Multilayer GMDH-neuro-fuzzy System Based on Extended Neo-fuzzy Neurons and Its Learning in Online Facial Expression Recognition

Yevgeniy Bodyanskiy
Control Systems Research Laboratory
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
yevgeniy.bodyanskiy@nure.ua

Yuriy Zaychenko
Institute for Applied System Analysis
Igor Sikorski Kyiv Polytechnic Institute
Kyiv, Ukraine
zaychenkoyuri@ukr.net

Nonna Kulishova
Media Systems and Technologies Department
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
nokuliaux@gmail.com

Galib Hamidov
Information Technologies Department
Igor Sikorski Kyiv Polytechnic Institute
Azershig, Azerbaijan
galib.hamidov@gmail.com

***Abstract*— Real-time image recognition is required in many important practical problems. Interaction with users in online mode requires flexibility and adaptability from applications. The Group Method of Data Handling (GMDH) allows changing the model structure and adjusting the system architecture to the characteristics of each task under consideration. Moreover, the approximating properties of neo-fuzzy neurons used as elements of the system provide high recognition accuracy under conditions of short data samples. This paper proposes a multilayer GMDH-neuro-fuzzy system based on extended neo-fuzzy neurons. The learning algorithm has filtering and tracking properties, guaranteed the required speed important for real-time applications. The effectiveness of the proposed system is confirmed for the human emotions recognition.**

Keywords*— *Group Method Data Handling; extended neo-fuzzy neuron; online image recognition; facial expression recognition

I. INTRODUCTION

Information technologies are actively being introduced into education, business, healthcare, entertainment and other spheres of human life. This requires technology interactivity, to conduct continuous two-way cooperation between person and computer or mobile device. One of the promising areas for such intellectual interfaces' development is the approach that uses the recognition of people, their age, sex, state of health, emotional status on the real time video. This complex technical problem already finds its own solutions [1-9]. Frequently, these decisions use the machine learning and neuro-fuzzy approach.

As a technical problem, the task of a user emotional status recognition by video is reduced to characteristic features detecting, and to the collected data classification. This problem is related to the fact that machine learning algorithms in this task require the training data sets in which the samples number can be tens or even hundreds of thousands. The forming of such sets is a serious, time-consuming task, significantly increasing the projects developing cost and implementation duration.

The prospective methods of recognition by short datasets are fuzzy systems and GMDH. Earlier it was provided that neural networks are universal approximators and have some remarkable properties, such as parallel information processing, ability to work with incomplete noisy input data and learning possibilities to achieve the desired output.

The GMDH, from the other side, uses the principle of self-organization that allows constructing an optimal structure of the mathematical model during the algorithm operation. It's very promising to combine advantages of these both approaches for the solution of the problem – development an efficient model structure. GMDH-neural networks whose nodes are active neurons [10-12], N-adalines [13], R-neurons [14, 15], Q-neurons [16] and MLP [17] are known. At the junction of the fuzzy GMDH [18, 19] and neural networks, the GMDH-neuro-fuzzy system [14, 20] and the GMDH -neo-fuzzy system [21] were created.

These systems have proven their effectiveness in solving a wide range of problems (for example, [19, 22]), but have lost the main advantages of the original GMDH: a small number of evaluated parameters in each node. In this regard, it seems advisable to develop a GMDH-neo-fuzzy system that combines the advantages of traditional GMDH and hybrid computational intelligence systems, and is trained using simple procedures to ensure high speed of online image recognition.

The aim of the present work is a synthesis of the GMDH neo-fuzzy system for the online image recognition.

II. THE EXTENDED NEO-FUZZY NEURON AS A NODE OF GMDH-NEURO-FUZZY SYSTEM

Takeshi Yamakawa and co-authors in [23-25] proposed the architecture of neo-fuzzy neuron (NFN). The authors of the NFN admit among its most important advantages, the high learning rate, computational simplicity, the possibility to find the learning criterion global minima in real-time processing. Besides, NFN is characterized by fuzzy linguistic "if-then" rules. The neo-fuzzy neuron is a nonlinear multi-input single-output system shown in fig. 1.

