UDC 004.032.26:004.93

# PIXEL-TO-SCALE STANDARD DEVIATIONS RATIO OPTIMIZATION FOR TWO-LAYER PERCEPTRON TRAINING ON PIXEL-DISTORTED SCALED 60-BY-80-IMAGES IN SCALED OBJECTS CLASSIFICATION PROBLEM

**V. Romanuke**
Khmelnytskyy National University
vul. Instytutska 11, 29000, Khmelnytskyy, Ukraine. E-mail: romanukevadimv@mail.ru

An optimization problem in identifying the two-layer perceptron classifier over scaled objects is solved. General totality is formed of monochrome 60-by-80-images of alphabet letters, so it is of 26 classes. The two-layer perceptron is modeled, trained and tested under 250 neurons in its hidden layer and “traingda” MATLAB training function. The optimization parameter is pixel-to-scale standard deviations ratio for forming training sets properly. The optimization criterion is classification error percentage minimization. The optimal ratio is statistically tracked while passing the training sets through the perceptron for 20 times. Then, increasing the passes number up to 50 at the optimal ratio, there has been statistically tracked the minimum of classification error percentage. For avoiding the fuzziness in minimization conclusion, there is applied a complementary criterion, where the classifier is tested with both scaled images and pixel-distorted scaled images. In verification procedure, when the trained-on-optimal-ratio classifier is used, the classification error percentage has been disclosed to be lesser than 0,44 on the scale range average.

**Key words:** scaled objects classification, two-layer perceptron, training set, classification error percentage, monochrome image, pixel-to-scale standard deviations ratio optimization.

# ОПТИМІЗАЦІЯ СПІВВІДНОШЕННЯ СЕРЕДНЬОКВАДРАТИЧНИХ ВІДХИЛЕНЬ ПІКСЕЛЬНИХ СПОТВОРЕНЬ І МАСШТАБУВАННЯ ДЛЯ НАВЧАННЯ ДВОШАРОВОГО ПЕРСЕПТРОНА НА МАСШТАБОВАНИХ ЗОБРАЖЕННЯХ ФОРМАТУ 60-НА-80 З ПІКСЕЛЬНИМИ СПОТВОРЕННЯМИ У ЗАДАЧІ КЛАСИФІКАЦІЇ МАСШТАБОВАНИХ ОБ’ЄКТІВ

**В. В. Романюк**
Хмельницький національний університет
вул. Інститутська, 11, 29000, м. Хмельницький, Україна. E-mail: romanukevadimv@mail.ru

Розв’язується оптимізаційна задача для ідентифікації класифікатора масштабованих об’єктів на основі двошарового персептрона. Генеральна сукупність формується з монохромних зображень формату 60-на-80 літер алфавіту, і вона складається із 26 класів. Двошаровий персептрон моделюється, навчається та тестується за 250 нейронів у його прихованому шарі та MATLAB-функції для навчання “traingda”. Оптимізаційним параметром є співвідношення середньоквадратичних відхилень піксельних спотворень і масштабування для належного формування навчальних множин. Оптимізаційним критерієм є мінімізація відсотка помилок класифікації. Оптимальне співвідношення статистично відслідковується за проходження навчальних множин через персептрон протягом 20 разів. Згодом, збільшуючи число проходжень до 50 за оптимального співвідношення, статистично відслідковано мінімум відсотка помилок класифікації. Для того, щоб позбутися нечіткості у висновку щодо мінімізації, застосовується додатковий критерій, де класифікатор тестується як на виключно масштабованих зображеннях, так і на масштабованих зображеннях з піксельними спотвореннями. У процедурі верифікації, коли застосовується навчений за оптимального співвідношення класифікатор, виявлено, що відсоток помилок класифікації є меншим за 0,44 у середньому за діапазоном масштабування.

**Ключові слова:** класифікація масштабованих об’єктів, двошаровий персептрон, навчальна множина, відсоток помилок класифікації, монохромне зображення, оптимізація співвідношення середньоквадратичних відхилень піксельних спотворень і масштабування.

PROBLEM STATEMENT. Problem of scaled objects classification is plain. An effect of scaling occurs usually when the being watched or controlled object is distanced from a camera or scanner of the classification system. One of the greatest problems of scaled objects classification is the classifier operation speed [1, 2]. Neocognitron and its derivatives, performing perfectly over scaled objects for their classification, take pretty long periods for the performance [3, 4]. Gigantic resources are consumed as well. Perceptron could have substituted neocognitrons to impart its swiftness to the scaled objects classification, but this type of classifiers needs substantial optimization of its parameters to become smarter after the training process [5, 6].

