

Matrix Neo-Fuzzy-System and Its Online Learning in Image Recognition Task

Yevgeniy Bodyanskiy¹, Olha Chala^{2*} ^{1,2} Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

Abstract - The paper proposes a 2D-hybrid system of computational intelligence, which is based on the generalized neofuzzy neuron. The system is characterised by high approximate abilities, simple computational implementation, and high learning speed. The characteristic property of the proposed system is that on its input the signal is fed not in the traditional vector form, but in the image-matrix form. Such an approach allows getting rid of additional convolution-pooling layers that are used in deep neural networks as an encoder. The main elements of the proposed system are a fuzzified multidimensional bilinear model, additional softmax layer, and multidimensional generalized neo-fuzzy neuron tuning with cross-entropy criterion. Compared to deep neural systems, the proposed matrix neo-fuzzy system contains gradually fewer tuning parameters - synaptic weights. The usage of the time-optimal algorithm for tuning synaptic weights allows implementing learning in an online mode.

Keywords – 2D signals, generalized neo-fuzzy neuron, image recognition, membership functions data, optimal learning algorithm, stream mining.

I. INTRODUCTION

Currently, neural networks are spread worldwide to solve various tasks within data mining, such as image pattern recognition, the characteristic property of which is the input data that are fed to large-scale processing. Here the most famous networks are deep convolutional neural networks (CNNs), architecture of which includes relatively independent components: autoencoder – that transforms the initial imagematrix into a vector-signal of relatively low-dimension. After that this vector signal is fed to the input of multilayer fully connected perceptron that solves recognition-classification task. Usually, these systems are quite bulky, consist of an enormous number of tuned parameters-synaptic weights, and require a lot of time as well as huge training datasets for learning.

In order to speed up the learning process, the matrix neural networks [1]–[5] are used. Even though such systems surpass traditional CNNs by speed, they still need a big amount of training data.

The speed can be essentially improved by using probabilistic neural networks [6], which are tuned with "lazy learning". Additionally, the input information is fed to the network in the matrix form [7]. Unfortunately, the recognition results of such system, that uses kernel approximation according to Nadaraya-Watson estimates, are yield to the CNN's results by accuracy.

To find a compromise between speed, accuracy and learning speed, the Hybrid Systems of Computational Intelligence, among them neo-fuzzy neuron (NFN) [8]–[13], should be considered. NFN is characterised by simple architecture and learning algorithms, having reasonable approximation abilities, implementing nonlinear mapping $R^n \rightarrow R^1$, so in order to develop NFN for image processing the study [14] proposed a 2D-neo-fuzzy neuron that implements mapping $R^{n\times m} \rightarrow R^1$. It is clear that such system can solve only binary classification problem of matrix signals-images.

It is important to notice that in [15], a multilayer architectures were discussed, where instead of traditional elementary Rosenblatt perceptron [16], neo-fuzzy neurons were used. Even though these networks turned out to be faster than traditional CNNs, they still required a huge amount of training data.

This situation can be improved by using the idea of stacking approach [17], [18] when in each layer instead of usual neurons simplified shallow neural networks that realise nonlinear mapping $R^n \to R^p$ are used. Therefore, in [19]–[21] stacking multilayer neo-fuzzy systems were proposed, each layer of which was formed with the so-called generalized neo-fuzzy neuron (GNFN) [22], which was represented as generalisation of NFN for the case of multiple outputs, implementing mapping $R^n \to R^p$. Technically this task can be solved with *p*, connected in parallel NFN with one output. However, GNFN contains a fewer number of membership functions at the autoencoder layer.