It realizes the following mapping

$$f_i\left(x_i\right)=\sum_{j=1}^{h} w_{ji}\mu_{ji}\left(x_i\right) \tag{1}$$

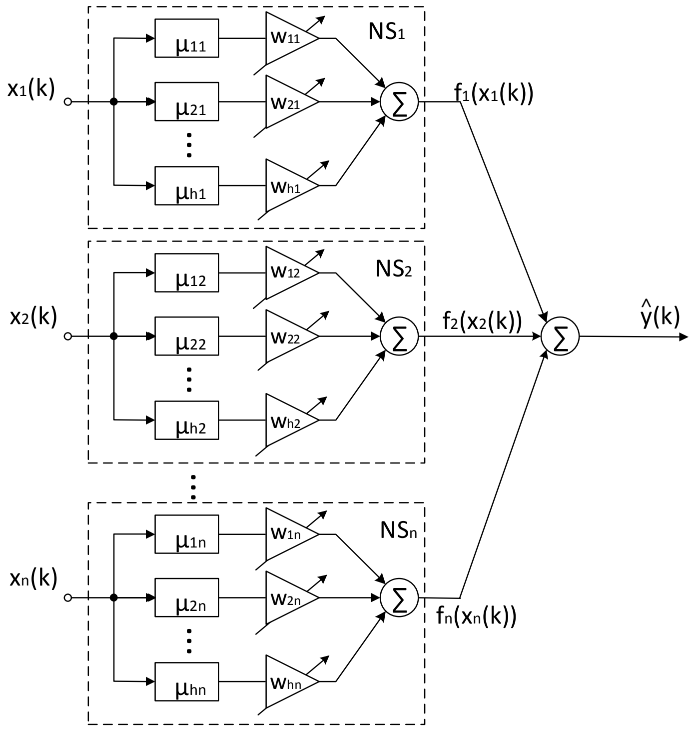

Fig. 1. Neo-fuzzy neuron

and implements fuzzy inference

IF x_i IS x_{ji} THEN THE OUTPUT IS w_{ji}

where x_{ji} is a fuzzy set with membership function $\mu_{ji}\left(x_i\right)$, w_{ji} is a singleton synaptic weight in consequent [26]. As it can be seen nonlinear synapse in fact actualize Takagi-Sugeno fuzzy inference of zero order. The membership functions $\mu_{ji}\left(x_i\right)$ in the antecedent could be B-splines or triangular functions, for example, like this

$$\mu_{ji}=\begin{cases}\dfrac{x_i-c_{j-1,i}}{c_{j,i}-c_{j-1,i}}, & if \quad x_i\in\left[c_{j-1,i},c_{j,i}\right],\\ \dfrac{c_{j+1,i}-x_i}{c_{j+1,i}-c_{j,i}}, & if \quad x_i\in\left[c_{j,i},c_{j+1,i}\right],\\ 0, & otherwise\end{cases}$$

where c_{ji} - the centers of membership functions, usually distributed on interval [0, 1]. This contributes to simplify the fuzzy inference. An input signal x_i activates only two neighboring membership functions simultaneously and the sum of the grades equals to unity, providing Ruspini partition:

$$\mu_{j-1,i}\left(x_i\right)+\mu_{ji}\left(x_i\right)=\mu_{ji}\left(x_i\right)+\mu_{j+1,i}\left(x_i\right)=1. \tag{2}$$

The inference result can be produced by arbitrary defuzzyfication method. Center-of-Gravity method gives output in the simple form:

$$f_i\left(x_i\right)=w_{ji}\mu_{ji}\left(x_i\right)+w_{j+1,i}\mu_{j+1,i}\left(x_i\right).$$

It is possible to improve approximating properties of such a system can by using a structural unit, called by authors as "extended nonlinear synapse" (ENS_i) (fig. 2) and synthesized on its basis the "extended neo-fuzzy neuron" [26-29] (ENFN). ENFN contains ENS_i as elements instead of usual nonlinear synapses NS_i.

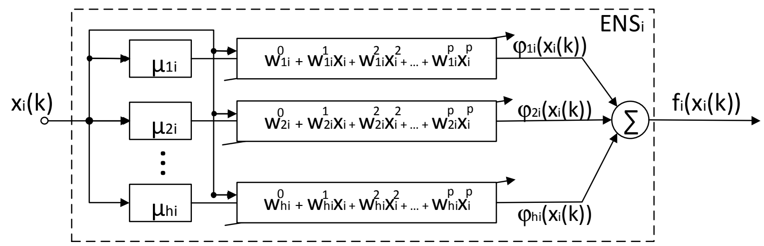

Fig. 2. Extended non-linear synapse

By introducing the additional variables

$$y_{li}\left(x_i\right)=\mu_{li}\left(x_i\right)\left(w_{li}^0+w_{li}^1x_i+w_{li}^2x_i^2+\ldots+w_{li}^px_i^p\right),$$

$$f_i\left(x_i\right)=\sum_{l=1}^{h}\mu_{li}\left(x_i\right)\left(w_{li}^0+w_{li}^1x_i+w_{li}^2x_i^2+\ldots+w_{li}^px_i^p\right)=$$