Optimization of the classifier parameters for improving the classification process lies in the following. For improving the classification process, the classifier parameters should be assigned rationally before running its training process [5, 7, 8]. Rationality is proved via batch testings of the trained classifier. Any other parameters, being assigned after the perceptron is trained, cannot change its performance characteristics. The perceptron training process, defining the classifier performance characteristics, has a great many of its parameters [7, 9, 10]. Parameters of the training set in the training process are specific for classification problems, provoked with types of distortion of objects (shift, turn or rotation, scale, nonlinear distortion). Particularly, in classifying scaled images the training set is formed as addition of matrix of pixel-distortion [11, 12] and matrix of scaled images [1, 2, 13]. And the question is what ratio of pixel-distortion share and scaling-distortion share shall be for forming the training set rationally, not heuristically. Just as an instance, shall it be $1:1$, or $2:1$, or $1:2$ ? This ratio $r$, if it is optimized to $r^*$, will allow to form the training set optimally and to identify the perceptron for classifying the scaled objects with much lower classification error percentage (CEP) over them.

Tasks and goal of article are as follows. The idea of this investigation is to plot CEP against $r$ with further solving the simple minimization problem. The investigation is going to be carried out within MATLAB, having powerful Neural Network Toolbox [12, 14, 15]. So, goal of the article is to define the spoken ratio $r$ and find its optimum on the corresponding minimization problem solution, at which CEP over scaled objects would be minimized for the trained perceptron, regardless of the traintime duration. To accomplish that, the object model should be taken firstly along with the general totality and number of classes in it. Then the number of hidden layers in perceptron with their neurons ought to be grounded [12, 16, 17], whereupon the training MATLAB function will be selected [12, 18, 19]. The next items are: to select (heuristically, believably) boundaries $r_{\min}$ and $r_{\max}$ for the ratio $r$, based on the object model, and select the step within the range $[r_{\min}; r_{\max}]$ of this ratio for running through this range as fast as possible, but the results would be adequate. After having run through the range $[r_{\min}; r_{\max}]$ there must be plotted the averaged CEP $p_{\mathrm{error}}(r)$ over this range, what would let solve the problem

$$r^* \in \arg \min_{r \in [r_{\min};\, r_{\max}]} p_{\mathrm{error}}(r). \qquad (1)$$

Certainly, it ought to be verified whether CEP $p_{\mathrm{error}}(r^*)$ is minimal over the being modeled scaled objects.

EXPERIMENTAL PART AND RESULTS OBTAINED. General totality of monochrome images and number of classes are as follows. May the monochrome image be the object model. It is convenient because the image distortions can be rendered and watched easily. The monochrome image shouldn't be of small format, otherwise the classification results will just stay theoretical, without power to be propagated on other scaled objects. And for accelerating the investigation procedures and acquiring the classification results the monochrome image shouldn't be of large format. Thus an appropriate format for the monochrome image is $60 \times 80$. The file format is bitmap, which is naturally coded by MATLAB with ones and zeros. So, now there is the finite general totality of $60 \times 80$ matrices of ones and zeros, containing altogether $2^{4800}$ monochrome images.

For defining the number of classes and their representatives within the said general totality, a real image should rather be selected, that would contain some generalized characteristics of any $60 \times 80$ monochrome image (horizontal and vertical lines, slants, circles, squares, curves or serpentine lines, and so on). So, let's take the enlarged English alphabet capital letter, wherein the number of classes is 26 (figure 1).