Fig. 1 shows the architecture of GNFN with n inputs and pAn input signal in a vector outputs. form $x(k) = (x_1(k), ..., x_i(k), ..., x_n(k))^T \in \mathbb{R}^n$ (k = 1, 2, ..., N - or the number of observations in the training dataset, or the unit of discrete time in an online mode) is fed to a fuzzification layer which is formed with hn membership functions μ_{li} , l = 1, 2, ..., h; i = 1, 2, ..., n, implementing fuzzification of input data. The tuning synaptic weights w_{ili} , j = 1, 2, ...p the amount of which is defined by number phn are connected with membership functions μ_{li} . The output signal of GNFN is a vector $\hat{y}(k) = (\hat{y}_1(k), ..., \hat{y}_i(k), ..., \hat{y}_n(k))^T \in \mathbb{R}^p$ where on the *j*-th output the following value appears:

©2021 Yevgeniy Bodyanskiy, Olha Chala.

^{*} Corresponding author. E-mail: chala.olha@nure.ua

This is an open access article licensed under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

$$\hat{y}_j(k) = \sum_{i=1}^n \sum_{l=1}^h w_{jli} \mu_{li}(x_i(k)), j = 1, 2, \dots, p. \quad (1)$$



Fig. 1. Generalized neo-fuzzy neuron.

By introducing $(hn \times 1)$ vectors of membership functions $\mu(x(k)) = (\mu_{11}(x_1(k), \mu_{21}(x_1(k), ..., \mu_{h1}(x_1(k), \mu_{12}(x_2(k), ..., \mu_{li}(x_i(k), ..., \mu_{hn}(x_n(k)))^T \in \mathbb{R}^{hn}$ and $(p \times hn)$ – matrix of synaptic weights

$$W = \begin{pmatrix} w_{111} & w_{121} & \cdots & w_{1hn} \\ w_{211} & w_{221} & \cdots & w_{2hn} \\ \vdots & \vdots & \ddots & \vdots \\ w_{p11} & w_{p21} & \cdots & w_{phn} \end{pmatrix} \in R^{p \times hn}$$
(2)

it is simply to write the transformation that is realised by the GNFN in the form:

$$\hat{y}(k) = w\mu(x(k)) . \tag{3}$$

It is important to notice that the same result can be achieved using *p* connected in parallel single-output NFN; however, in such a system the number of membership functions is *p* times greater than in GNFN. Also, in (2) synaptic weights w_{jli} are fed linearly that allows using time optimal algorithms such as Kaczmarz-Widrow-Hoff [23]–[25] for their learning. Some bell-shaped functions can be used as membership functions, which are used in the neuro-fuzzy system theory. However, from the computational point of view, triangular constructions are the most useful ones.

$$\mu_{li}(x_{i}(k)) = \begin{cases} \frac{x_{i}(k) - x_{l-1,i}^{*}}{x_{1i}^{*} - x_{l-1,i}^{*}} & \text{if } x_{i}(k) \in [x_{1i}^{*}, x_{l-1,i}^{*}], \\ \frac{x_{l+1,i}^{*} - x_{i}(k)}{x_{l+1,i}^{*} - x_{1i}^{*}} & \text{if } x_{i}(k) \in [x_{1i}^{*}, x_{l+1,i}^{*}], \\ 0 & \text{otherwise} \end{cases}$$
(4)

satisfying Ruspini unity partition condition:

$$\begin{cases} \mu_{l-i,i}(x_i(k)) + \mu_{li}(x_i(k)) = 1 i f x_i(k) \in [x_{l-i,i}^*, x_{li}^*] \\ \mu_{li}(x_i(k)) + \mu_{l+i,i}(x_i(k)) = 1 i f x_i(k) \in [x_{li}^*, x_{l+i,i}^*] \end{cases}$$
(5)

where $\mathbf{x}_{l-i,i}^*, \mathbf{x}_{li}^*, \mathbf{x}_{l+i,i}^*$ are centres of neighbouring membership functions, which usually are chosen empirically.

An advantage of functions of Type (4) is when the signal $x_i(k) \in [x_{l-i,i}^*, x_{li}^*]$ is fed to the system, only two neighbouring functions $\mu_{l-i,i}(x_i(k))$ and $\mu_{li}(x_i(k))$ are fired. Moreover, only two synaptic weights $w_{j,l-1,i}$ and w_{jli} are tuned that gradually simplify the learning process as well as speed.