$$=w_{1i}^0\mu_{1i}\left(x_i\right)+w_{1i}^1x_i\mu_{1i}\left(x_i\right)+\ldots+w_{1i}^px_i^p\mu_{1i}\left(x_i\right)+$$

$$+w_{2i}^0\mu_{2i}\left(x_i\right)+\ldots+w_{2i}^px_i^p\mu_{2i}\left(x_i\right)+\ldots+w_{hi}^px_i^p\mu_{hi}\left(x_i\right),$$

$$w_i=\left(w_{1i}^0,w_{1i}^1,\ldots,w_{1i}^p,w_{2i}^0,\ldots,w_{2i}^p,\ldots,w_{hi}^p\right)^T,$$

$$\tilde{\mu}_i\left(x_i\right)=\Big(\mu_{1i}\left(x_i\right),x_i\mu_{1i}\left(x_i\right),\ldots,x_i^p\mu_{1i}\left(x_i\right),$$

$$\mu_{2i}\left(x_i\right),\ldots,x_i^p\mu_{2i}\left(x_i\right),\ldots,x_i^p\mu_{hi}\left(x_i\right)\Big)^T,$$

we can write

$$f_i\left(x_i\right)=w_i^T\tilde{\mu}_i\left(x_i\right), \tag{4}$$

$$\hat{y}=\sum_{i=1}^{n}f_i\left(x_i\right)=\sum_{i=1}^{n}w_i^T\tilde{\mu}\left(x_i\right)=\tilde{w}^T\tilde{\mu}\left(x\right), \tag{5}$$

where $\tilde{w}^T=\left(w_1^T,\ldots,w_i^T,\ldots,w_n^T\right)^T$

$$\tilde{\mu}\left(x\right)=\left(\tilde{\mu}_1^T\left(x_1\right),\ldots,\tilde{\mu}_i^T\left(x_i\right),\ldots,\tilde{\mu}_n^T\left(x_n\right)\right)^T.$$

It's easy to see that ENFN contains $(p+1)hn$ adjusting synaptic weights and fuzzy output, implemented by each ENSi, has the form

IF x_i IS x_{li} THEN THE OUTPUT IS

$$w_{li}^0 + w_{li}^1 x_i + \ldots + w_{li}^p x_i^p, l = 1, 2, \ldots, h, \quad (6)$$

i.e. essentially coincides with p-order Takagi-Sugeno inference.

Fig. 3 shows the architecture of an extended neo-fuzzy neuron.

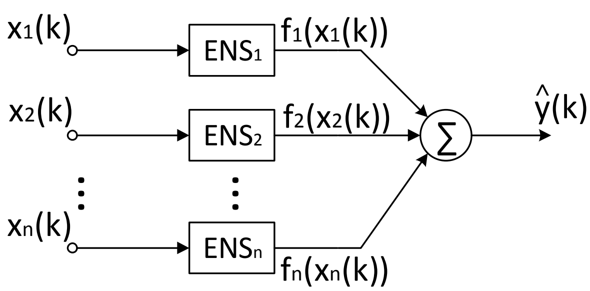

Fig. 3. Extended neo-fuzzy neuron

III. The Neo-Fuzzy Neuron Learning Algorithm

As a goal function for NFN learning the local quadratic error function is used:

$$E(k) = \frac{1}{2}\left(y(k) - \hat{y}(k)\right)^2 = \frac{1}{2}e^2(k) = \frac{1}{2}\left(y(k) - \sum_{i=1}^{n}\sum_{j=1}^{h} w_{ji}\mu_{ji}\left(x_i(k)\right)\right)^2.$$

It is minimized in the gradient stepwise algorithm in the form:

$$w_{ji}(k) = w_{ji}(k-1) + \eta e(k)\mu_{ji}\left(x_i(k)\right) =$$

$$= w_{ji}(k-1) + \eta\left(y(k) - \sum_{i=1}^{n}\sum_{j=1}^{h} w_{ji}(k-1)\mu_{ji}\left(x_i(k)\right)\right)\mu_{ji}\left(x_i(k)\right)$$

where $y(k)$ – external reference signal, $e(k)$ – learning error, η – the scalar learning rate parameter.

