural network, sensor parameters can be normalized and calibrated to increase selectivity and to predict complex integral features in alcohol solution samples.

When considering a component that is recognized by a neural network, not only the winner neurons from the set of source neurons but also the source neurons that received the second result should be considered.

An increase in the number of measurement samples leads to an increase in neural network complexity, as well as an increase in the number of sensors needed to accurately classify the test substances.

Many substances have a specific range of odors that can confuse the neural network and generate incorrect neuron at the output, and therefore the correct classification of substances by constituents is an important technical task.

In the multi-component analytical system implemented, optimal conditions for recognizing the concentration of monatomic phenolic compounds in alcohol solutions were determined and formulated.

The main areas of further research are the development of methods for determining the age of strong drinks, including the justification for the use of markers and their linear combinations to improve the accuracy and reliability of the result, as well as the transfer of the measuring complex to the platform of programmable digital integrated circuits FPGA.

References

1. Lvova, L., Kirsanov, D.: Editorial: Multisensor Systems for Analysis of Liquids and Gases: Trends and Developments. Front. Chem. 6, 591 (2018). https://doi.org/10.3389/fchem.2018.00591.
2. Cuypers, W., Lieberzeit, P.A.: Combining Two Selection Principles: Sensor Arrays Based on Both Biomimetic Recognition and Chemometrics. Front. Chem. 6, 268 (2018). https://doi.org/10.3389/fchem.2018.00268.
3. Macías, M., Manso, A., Orellana, C., Velasco, H., Caballero, R., Chamizo, J.: Acetic Acid Detection Threshold in Synthetic Wine Samples of a Portable Electronic Nose. Sensors. 13, 208–220 (2012). https://doi.org/10.3390/s130100208.
4. Waterhouse, A.L., Ebeler, S.E. eds: Chemistry of wine flavor. American Chemical Society ; Distributed by Oxford University Press, Washington, DC : [New York] (1998).
5. Ghafar-Zadeh, E., Sawan, M.: CMOS capacitive sensors for lab-on-chip applications: a multidisciplinary approach. Springer, Dordrecht ; New York (2010).
6. Guadarrama, A., Fernández, J.A., Íñiguez, M., Souto, J., de Saja, J.A.: Array of conducting polymer sensors for the characterisation of wines. Analytica Chimica Acta. 411, 193–200 (2000). https://doi.org/10.1016/S0003-2670(00)00769-8.
7. Litvinenko, S.V., Kozinetz, A.V., Skryshevsky, V.A.: Concept of photovoltaic transducer on a base of modified p–n junction solar cell. Sensors and Actuators A: Physical. 224, 30–35 (2015). https://doi.org/10.1016/j.sna.2015.01.014.
8. Rumelhart, D.E., McClelland, J.L., University of California, S.D., PDP Research Group: Parallel distributed processing. explorations in the microstructure of cognition. (1987).
9. Park, J., Jang, K., Yang, S.-B.: Deep neural networks for activity recognition with multi-sensor data in a smart home. In: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT). pp. 155–160. IEEE, Singapore (2018). https://doi.org/10.1109/WF-IoT.2018.8355147.
10. Gupta, Y.: Selection of important features and predicting wine quality using machine learning techniques. Procedia Computer Science. 125, 305–312 (2018). https://doi.org/10.1016/j.procs.2017.12.041.
11. Ebeler, S.E.: Linking Flavor Chemistry to Sensory Analysis of Wine. In: Teranishi, R., Wick, E.L., and Hornstein, I. (eds.) Flavor Chemistry. pp. 409–421. Springer US, Boston, MA (1999). https://doi.org/10.1007/978-1-4615-4693-1_35.

12. Tmienova, N., Sus, B.: Hardware data encryption complex based on programmable microcontrollers. In: CEUR Workshop Proceedings, pp. 199–208 (2018). http://ceur-ws.org/Vol-2318/paper17.pdf
13. Chaikivskyi, T., Bauzha, O., Sus, B. B., Tmienova, N., Zagorodnyuk, S.: 3D simulation of virtual laboratory on electron microscopy. In: CEUR Workshop Proceedings 2533, pp. 282-291. (2019). http://ceur-ws.org/Vol-2533/paper26.pdf
14. Sus, B., Tmienova, N., Revenchuk, I., Vialkova, V.: Development of Virtual Laboratory Works for Technical and Computer Sciences. In: Damaševičius, R. and Vasiljevienė, G. (eds.) Information and Software Technologies. pp. 383–394. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-30275-7_29.
15. Bauzha, O., Sus, B., Zagorodnyuk, S., Stuchynska, N.: Electrocardiogram Measurement Complex Based on Microcontrollers and Wireless Networks. In: International Scientific-Practical Conference on Problems of Infocommunications Science and Technology, PIC S and T, pp. 345-349. (2019).