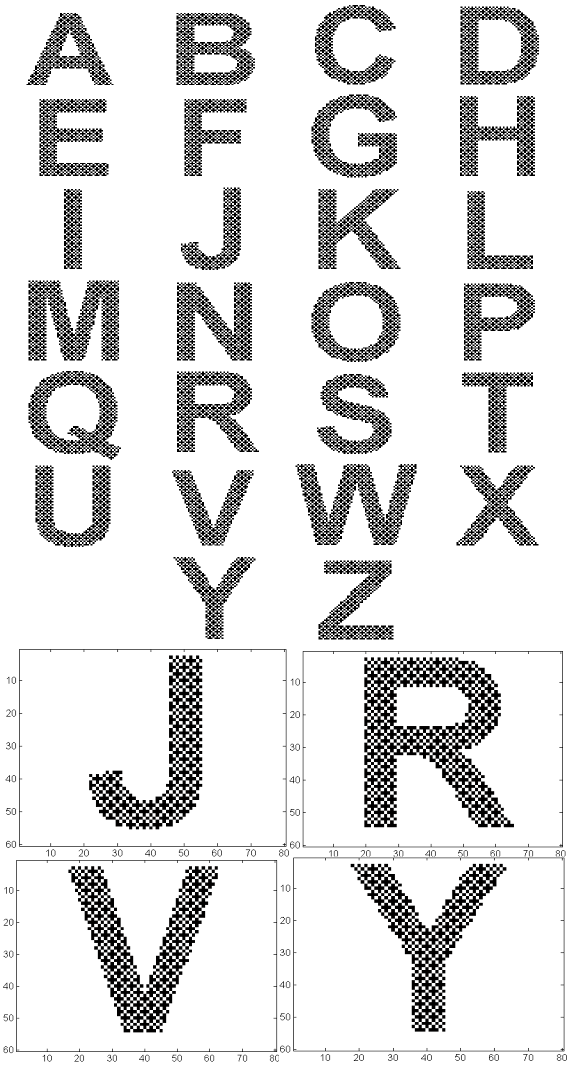


Figure 1 – Non-distorted representatives of 26 classes in the general totality of $60 \times 80$ monochrome images, viewed as bitmap files, and letters "J", "R", "V", "Y", viewed from within MATLAB

One can see in figure 1 that those non-distorted representatives of 26 classes in the general totality have regular crosshatching with white specks on the letter black cast. Such crosshatching is probably better for a classifier rather than the continuous cast on the continuous background.

*Number of neurons in perceptron hidden layers and the training MATLAB function.* Clearly that perceptron for classifying scaled images must have more than the single layer. There is a strong proof [20, 21] that for multilayer perceptron it is sufficient to have two layers (single hidden layer) for any classification problem. The only matter is of the traintime. And some earlier explorations allude that with two or three hidden layers in perceptron the traintime duration doesn't shorten, and CEP doesn't decrease in the being considered classification problem [20, 22]. Hence, there will be the single hidden layer in perceptron. Its size can be set to 250 neurons, experimentally used before over 26 classes of $60 \times 80$ monochrome images. To that, the fastest method in backpropagation algorithm for training the two-layer perceptron (TLP) is implemented as the training

MATLAB function “traingda”. It’s a training function that updates weight and bias values according to gradient descent with adaptive learning rate [18, 23, 24]. Henceforward the training MATLAB function “traingda” will be used in any training process.

*Range of pixel-to-scale standard deviations ratio.* As it was said above, the training set for classifying scaled monochrome images is formed as addition of matrix of pixel-distortion and matrix of scaled images. Namely, at the $k$-th stage of formation of pixel-distorted scaled monochrome 60-by-80-images (PDSM6080I) set, this addition is

$$\mathbf{A}_{\mathrm{PDSM6080I}}^{\langle k \rangle} = \mathbf{A}_{\mathrm{SM6080I}}^{\langle k \rangle} + \sigma_{\mathrm{PD}}^{\langle k \rangle} \cdot \mathbf{\Xi} \tag{2}$$

by standard deviation (SD)

$$\sigma_{\mathrm{PD}}^{\langle k \rangle} = \frac{k}{F} \cdot \sigma_{\mathrm{PD}}^{\langle \max \rangle} \quad \forall\, k = \overline{1, F} \tag{3}$$

and its maximum $\sigma_{\mathrm{PD}}^{\langle \max \rangle} > 0$ at $4800 \times 26$ matrix $\mathbf{\Xi}$ of values of normal variate (NV) with zero expectation and unit variance (ZEUV), where $\mathbf{A}_{\mathrm{PDSM6080I}}^{\langle k \rangle}$ is $4800 \times 26$ matrix of PDSM6080I, whose $q$-th column is the $q$-th class representative, reshaped into 4800-length-column, and $\mathbf{A}_{\mathrm{SM6080I}}^{\langle k \rangle}$ is $4800 \times 26$ matrix of scaled monochrome 60-by-80-images (SM6080I), whose $q$-th column is the same $q$-th class representative, 4800-length-column-reshaped, $q = \overline{1, 26}$. In the assignment (3) for the $k$-th SD the number $F$ indicates at smoothness in forming the training set [11] with (2): the greater this number, the smoother training process is. The matrix $\mathbf{A}_{\mathrm{SM6080I}}^{\langle k \rangle} = \left[ \tilde{\tilde{a}}_{jq}(k) \right]_{4800 \times 26}$ is formed by concatenating horizontally 4800-length-column-reshaped matrices $\left\{ \tilde{\mathbf{A}}_q(k) \right\}_{q=1}^{26}$, where $q$-th SM6080I as the matrix $\tilde{\mathbf{A}}_q(k) = \left[ \tilde{a}_{uv}^{\langle q \rangle}(k) \right]_{60 \times 80}$ is the $q$-th class representative.