To increase the speed of training process Kaczmarz-Widrow-Hoff one-step learning algorithm [30-34] can be used:

$$w(k) = w(k-1) + \frac{y(k) - w^T(k-1)\mu\left(x(k)\right)}{\left\|\mu\left(x(k)\right)\right\|^2}\mu\left(x(k)\right)$$

where

$$\mu\left(x(k)\right) = \left(\mu_{11}\left(x_1(k)\right), \ldots, \mu_{h1}\left(x_1(k)\right), \ldots, \mu_{h2}\left(x_2(k)\right), \ldots,\right.$$

$$\left.\mu_{ji}\left(x_i(k)\right), \ldots, \mu_{hn}\left(x_n(k)\right)\right)^T,$$

$$w(k-1) = \left(w_{11}(k-1), \ldots, w_{h1}(k-1), \ldots,\right.$$

$\left.w_{h2}(k-1), \ldots, w_{ji}(k-1), \ldots, w_{hn}(k-1)\right)^T$ - $(nh\times 1)$ – vectors generated by input variables. The learning algorithm exponentially weighted form:

$$\begin{cases} w(k) = w(k-1) + r^{-1}(k)\left(y(k) - w^T(k-1)\mu\left(x(k)\right)\right)\mu\left(x(k)\right), \\ r(k) = \alpha r(k-1) + \left\|\mu\left(x(k)\right)\right\|^2, 0 \le \alpha \le 1 \end{cases}$$

which has filtering and tracking properties can be effectively used in stochastic and nonstationary situation.

IV. The Neuro-Fuzzy System and Its Architecture Optimization Using the Group Method of Data Handling

The neuro-fuzzy system under consideration is a multilayer feedforward architecture that consists of extended neo-fuzzy neurons and shown on fig. 4.

A $(n\times 1)$ -dimensional input signals vector $x = \left(x_1, x_2, \ldots, x_n\right)^T$ arrives at system zero receptive layer, and is transmitted then to first hidden layer containing $n_1 = C_n^2$ neuron nodes with only two inputs. At the outputs of first hidden layer nodes $N^{[1]}$, output signals $\hat{y}_l^{[1]}, l = 1, 2, \ldots, \frac{n}{2}(n-1) = C_n^2$ are formed. Further, these signals are sent to selection block $SB^{[1]}$ of first hidden layer, which selects from output signals set $\hat{y}_l^{[1]}$ the most accurate $n_1^*\left(n_1^* < n_1\right)$ of them in the accepted criterion sense, most often the mean square error $\sigma^2_{y_l^{[1]}}$.

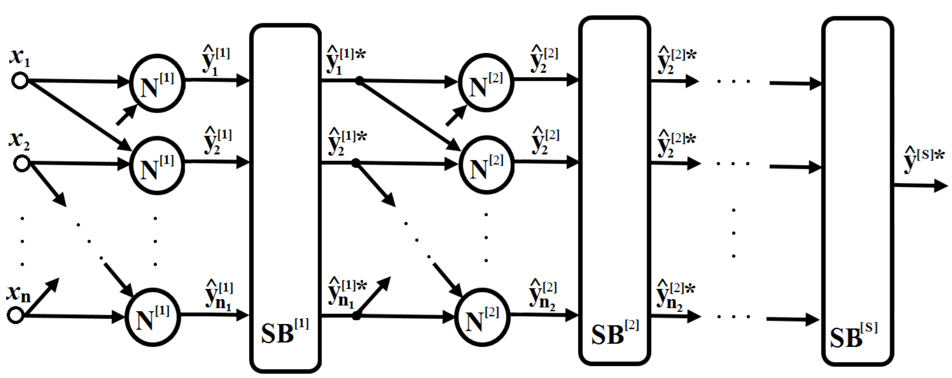

Fig. 4. GMDH-neuro-fuzzy system based on extended neo-fuzzy neurons

From these n_1^* best outputs of first hidden layer $\hat{y}_l^{[1]*}$, n_2 ($n_2 = C_n^2$ usually) pairwise combinations are formed, which are fed to second hidden layer formed by neurons $N^{[2]}$ similar to neurons $N^{[1]}$. From output signals of this layer $\hat{y}_l^{[2]}$, the selection block of the second hidden layer $SB^{[2]}$ selects the best F signals of the first hidden layer $\hat{y}_l^{[1]*}$, where F is a number of the best models choice according to an external criterion, and transfer them to the third layer.

The third hidden layer works similarly. The system evolution process continues until the best signal of the selection block $SB^{[s+1]}$ occurs to be worse than the best signal of the previous s-th layer, that is $\sigma^2_{y_l^{[s+1]}} > \sigma^2_{y_l^{[s]}}$. In this case, need to return to the previous row (iteration) and select the best model. This will be the optimal model of minimal complexity.