In case of the necessity of image processing which are fed in a form $(n \times m)$ - matrix $x(k) = \{x_{ili2}(k)\} \in \mathbb{R}^{n \times m}$, $i_1 =$ **1**, **2**, ..., **n**; $i_2 = 1, 2, ..., m$. This signal can be transformed into $(nm \times 1)$ - vector $\vec{x}(k)$, however, the number of tuned parameters in this situation grows rapidly that makes the system abnormally cumbersome. Thus, it is effectively basing on GNFN to introduce into consideration matrix neo-fuzzy system that is purposed to solve pattern-image recognition task.

II. ARCHITECTURE OF MATRIX NEO-FUZZY SYSTEM

Architecture of matrix neo-fuzzy system is a generalisation of GNFN in case of 2D output signals presented in Fig. 2. It consists of four sequentially connected layers: a fuzzification layer, two layers of synaptic weight tuning and output layer of membership functions.

The first layer is processed with *hnm* membership functions of Type (4) $\mu_{ili2}(x_{ili2}(k))$, *h* functions for each component $(n \times m)$ of matrix signal $x(k) = \{x_{ili2}(k)\}$, l=1,2,...,h, $i_1 =$ **1**, **2**, ..., **n**; $i_2 =$ **1**, **2**, ..., **m**. At this layer, the input signal is fuzzified, in other words, on its output $(hn \times m)$ matrix signal of fuzzified inputs $\mu(\mathbf{x}(\mathbf{k}))$ is formed.

$$\hat{y}(k) = W_L^{(k-1)} \mu(x(k)) W_R(k-1) , \qquad (6)$$

where $W_L(k-1), W_R(k-1) - (p \times hn)$ and (m+1) are matrices of tuned weights which are obtained from data of previous (k-1) observations and specified with online learning algorithms; $(p \times 1)$ is an output signal $\hat{y}(k)$ (here p – the number of classes in the training dataset). Hereafter, signal is fed to the output layer of softmax membership, forming output $(p \times 1)$ signal $y^*(k)$, components of which $y_j^*(k), j = 1, 2, ... p$ essentially are fuzzy membership levels of the observation x(k) to the *j*-th class.

III. LEARNING OF MATRIX NEO-FUZZY SYSTEM

A learning $(p \times 1)$ vector signal y(k) in pattern recognition tasks is formed based on "one-hot coding" and consists of p-1 zeroes and single one, which is related to the "correct" class.

As a training criterion the "cross-entropy" in the following form can be used:

$$E = -\sum_{k=1}^{N} \sum_{j=1}^{p} y_{j}(k) \ln y_{j}^{*}(k) =$$

= $-\sum_{k=1}^{N} \sum_{j=1}^{p} y_{j}(k) \ln \operatorname{softmax} \hat{y}(k) = ,$ (7)
= $-\sum_{k=1}^{N} \sum_{j=1}^{p} y_{j}(k) \ln \frac{e^{\hat{y}_{j}}(k)}{\sum_{q=1}^{p} e^{\hat{y}_{q}(k)}}$

or (what is more appropriate for work in an online mode) it is possible to use the goal function, which is based on an instant mean-square error.

Let us introduce into consideration three output signals

$$\begin{cases} \hat{y}(k) = W_L(k-1)\mu(x(k))W_R(k-1) = \\ = W_L(k-1)\mu_R(k), \\ \hat{y}(k) = W_L(k)\mu(x(k))W_R(k-1) = \\ = \mu_L(k)W_R(k-1), \\ \hat{y}_R(k) = W_L(k)\mu(x(k))W_R(k) \end{cases}$$
(8)

(here $W_L(k-1)$, $\mu_R(\mathbf{k})$, $\mu_L(\mathbf{k})$, $W_R(k-1) - (p \times hn)$, $(p \times m)$, $(hn \times 1)$, $(m \times 1)$ – matrices respectfully) of the matrix neo-fuzzy systems.