An SM6080I can be formed within MATLAB by means of the function “imresize”. The map $\rho$, implementing this function, is applied to the image $\mathbf{A}_q = \left( a_{uv}^{\langle q \rangle} \right)_{60 \times 80}$ of the $q$-th class non-distorted representative of elements $a_{uv}^{\langle q \rangle} \in \{0, 1\}$ as

$$\tilde{\mathbf{A}}_q(k) = \rho\left( \mathbf{A}_q, \varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right), 60, 80 \right) \tag{4}$$

with the scale coefficient $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)$ by SD

$$\sigma_{\mathrm{scale}}^{\langle k \rangle} = \frac{k}{F} \cdot \sigma_{\mathrm{scale}}^{\langle \max \rangle} \quad \forall\, k = \overline{1, F} \tag{5}$$

for $\sigma_{\mathrm{scale}}^{\langle \max \rangle} > 0$ at $k$-th part of forming the training set that will feed the input of TLP, $q = \overline{1, 26}$. The scale coefficient

$$\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) = \sigma_{\mathrm{scale}}^{\langle k \rangle} \xi(k) + 1 \tag{6}$$

is determined by the value $\xi(k)$ of NV with ZEUV, raffled at the $k$-th stage of SM6080I set formation. If occurs $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) \leq 0$ then the corresponding NV with ZEUV is re-raffled until $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) > 0$. The input image $\mathbf{A}_q$ is enlarged by $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)$ times within the map (4) if $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) > 1$; the input image $\mathbf{A}_q$ is reduced by $\dfrac{1}{\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)}$ times within the map (4) if $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) < 1$; the input image $\mathbf{A}_q$ remains non-scaled if $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) = 1$ or (6) is rounded to 1 due to that

$$\left| \varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) - 1 \right| < 0.006,$$

and $\tilde{\mathbf{A}}_q(k) = \mathbf{A}_q$ as corollary. The scaled by $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) \neq 1$ image is the matrix $\mathbf{S}\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)$ of the intermediary format $V \times H$. If $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) > 1$ then the scaled image is cropped as lines of their numbers

$$\left\{ \left\{ \overline{1, N_V} \right\}, \left\{ \overline{61 + N_V, V} \right\} \right\} \tag{7}$$

and columns of their numbers

$$\left\{ \left\{ \overline{1, N_H} \right\}, \left\{ \overline{81 + N_H, H} \right\} \right\} \tag{8}$$

in the matrix $\mathbf{S}\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)$ are discarded, where $\eta(x)$ is a function, returning the integer part of the number $x$, for integers

$$N_V = \eta\left( \frac{V - 60}{2} \right) + \left( \frac{1 + \operatorname{sign} \zeta_V}{2} \cdot \operatorname{sign}|\zeta_V| \right) \cdot \operatorname{sign}\left[ \frac{V}{2} - \eta\left( \frac{V}{2} \right) \right] =$$

$$= \eta\left( \frac{V}{2} \right) - 30 + \left( \frac{1 + \operatorname{sign} \zeta_V}{2} \cdot \operatorname{sign}|\zeta_V| \right) \cdot \operatorname{sign}\left[ \frac{V}{2} - \eta\left( \frac{V}{2} \right) \right]; \tag{9}$$

$$N_H = \eta\left( \frac{H}{2} \right) - 40 + \left( \frac{1 + \operatorname{sign} \zeta_H}{2} \cdot \operatorname{sign}|\zeta_H| \right) \cdot \operatorname{sign}\left[ \frac{H}{2} - \eta\left( \frac{H}{2} \right) \right], \tag{10}$$

calculated by the values $\{\zeta_V, \zeta_H\}$ of two independent NV with ZEUV, raffled every time, when the function $\eta(x)$ is applied. If $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) < 1$ then the matrix $\mathbf{S}\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)$ intermediary format $V \times H$ is lesser than $60 \times 80$, that is $V < 60$ and $H < 80$. Therefore by $\varsigma\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right) < 1$ the scaled image is contoured rectangularly with the background white color: the matrix $\mathbf{S}\left( \sigma_{\mathrm{scale}}^{\langle k \rangle} \right)$ is padded from left for