As the nodes of the GMDH system, one can use various neurons types with necessary approximating capabilities. However, at the same time, the main advantage of the original GMDH method - the ability to real-time work in the presence of short training sets - may be lost. This paper proposes to use the extended multidimensional neo-fuzzy neurons [23-26] as GMDH system nodes. Its architecture is a Takagi-Sugeno-Kang neuro-fuzzy system [35 - 37] with two inputs x_1 and x_2, five sequentially connected information processing layers and one output $\hat{y}_l$. A two-dimensional vector of input signals $x(k) = (x_1(k), x_2(k))^T$ to be processed is fed to the input of the node ($k = 1,2,...,N$ — the observation number in the training set or the current discrete time).The node first layer contains *2h* membership functions $\mu_{pi}(x_i), \mu_{pj}(x_j)$, $p = 1,2,...,h$ and implements fuzzyfication of input variables. The second layer provides aggregation of membership levels calculated in the first layer, contains h multiplication blocks and forms two-dimensional radial basis activation functions $\mu_{pi}(x_i), \mu_{pj}(x_j)$. The third layer is a layer of synaptic weights to be adjusted during the training process, while layer outputs are values $w_{lp}^{ij}\mu_{pi}(x_i), \mu_{pj}(x_j)$, and weights number is determined by number of membership functions at each input h. The fourth layer is formed by two adders and calculates the sum of second and third layers output signals. Finally, in the fifth neuron output layer normalization is performed, as a result of which the node output signal $\hat{y}_l$ is calculated.

Thus, when a signal $x(k)$ applying to the neuron node input, elements of the first layer calculate membership levels $0 < \mu_{pi}(x_i) \le 1, 0 < \mu_{pj}(x_j) \le 1$. As membership functions, bell-shaped constructions with a not strictly local receptor field are usually used. They avoids the appearance of “holes” in the fuzzyfied space when using the scattered partition spaces of input variables [37]. The membership functions of the first layer are Gaussian

$$\mu_{pi}(x_i(k)) = \exp\left(-\frac{(x_i(k) - c_{pi})^2}{2\sigma_i^2}\right),\ \mu_{pj}(x_j) = \exp\left(-\frac{(x_j(k) - c_{pj})^2}{2\sigma_j^2}\right)$$

where c_{pi}, c_{pj} - parameters defining the centers of membership functions, σ_i, σ_j - width parameters of these functions.

The *h* aggregated signals appears on second layer outputs

$$\tilde{x}_p(k) = \mu_{pi}(x_i(k))\mu_{pj}(x_j(k))$$

where $\tilde{x}_p(k) = \exp\left(-\frac{(x_i(k) - c_{pi})^2}{2\sigma_i^2}\right)\exp\left(-\frac{(x_j(k) - c_{pj})^2}{2\sigma_j^2}\right) =$

$$= \exp\left(-\frac{\|x(k) - c_p\|^2}{2\sigma^2}\right),\ c_p = (c_{pi}, c_{pj})^T.$$

The third layer outputs are the values

$$w_{lp}^{ij}\mu_{pi}(x_i(k))\mu_{pj}(x_j(k)) = w_{lp}^{ij}\tilde{x}_p(k),$$

the fourth:

$$\sum_{p=1}^{h} w_{lp}^{ij}\mu_{pi}(x_i(k))\mu_{pj}(x_j(k)) = \sum_{p=1}^{h} w_{lp}^{ij}\tilde{x}_p(k),$$

$$\sum_{p=1}^{h} \mu_{pi}(x_i(k))\mu_{pj}(x_j(k)) = \sum_{p=1}^{h} \tilde{x}_p(k)$$

and finally, at node output (fifth layer), a signal is formed

$$\hat{y}_l(k) = \frac{\sum_{p=1}^{h} w_{lp}^{ij}\mu_{pi}(x_i(k))\mu_{pj}(x_j(k))}{\sum_{p=1}^{h} \mu_{pi}(x_i(k))\mu_{pj}(x_j(k))} = \frac{\sum_{p=1}^{h} w_{lp}^{ij}\tilde{x}_p(k)}{\sum_{p=1}^{h} \tilde{x}_p(k)} =$$

$$= \sum_{p=1}^{h} w_{lp}^{ij}\frac{\tilde{x}_p(k)}{\sum_{p=1}^{h} \tilde{x}_p(k)} = \sum_{p=1}^{h} w_{lp}^{ij}\varphi_p^{ij}(x(k)) = (w_l^{ij})^T \varphi^{ij}(x(k))$$

where

$$\varphi_p^{ij}(x(k)) = \mu_{pi}(x_i(k))\mu_{pj}(x_j(k))\left(\sum_{p=1}^{h} \mu_{pi}(x_i(k))\mu_{pj}(x_j(k))\right)^{-1},$$

$$w_l^{ij} = (w_{l1}^{ij}, w_{l2}^{ij}, ..., w_{lh}^{ij})^T,$$

$$\varphi^{ij}(x(k)) = (\varphi_1^{ij}(x(k)), \varphi_2^{ij}(x(k)), ..., \varphi_p^{ij}(x(k)))^T.$$

It can be noticed that the node implements a non-linear mapping of input signals to output.