It can be seen that the second and third hidden layers W_L and W_R form adaptive bilinear model [26].



Fig. 2. Matrix neo-fuzzy system for image recognition.

By introducing three learning errors

$$\begin{cases}
e(k) = y(k) - \hat{y}(k) = \\
= y(k) - W_L(k - 1)\mu_R(k), \\
e_L(k) = y(k) - \hat{y}_L(k) = \\
= y(k) - \mu_L(k - 1)W_R(k - 1), \\
e_R(k) = y(k) - \hat{y}_R(k) = \\
= y(k) - W_L(k)\mu(x(k))W_R(k)
\end{cases}$$
(9)

and two locally optimal learning criteria

$$\begin{cases} E(k) = ||e(k)||^2, \\ E_L(k) = ||e_L(k)||^2. \end{cases}$$
(10)

the learning algorithm of synaptic weight matrices can be written in the following form:

$$\begin{cases} W_L(k) = W_L(k-1) + \frac{e(k)\mu_R^T(k)}{\|\mu_R(k)\|^2}, \\ W_R(k) = W_R(k-1) + \frac{e_L(k)\mu_L^T(k)}{Tr(\mu_L(k)\mu_L^T(k))} \end{cases}$$
(11)

or

$$\begin{pmatrix}
W_{L}(k) = W_{L}(k-1) + \\
+ \frac{y(k) - W_{L}(k-1)\mu(x(k))W_{R}(k-1)}{\|\mu(x(k))W_{R}(k-1)\|^{2}} \\
(\mu(x(k))W_{R}(k-1))^{T}, \\
W_{R}(k) = W_{R}(k-1) + (W_{L}(k)\mu(x(k)))^{T} \\
\frac{y(k) - W_{L}(k)\mu(x(k))W_{R}(k-1)}{Tr(W_{L}(k)\mu(x(k))W_{L}^{T}(k)\mu^{T}(k))}.
\end{cases}$$
(12)

It is easy to notice that algorithms (11), (12) are a generalization of optimal by speed adaptive algorithm Kaczmarz-Widrow-Hoff [23]–[25] and in matrix case [27].

Hence, the learning process of the proposed matrix neo-fuzzy system is characterised by the optimal speed in the class of gradient procedures, which allows implementing the system in online mode.

This characteristic is of prime importance in the learning process in the case of short training datasets where the speed of convergence is at the fore, additionally, in Big Data Mining tasks, where data are fed to the processing sequentially, and the size of the dataset is unknown.

IV. THE RESULTS OF A COMPUTATIONAL EXPERIMENT

The effectiveness of the proposed system was tested using a dataset – Fashion MNIST [28]. There are 60 000 samples in the training dataset and 10 000 in the test dataset where each sample is an image 24×24 .

The first experiment was time and accuracy comparison of the proposed GNFN and simple custom CNN, architecture of which is represented in Fig. 3.



Fig. 3. Architecture of CNN.

TABLE I THE ACCURACY AND TIME CONSUMPTION OF ALGORITHMS WITH FASHION MNIST DATASET

Algorithms for the comparison	Classification accuracy, %	Time, s
CNN	97.91	182
Proposed approach	92.04	78

From Table I, it can be seen that the accuracy of the proposed system yields to CNN; however, here it is important to notice that it has drastically different time consumption, in particular, CNN requires twice time of the NFN system.

Next step was practical research of the triangular membership function (Fig. 4) effectiveness, and for this task the original architecture of the proposed system was taken and compared to the absolutely identical one, except the membership function was changed to the traditional Gaussian membership function, and results can be seen in Table II.



Fig. 4. Triangular membership function.