$$N_{\mathrm{left}} = \eta\left( \frac{80 - H}{2} \right) + \left( \frac{1 + \operatorname{sign} \zeta_H}{2} \cdot \operatorname{sign}|\zeta_H| \right) \cdot \operatorname{sign}\left[ \frac{H}{2} - \eta\left( \frac{H}{2} \right) \right] \tag{11}$$

columns of ones (in MATLAB the white color is coded with ones, and the black is coded with zeros) and from right for

$$N_{\mathrm{right}} = 80 - H - N_{\mathrm{left}} \tag{12}$$

columns of ones, and it is padded from the top for

$$N_{\text{top}} = \eta\left(\frac{60-V}{2}\right)+\left(\frac{1+\operatorname{sign}\zeta_V}{2}\cdot\operatorname{sign}\left|\zeta_V\right|\right)\cdot\operatorname{sign}\left[\frac{V}{2}-\eta\left(\frac{V}{2}\right)\right] \quad (13)$$

lines of ones and from bottom for

$$N_{\text{bottom}} = 60 - V - N_{\text{top}} \quad (14)$$

lines of ones. After having been cropped, the map (4) finally returns SM6080I as $60\times 80$ matrix $\tilde{\mathbf{A}}_q(k)$.

The counter number $k$ for SD (2) and (5) increases simultaneously, so pixel-to-scale SD ratio must bind these SD:

$$\sigma_{\text{PD}}^{\langle\max\rangle} = r\sigma_{\text{scale}}^{\langle\max\rangle} \quad \text{at} \quad r>0\,. \quad (15)$$

The determinative values in (15) are the ratio $r$ and the scale maximal SD $\sigma_{\text{scale}}^{\langle\max\rangle}$. Within the stated model (4) — (14) for making randomized SM6080I, SD $\sigma_{\text{scale}}^{\langle\max\rangle}=0.2$ is sufficient. Heuristically, the ultimate maximal pixel-distortion SD is $\sigma_{\text{PD}}^{\langle\max\rangle}=2$. Therefore, the upper boundary for the ratio $r$ for (15) is $r_{\max}=10$. The lower boundary may be $r_{\min}=0.025$ as at maximal pixel-distortion SD $\sigma_{\text{PD}}^{\langle\max\rangle}=0.005$ values in the matrix $\Xi$ or value $\xi(k)$ of NV with ZEUV are too insignificant, distorting pixels in SM6080I sparsely. Thus range of pixel-to-scale SD ratio $[0.025;10]$ should primarily be run through with a rough step, what will help in determining the shape of the averaged CEP $p_{\text{error}}(r)$ to solve the problem

$$r^* \in \arg\min_{r\in[0.025;\,10]} p_{\text{error}}(r)\,. \quad (16)$$

Let the rough steps within the range $[0.025;10]$ be $\{0.025,\,0.05,\,0.1,\,1\}$, giving 30 points

$$\left\{0.025,\,0.05,\,\{0.1h\}_{h=1}^{20},\,\{2+h\}_{h=1}^{8}\right\} \quad (17)$$

within the segment $[0.025;10]$ to compute CEP in them. Once local minima are approximately determined, the range $[0.025;10]$ is narrowed to a subsegment of $[0.025;10]$ and a new step within the subsegment will be selected.

*Running through the range of pixel-to-scale SD ratio.* In the training process by PDSM6080I the input of TLP is fed with the training set

$$\left\{\tilde{\mathbf{P}}_i^{\langle\text{PDSM6080I}\rangle}\right\}_{i=1}^{C+F} = \left\{\{\mathbf{A}\}_{l=1}^{C},\,\left\{\mathbf{A}_{\text{PDSM6080I}}^{\langle k\rangle}\right\}_{k=1}^{F}\right\} \quad (18)$$

of $C$ replicas of all 26 classes non-distorted representatives and $F$ matrices of PDSM6080I through $k=\overline{1,\,F}$ by the set of identifiers

$$\{\mathbf{T}_i\}_{i=1}^{C+F} = \{\mathbf{I}\}_{i=1}^{C+F} \quad (19)$$

with identity $26\times 26$ matrix $\mathbf{I}$. The set (18), being formed by (2) — (14), is passed through TLP with identifiers (18) for $Q_{\text{pass}}$ times. After this TLP is tested with SM6080I at the range of scale SD $\sigma_{\text{scale}}$ from the minimal one up to $\sigma_{\text{scale}}^{\langle\max\rangle}$, that is $\sigma_{\text{scale}}\in[0;\,0.2]$. The range $[0;\,0.2]$ is going to be run through with the step $0.02$, which lets evaluate CEP in those 11 different points of the scale SD.