V. The Experiments

The task of a person's facial expression recognition is complex and multi-stage. It includes pre-processing of the image and searching for the face area within the image. After the face area is distinguished, it is possible to recognize the emotion by the face features set. In the practice of faces expressions recognition, several descriptor principles are used. The most common are adaptive appearance models, which uses the descriptions based on face image feature points and contours. It is established that such descriptions convey complete information about the person emotional state, even if it is expressed weakly.

Under the emotions influence, the facial muscles reduction leads to the displacement of feature points and this movement can serve as an indicator of basic facial actions. The most commonly used facial expressions are some basic emotions (fear, sadness, happiness, anger, disgust, surprise) and neutral state.

As a base for the feature vector, in [38] it was proposed to use a set of 35 characteristic points that can be localized in the facial area using contour detectors (Fig. 5).

Fig. 5. Examples of training images and position of characteristic points

The neo-fuzzy neurons has one output as the dimensionality of the output data vector. Seven basic emotions were selected for recognition: anger, disgust, fear, surprise, happiness, sadness, neutral expression. Therefore, the output values are {1; 2; 3; 4; 5; 6; 7}. The character features vector contains the two-dimensional coordinates of feature points position (Fig. 5). So, the system input is vector $\{x_i\}_{1\times 70}$. The order of the polynomial in nonlinear synapse membership function was chosen equal to 4, the number of synapses in the neo-fuzzy neuron was equal to 5.

The proposed architecture ability to recognize individual emotions was investigated using photographs from two open bases - Psychological Image Collection at Stirling (PICS) [39], partly from the Extended Cohn-Kanade (CK +) database [40]. Some images are in public use as objects for recognition.

In this set of photographs, those were selected differs in the person emotional state expression degree - from weakly noticeable to very noticeable.

In this task, special attention was paid to a learning data set small size. To examine how the proposed architecture and learning algorithm will recognize facial expressions, small photo sets are used. Their dimensions are given in Table 1.

A GMDH-neuro-fuzzy system based on extended neo-fuzzy neurons configured a task model from two rows. Learning error change is shown in Fig. 6.

Then the architecture ability to learn from a mixed set was examined, and sets total size was 344 photos. The number of unrecognized emotions is given in Table 2.

TABLE I. Dimensions of training sets of photos for individual emotions

Emotion	Anger	Disgust	Fear	Happiness	Sorrow	Surprise	Neutral
Data set size	49	66	35	45	19	50	80

TABLE II. The number of unrecognized emotions as a result of GMDH-neuro-fuzzy system based on extended neo-fuzzy neurons learning from a mixed set

	Primary emotions						
	Anger	*Disgust*	*Fear*	*Happiness*	*Sorrow*	*Surprise*	*Neutral*
The percentage of unrecognized images,%	2.04	0	5.71	4.44	0	0	2.5

VI. Conclusions

The article proposes the GMDH system with extended neo-fuzzy neurons as nodes. The system architecture allows modifying the process model structure in real time due to proposed neo-fuzzy nodes-neurons synaptic weights adjusting. A feature of the proposed architecture and its learning algorithm is the ability to work with small training samples.

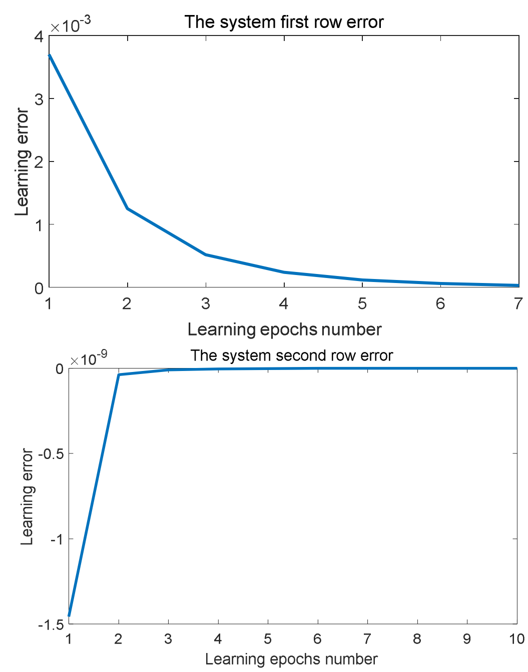

Fig. 6. A GMDH-neuro-fuzzy system based on extended neo-fuzzy neurons learning error

References

[1] A. Kołakowska, A. Landowska, M. Szwoch, W.Szwoch, M. R. Wrobel, "Human-Computer Systems Interaction: Backgrounds and

Applications," ch. 3, Emotion Recognition and Its Applications. Cham: Springer International Publishing, 2014, pp. 51 – 62.