TABLE II THE ACCURACY AND TIME CONSUMPTION FOR SYSTEMS WITH DIFFERENT MEMBERSHIP FUNCTIONS

Used membership functions	Classification accuracy, %	Time, s
Triangular	92.04	78
Gaussian	91.58	128

It can be seen that the original architecture with the triangular membership function works faster compared to the one with the sigmoid membership function even though their accuracy is comparable. Such a situation occurs because during classification only two neighbouring functions are activated leading to the simplification of calculations, thus increasing the processing speed.

This experiment was followed by the comparison of the weight tuning procedures. We have compared the proposed modified optimal by speed Kaczmarz-Widrow-Hoff algorithm with classical stochastic gradient procedure for synaptic weights tuning in the system under consideration. The obtained results are shown in Table III.

TABLE III TIME CONSUMPTION FOR PROCESSING USING DIFFERENT LEARNING ALGORITHMS

Learning algorithms for comparison	Time, s
Modified Kaczmarz- Widrow-Hoff	78
Stochastic gradient descent	135

It can be seen that the proposed learning algorithm is faster in terms of speed, because it has matrix modification that allows skipping the vectorization part of the image processing. Additionally, the original Kaczmarz-Widrow-Hoff algorithm is optimal by time by its nature.

Hence, the results of the experiment confirm effectiveness of the proposed system in image online classification.

V. CONCLUSION

In the article, the matrix neo-fuzzy system has been proposed. It has been designed to solve a pattern-image recognition task without vectorization of the initial data that are represented in matrix form. The system is based on a generalized neo-fuzzy neuron, which is modified for the case of 2D-input signals.

Learning of the proposed system is implemented with optimal by speed gradient algorithm, which is adapted for matrix signals. The proposed system is characterised by the high speed that allows processing data streams in online mode. Additionally, it has good approximation properties and allows solving the classification task of difficult form objects. The proposed system is simple enough in calculation implementation, and the results of the computational experiment have proven its effectiveness.

REFERENCES

- P. Daniušis and P. Vaitkus, "Neural network with matrix inputs," *Informatica*, vol. 19, no. 4, pp. 477–486, Jan. 2008. <u>https://doi.org/10.15388/Informatica.2008.225</u>
- [2] P. Stubberud, "A vector matrix real time backpropagation algorithm for recurrent neural networks that approximate multi-valued periodic functions," *International Journal on Computational Intelligence and Application*, vol. 8, no. 4, pp. 395–411, Dec. 2009. https://doi.org/10.1142/S1469026809002667
- [3] M. Mohamadian, H. Afarideh, and F. Babapour, "New 2D matrix-based neural network for image processing applications," *IAENG International Journal of Computer Science*, vol. 42, no. 3, pp. 265–274, 2015.
- [4] J. Yao, Y. Guo, and Z. Wang, "Matrix neural networks," *Lecture Notes in Computer Science*, vol. 10261, May 2017, pp. 313–320. https://doi.org/10.1007/978-3-319-59072-1_37
- [5] J. Pliss, O. Boiko, V. Volkova, and Y. Bodyanskiy, "Matrix deep neural network and its rapid learning in Data Science tasks"," *Proceedings of the International Conference "Advanced Computer Information Technologies*," Czech Republic, June 2018, pp. 141–144.