Parameters $C=2$, $F=8$, $Q_{\text{pass}}=20$ are acceptable on TLP with 250 neurons in its hidden layer to be trained with PDSM6080I for classifying SM6080I. And having run through the points (17) of the range $[0.025;10]$, a rough graph of the function $p_{\text{error}}(r)$ can be seen (figure 2).

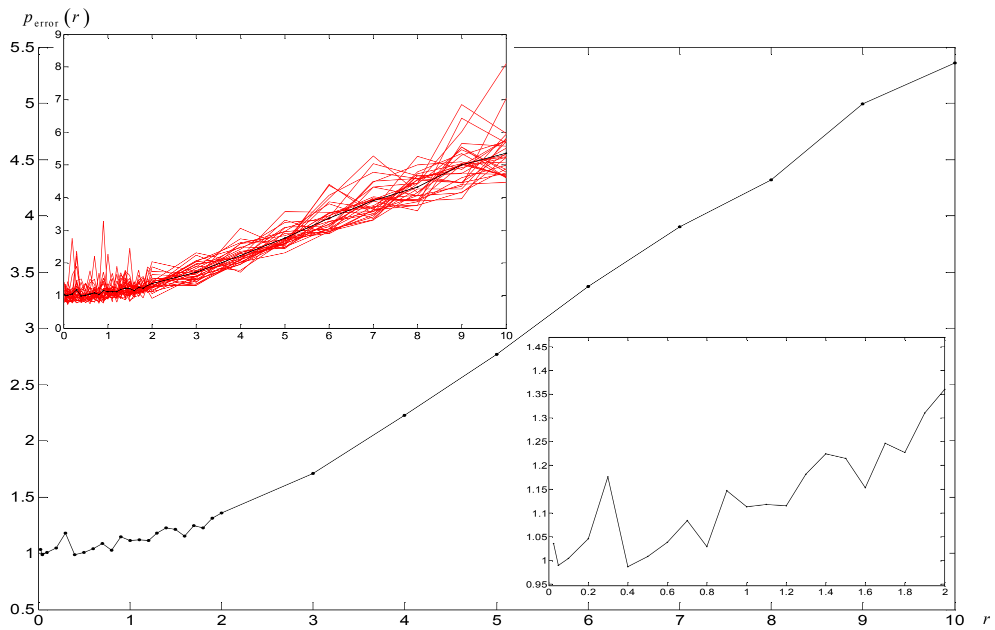


Figure 2 – A rough graph of the function $p_{\text{error}}(r)$ as 29-segmented polyline, derived from 30 series of 400 batch testings of PDSM6080I-trained at $\sigma_{\text{scale}}^{\langle\max\rangle}=0.2$ TLP for each of 30 points (17) of the range $[0.025;10]$ (CEP polylines in every series are shown over)

A zoom of the rough graph of the function $p_{\text{error}}(r)$ in figure 2 leaves an apparent motive for conclusion that $r^* < 1$. The range $[0.025; 10]$ is narrowed to the subsegment $[0.025; 1.1]$ and the new step within this subsegment is $0.01$, giving 109 points

$$\left\{0.025, \{0.02 + 0.01h\}_{h=1}^{108}\right\} \qquad (20)$$

within the subsegment $[0.025; 1.1]$ to compute CEP in them. Notwithstanding the twice greater number of batch testings on each of 50 TLP in every of those 109 points (20), the locally refined graph of the function $p_{\text{error}}(r)$ as 108-segmented polyline looks statistically volatile (figure 3). It might perhaps be concluded that $r^* < 0.5$ because $p_{\text{error}}(r) < 0.97$ at eight points by $r < 0.5$, and $p_{\text{error}}(r) > 0.97$ by $r \in [0.5; 1.1]$. And thus the problem (1) could be said being solved intervally, like $r^* \in [0.025; 0.5)$. Yes, we could further narrow the subrange to find the statistical minimum, but it would have taken huge sizes of testings and their series. Moreover, there is a feeling that the function $p_{\text{error}}(r)$ may not have a distinct minimum by $r \in [0.025; 0.5]$. Consequently, on the narrowed subrange a complementary criterion of testing the PDSM6080I-trained TLP is better to be applied.