[2] Kaggle. Challenges in representation learning: Facial recognition challenge, 2013.

[3] G.U. Kharat, S.V. Dudul, "Emotion Recognition from Facial Expression Using Neural Networks," in Human-Computer Systems Interaction. Advances in Intelligent and Soft Computing, vol 60, Z.S. Hippe, J.L. Kulikowski, Eds. Berlin, Heidelberg: Springer, 2009.

[4] C. Shan, S. Gong, P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and Vision Computing, vol. 27, no. 6, 2009, pp. 803 – 816.

[5] Ch.-Yi Lee, Li-Ch. Liao, "Recognition of Facial Expression by Using Neural-Network System with Fuzzified Characteristic Distances Weights," IEEE Int. Conf. Fuzzy Systems FUZZ-IEEE 2008. [IEEE World Congress on Computational Intelligence, pp. 1694 – 1699, 2008].

[6] N. Kulishova, "Emotion Recognition Using Sigma-Pi Neural Network," Proc. of 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2016, pp. 327 –331.

[7] A. Graves, J. Schmidhuber, C. Mayer, M. Wimmer, B. Radig, "Facial Expression Recognition with Recurrent Neural Networks," International Workshop on Cognition for Technical Systems, Munich, Germany, October 2008.

[8] S. Ouelett, "Real-time emotion recognition for gaming using deep convolutional network features," CoRR, vol. abs./1408.3750, 2014.

[9] B. Kim, J. Roh, S. Dong, S. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition," Journal on Multimodal User Interfaces, 2016, pp. 1–17.

[10] A.G. Ivakhnenko, G.A. Ivakhnenko, J.A Mueller, "Self-organization of the neural networks with active neurons", Pattern Recognition and Image Analysis, 4, №2, 1994, pp. 177 – 178.

[11] G.A. Ivakhnenko, "Self-organization of neuronet with active neurons for effect of nuclear test explosion forecasting," System Analysis Modeling Simulation, 20, 1995, pp. 107 -116.

[12] A.G. Ivakhnenko, D. Wuensch, G.A. Ivakhnenko, "Inductive sorting-out GMDH algorithms with polynomial complexity for active neurons of neural networks," Neural Networks, 2, 1999, pp. 1169 – 1173.

[13] D.U. Pham, X Liu, Neural networks for Identification, Prediction and Control. London: Springer-Verlag, 1995.

[14] T. Ohtani, "Automatic variable selection in RBF network and its application to neurofuzzy GMDH," Proc. Fourth Int. Conf. on Knowledge-Based Intelligent Engineering Systems and Allied Technologies, vol. 2, 2000, pp. 840 – 843.

[15] Ye. Bodyanskiy, N. Teslenko, P Grimm, "Hybrid evolving neural network using kernel activation functions," Proc. 17th Zittau East West Fuzzy Colloquium, Zittau/ Goerlitz: HS, 2010, pp. 39 – 46.

[16] Ye. Bodyanskiy, O. Vynokurova, I Pliss, "Hybrid GMDH-neural network of computational intelligence," Proc. 3rd Int. Workshop on Inductive Modeling, Krynica, Poland, 2009, pp. 100 – 107.

[17] M. S.Elbisy and F.A. Osra, "Application of Group Method of Data Handling Type Neural Network for Significant Wave Height Prediction," American Journal of Neural Networks and Applications, vol. 5, No. 2, 2019, pp. 51-57. doi: 10.11648/j.ajnna.20190502.12

[18] Yu. Zaychenko, "The fuzzy Group Method of Data Handling and its application for economical processes forecasting," Scientific Inquiry, 7, №1, 2006, pp. 83 – 96.

[19] H. Harandizadeh, D.J. Armaghani and E.T. Mohamad, "Development of fuzzy-GMDH model optimized by GSA to predict rock tensile strength based on experimental datasets," Neural Comput. & Applic., 2020. doi.org/10.1007/s00521-020-04803-z

[20] T. Ohtani, H. Ichihashi, T. Miyoshi, K. Nagasaka, Y Kanaumi, "Structural learning of neurofuzzy GMDH with Minkowski norm," Proc. 1998 Second Int. Conf. on Knowledge-Based Intelligent Electronic Systems, vol. 2. 1998, pp. 100 – 107.