- [6] D. F. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3., no. 1, pp. 109–118, 1990. <u>https://doi.org/10.1016/0893-6080(90)90049-0</u>
- [7] Ye. Bodyanskiy, A. Deineko, I. Pliss, O. Chala, and A. Nortsova, "Matrix fuzzy-probabilistic neural network in image recognition task," 2020 IEEE Third International Conference on Data Stream Mining and Processing (DSMP), Lviv, Ukraine, Aug. 2020, pp. 33–36. https://doi.org/10.1109/DSMP47368.2020.9204236
- [8] J. Yamakawa, E. Uchino, J. Miki, and H. Kusanagi, "A neo-fuzzy neuron and its application to system identification and prediction of the system behavior," *Proceedings of the 2nd International Conference on Fuzzy Logic & Neural Networks*, lizuka, Japan, July 1992.
- [9] E. Uchino and T. Yamakawa, "Neo-fuzzy-neuron based new approach to system modeling, with application to actual system," *Proceedings Sixth International Conference on Tools with Artificial Intelligence, TAI 94*, New Orleans, LA, USA, Nov. 1994. pp. 564–570. https://doi.org/10.1109/tai.1994.346442
- [10] T. Miki, "Analog implementation of neo-fuzzy neuron and its on-board learning," in *IMACS/IEEE CSCC'99 Proceedings*, pp. 4401–4406, 1999.
- [11] A. M. Silva, W. Caminhas, A. Lemos, and F. Gomide, "A fast learning algorithm for evolving neo-fuzzy neuron," *Applied Soft Computing*, vol. 14, Part B, pp. 194–209, Jan. 2014. <u>https://doi.org/10.1016/j.asoc.2013.03.022</u>
- [12] D. Zurita, M. Delgado, J. A. Carino, J. A. Ortega, and G. Clerc, "Industrial time series modelling by means of the neo-fuzzy neuron," *IEEE Access*, vol. 4, pp. 6151–6160, Sept. 2016. https://doi.org/10.1109/ACCESS.2016.2611649
- [13] Y. Bodyanskiy, A. Deineko, I. Pliss, and O. Chala, "Fast probabilistic neuro-fuzzy system for pattern classification task," *Information Technology and Management Science*, vol. 23, no. 1, pp. 12–16, Dec. 2020. <u>https://doi.org/10.7250/itms-2020-0002</u>
- [14] Y. Bodyanskiy, O. Vynokurova, V. Volkova, and O. Boiko, "2D-Neo-Fuzzy Neuron and Its Adaptive Learning," *Information Technology and Management Science*, vol. 21, no. 1, pp. 24–28, Dec. 2018. https://doi.org/10.7250/itms-2018-0003
- [15] Y. Bodyanskiy and T. Antonenko, "Deep neo-fuzzy neural network and its accelerated learning," 2020 IEEE Third International Conference on Data Stream Mining Processing (DSMP), Lviv, Ukraine, Aug. 2020, pp. 67–71. https://doi.org/10.1109/DSMP47368.2020.9204068
- [16] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain.," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958. <u>https://doi.org/10.1037/h0042519</u>
- [17] D. H. Wolpert, "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241–259, Jan. 1992. <u>https://doi.org/10.1016/S0893-6080(05)80023-1</u>
 [18] L. Deng, D. Yu, and J. Platt, "Scalable stacking and learning for building
- [18] L. Deng, D. Yu, and J. Platt, "Scalable stacking and learning for building deep architectures," 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, Mar. 2012, pp. 2133–2136. <u>https://doi.org/10.1109/ICASSP.2012.6288333</u>
- [19] Y. Bodyanskiy, I. Pliss, O. Vynokurova, D. Peleshko, and Y. Rashkevych, "Neo-fuzzy encoder and its adaptive learning for big data processing," *Information Technology and Management Science*, vol. 20, no. 1, Jan. 2017. https://doi.org/10.1515/itms-2017-0001
- [20] Y. Bodyanskiy, O. Vynokurova, I. Pliss, D. Peleshko, and Y. Rashkevych, "Deep stacking convex neuro-fuzzy system and its on-line learning," *Advances in Dependability Engineering of Complex Systems*, vol. 582, W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, and J. Kacprzyk, Eds. Cham: Springer International Publishing, 2018, pp. 49–59. https://doi.org/10.1007/978-3-319-59415-6_5
- [21] Y. Bodyanskiy, G. Setlak, O. Vynokurova, I. Pliss, and O. Boiko, "Deep evolving stacking convex cascade neo-fuzzy network and its rapid learning," *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems*, Sep. 2018, pp. 29–33. <u>https://doi.org/10.15439/2018F200</u>
- [22] R. P. Landim, B. Rodrigues, S. R. Silva, and W. M. Caminhas, "A neofuzzy-neuron with real time training applied to flux observer for an induction motor," *Proceedings 5th Brazilian Symposium on Neural Networks (Cat. No.98EX209)*, Belo Horizonte, Brazil, 1998, pp. 67–72. <u>https://doi.org/10.1109/sbrn.1998.730996</u>
- [23] S. Kaczmarz, "Angenaherte Auflosung von Systemen linearer Gleichungen," in Bull. Acad. Polonaise Sci. et Lettres A, pp. 355–357, 1937.
- [24] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *IRE WESCON Convention Record*, pp. 96–104, 1960.
- [25] S. Kaczmarz, "Approximate solution of systems of linear equations," International Journal of Control, vol. 57, no. 6, pp. 1269–1271, Jun. 1993. <u>https://doi.org/10.1080/00207179308934446</u>