The complementary criterion is that whether and how the PDSM6080I-trained at $r \in [0.025; 0.5]$ TLP can classify PDSM6080I. For this let TLP be tested with PDSM6080I by scale SD $\sigma_{\text{scale}}^{\langle\max\rangle} = 0.2$ at the range of pixel-distortion SD $\sigma_{\text{PD}}$ from the minimal one up to $0.5\sigma_{\text{scale}}^{\langle\max\rangle}$, that is $\sigma_{\text{PD}} \in [0; 0.1]$. The range $[0; 0.1]$ is going to be run through with the step $0.01$, which lets evaluate CEP in those 11 different points of the pixel-distortion SD. Denoting the complementary averaged CEP by $p_{\text{error}}^{\langle +\text{PD}\rangle}(r)$, there are another two polylines in figure 4, zoomed in over the subrange $[0.025; 0.5]$. And they legibly give the minimum $r^* = 0.1$ accurate to $0.025$ (by 100 series of 800 batch testings), though the point $r = 0.075$ appears close to that just by 30 series of 800 batch testings.

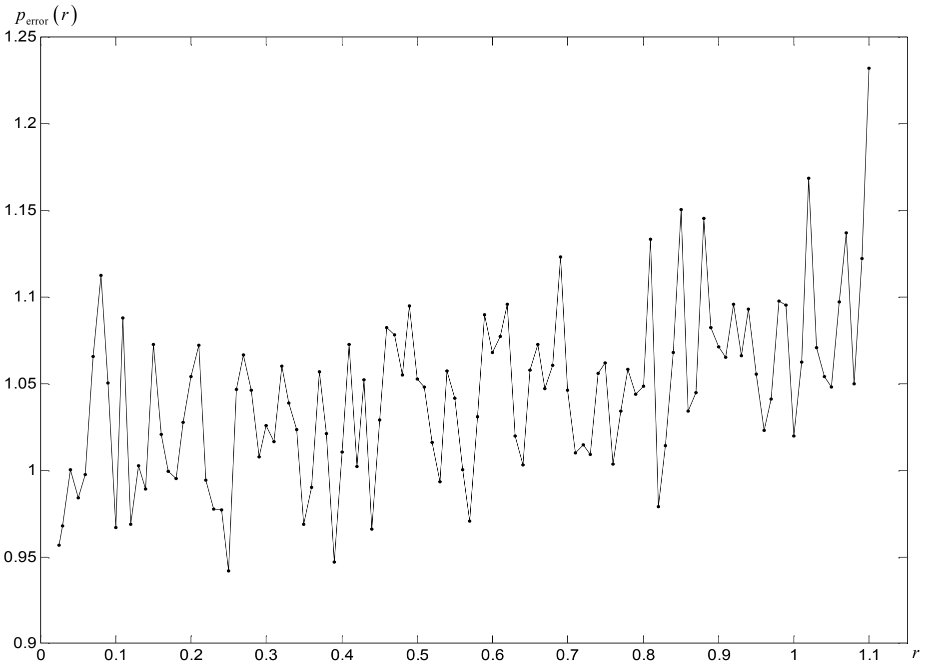


Figure 3 – The locally refined graph of the function $p_{\text{error}}(r)$ as 108-segmented polyline, derived from 50 series of 800 batch testings of PDSM6080I-trained at $\sigma_{\text{scale}}^{\langle\max\rangle} = 0.2$ TLP for each of 109 points (20) of the range $[0.025; 1.1]$

By the way, if it were to select $r^* \in \{0.075, 0.1\}$ then there would have been used another additional criterion of the traintime duration. The matter is that greater pixel-to-scale SD ratios make the training process faster [12, 20]. It's forced to be regarded if there is a multiple choice problem. In this way, having no polylines in figure 4, the optimal CEP would have been assigned to the right extremity of the segment $[0.025; 0.5]$.

*Solution of the problem (16).* Obviously, the problem (16) can be solved only graphically or numerically. Actually, the graph of the function $p_{\text{error}}(r)$ averaged polyline in figure 2, zoomed in to the polyline in figure 3, and the function $p_{\text{error}}^{\langle +\text{PD}\rangle}(r)$ averaged polylines in figure 4 clue on the minimum point $r^*$ inclines to be enclosed within the interval $(0.075; 0.125)$. Now, to affirm that $r^{*}=0.1$ is sufficiently accurate for practice, the verification procedure must be carried out. And the rounded upward CEP $p_{\text{error}}(0.1)<1.05$ and $p_{\text{error}}^{\langle +\text{PD}\rangle}(r)<4.45$ at $Q_{\text{pass}}=20$ are going to be verified below.