[21] Ye. Bodyanskiy, Yu. Zaychenko, E. Pavlikovskaya, M. Samarina, Ye. Viktorov, "The neo-fuzzy neural network structure optimization using GMDH fog the solving forecasting and classification problems," Proc. Int. Workshop on Inductive Modeling, Krynica, Poland, 2009, pp. 77 – 89.

[22] O. Dag, M. Kasikci, E. Karabulut and Reha Alpar, "Diverse classifiers ensemble based on GMDH-type neural network algorithm for binary classification," Communications in Statistics - Simulation and Computation, 2019, pp. 1-17, DOI: 10.1080/03610918.2019.1697451

[23] J. Miki, T. Yamakawa, "Analog implementation of neo-fuzzy neuron and its on-board learning," in Computational Intelligence and Applications, Ed. N.E. Mastorakis, Piraeus: WSES Press, 1999, pp. 144 – 149.

[24] T. Yamakawa, E. Uchino, J. Miki, H. Kusanagi, "A neo-fuzzy neuron and its application to system identification and prediction of the system behavior," Proc. 2-nd Int. Conf. on Fuzzy Logic and Neural Networks "IIZUKA-92", Iizuka, Japan, 1992, pp. 477 – 483.

[25] E. Uchino, T. Yamakawa, "Soft computing based signal prediction, restoration and filtering," in Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks and Genetic Algoritms, Ed. Da Ruan, Boston: Kluwer Academic Publishers, 1997, pp. 331 – 349.

[26] Ye.V. Bodyanskiy, N.Ye. Kulishova, "Extended neo-fuzzy neuron in the task of images filtering", Radioelectronics. Computer Science. Control, № 1(32), 2014, pp. 112 – 119.

[27] Z. Hu, Ye. Bodyanskiy, N. Kulishova, O. Tyshchenko, "Multidimensional Extended Neo-Fuzzy Neuron for Facial Expression Recognition", International Journal of Intelligent Systems and Applications (IJISA), vol.9, №9, 2017, pp.29 – 36.

[28] Ye. Bodyanskiy, N. Kulishova, D. Malysheva, "The Extended Neo-Fuzzy System of Computational Intelligence and its Fast Learning for Emotions Online Recognition," Proc. of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine August 21–25, 2018, pp. 473 – 478.

[29] Ye. Bodyanskiy, N. Kulishova, O. Chala, "The Extended Multidimensional Neo-Fuzzy System and its Fast Learning in Pattern Recognition Tasks," Data, №3, 2018, pp. 63 – 73.

[30] S. Kaczmarz, "Angenaeherte Ausloesung von Systemen Linearer Gleichungen," Bull. Int. Acad. Polon. Sci, Let. A, 1937, ss. 355 – 357.

[31] S. Kaczmarz, "Approximate solution of systems of linear equations," Int. J. Control,53, 1993, pp. 1269-1271.

[32] Ye. Bodyanskiy, I. Kokshenev, V. Kolodyazhniy, "An adaptive learning algorithm for a neo-fuzzy neuron," Proc. of the 3rd Conference of the European Society for Fuzzy Logic and Technology, pp. 375 – 379, 2005.

[33] V. Kolodyazhniy, Ye. Bodyanskiy, P Otto, "Universal approximator employing neo-fuzzy neurons," Computational Intelligence: Theory and Applications. Ed. By B. Reusch, Berlin-Heidelberg: Springer, 2005, pp. 631 – 640.

[34] B. Widrow, Jr.M.E. Hoff, "Adaptive switching circuits," 1960 URE WESCON Convention Record, N.-Y.: IRE, part 4, 1960, pp. 96 – 104.

[35] L.-X. Wang, J.M. Mendel, "Fuzzy basis functions, universal approximation and orthogonal least-squares learning," IEEE Trans. on Neural Networks, 3, № 5, 1992, pp. 807 – 814.

[36] L.-X. Wang, Adaptive Fuzzy Systems and Control. Design and Statistical Analysis. Upper Saddle River: Prentice Hall, 1994.

[37] R. J.-S. Jang, C.-T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Upper Saddle River: Prentice Hall, 1997.

[38] Ye. V. Bodyanskiy, N. Ye. Kulishova and V. Ph. Tkachenko, "Feature vector generation for the facial expression recognition using neo-fuzzy system," Radio Electronics, Computer Science, Control, 2018, № 3, pp. 88-96. doi.org/10.15588/1607-3274-2018-3-10

[39] http://pics.psych.stir.ac.uk/2D_face_sets.htm

[40] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," Proceedings of IEEE workshop on CVPR for Human Communicative Behavior Analysis, San Francisco, USA, 2010