- [26] Ye. Bodyanskiy, I. Pliss, and V. A. Timofeev, "Discrete adaptive identification and extrapolation of two-dimensional fields," *Pattern Recognition and Image Analysis*, vol. 5, no. 3, pp. 410–416, 1995.
- [27] O. Rudenko, Ye. Bodyanskiy, and I. Pliss, "Adaptivnyy algoritm prognozirovaniya sluchaynykh posledovatel'nostey," *Avtomatika*, vol. 12, no. 1, pp. 46–48, 1979.
- [28] Fashion-MNIST. Zalando Research. (2021). [Online]. Available: https://github.com/zalandoresearch/fashion-mnist, Accessed on: Oct. 15, 2021.

Yevgeniy Bodyanskiy. In 1971, he graduated with honour from Kharkiv National University of Radio Electronics (KhNURE). In 1980, he defended the Doctoral Thesis. In 1984, he was awarded the academic title of Senior Researcher. In 1990, he was awarded Dr. habil. sc. ing. degree. In 1994, he was awarded the academic title of Professor. His major fields of research are evolving hybrid systems of computational intelligence, data stream mining, data science, and big data. Since 1974, he has been working at Kharkiv National University of Radio Electronics. In 1974–1976, he was a Researcher; in 1977– 1983, he was a Senior Researcher; in 1986-1991, he was a Scientific Head of Control Systems Research Laboratory; in 1991-1992, he was a Research Fellow. Since 1992, he has been a Professor of Artificial Intelligence Department at KhNURE, Scientific Head of Control Systems Research Laboratory at KhNURE. He has more than 660 scientific publications, including 42 inventions and 16 monographs. Research interests include hybrid systems of computational intelligence: adaptive, neuro-, wavelet-, neo-fuzzy-, real-time systems, including problems connected with control, identification, forecasting, clustering, diagnostics, fault detection in technical, economical, medical and ecological objects. He is the IEEE senior member, member of four scientific and seven editorial boards.

Contact data: Office 511, 14 Nauky Ave., Kharkiv, 61166, Ukraine.

ORCID iD: https://orcid.org/0000-0001-5418-2143

Olha Chala. In 2019, she obtained a Bachelor degree in Computer Science at Kharkiv National University of Radio Electronics. In 2020, she obtained a Master degree in Artificial Intelligence at KhNURE. Since 2018, she has been working at the Control Systems Research Laboratory of Kharkiv National University of Radio Electronics. In 2021, she started working as an Assistant Professor at the Artificial Intelligence Department of KhNURE. She has more than 10 scientific publications. Her current scientific interests are neuro-fuzzy systems, deep learning, neo-fuzzy systems, and pattern recognition task. Contact data: Office 511, 14 Nauky Ave., Kharkiv, 61166, Ukraine. E-mail: chala.olha@nure.ua

ORCID iD: https://orcid.org/0000-0001-7918-7362

E-mail: yevgeniy.bodyanskiy@nure.ua