*Verification procedure.* Acceptability of the ratio $r^{*}=0.1$ is verified via testing the PDSM6080I-trained TLP over both SM6080I and PDSM6080I. For verifying whether CEP $p_{\text{error}}^{\langle +\text{PD}\rangle}(0.1)$ is minimal over PDSM6080I, the averaged CEP $p_{\text{error}}^{\langle +\text{PD}\rangle}(r)$ should be plotted over the interval $(0.075;\,0.125)$, enclosing the point $r^{*}=0.1$, wherein that point is its mean. Over this interval CEP $p_{\text{error}}(0.1)$ is expected to be not greater than CEP under other ratios from $(0.075;\,0.125)$. Indeed, figure 5 and figure 6 leave the verification of that

$$p_{\text{error}}^{\langle +\text{PD}\rangle}(0.1)\le p_{\text{error}}^{\langle +\text{PD}\rangle}(r)+0.03 \quad \forall\, r\in(0.075;\,0.125)\setminus\{0.1\} \tag{21}$$

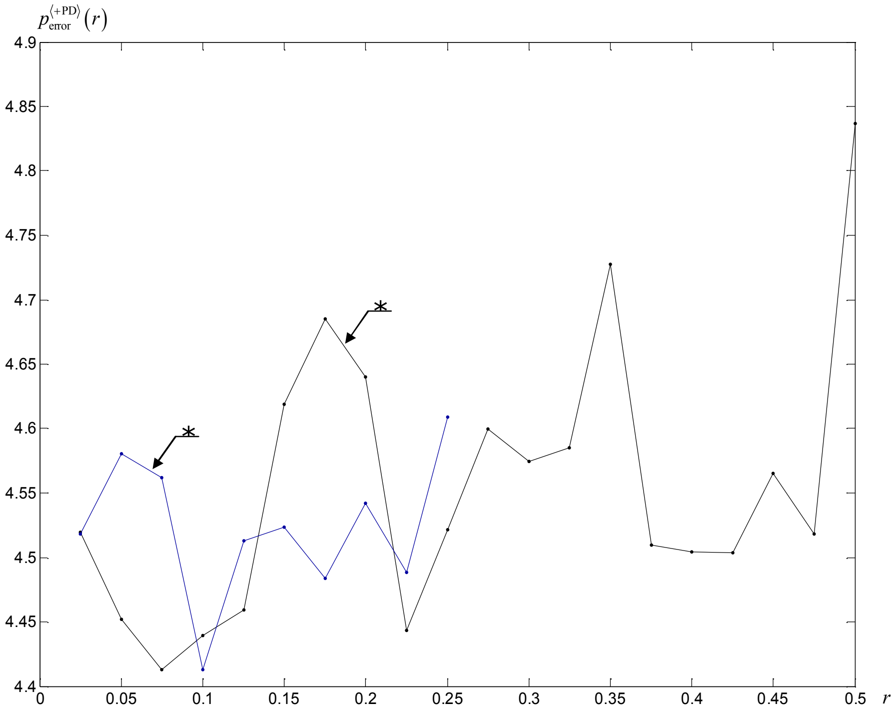


Figure 4 – The locally refined graphs of the function $p_{\text{error}}^{\langle +\text{PD}\rangle}(r)$ as 19-segmented polyline (*), derived from 30 series of 800 batch testings of PDSM6080I-trained at $\sigma_{\text{scale}}^{\langle\max\rangle}=0.2$ TLP for each of 20 points $\{0.025h\}_{h=1}^{20}$ of the range $[0.025;\,0.5]$, and as 9-segmented polyline (**), derived from 100 series of 800 batch testings of PDSM6080I-trained at $\sigma_{\text{scale}}^{\langle\max\rangle}=0.2$ TLP for each of 10 points $\{0.025h\}_{h=1}^{10}$ of the half-range $[0.025;\,0.25]\subset[0.025;\,0.5]$

and

$$p_{\text{error}}(0.1)\le p_{\text{error}}(r)+0.03 \quad \forall\, r\in(0.075;\,0.125)\setminus\{0.1\} \tag{22}$$

by $Q_{\text{pass}}\in\{20,\,30,\,40,\,50\}$. And with figures 2 — 4 the statements (21) and (22) extend to that

$$p_{\text{error}}(0.1)<p_{\text{error}}(r)+\varepsilon, \quad \forall\, r\in[0.025;\,10]\setminus\{0.1\} \text{ for } \varepsilon\le 0.03, \tag{23}$$

where $\varepsilon$ is considered as an ultimate rebate owing to statistical noise. Consequently, the solution of the problem (16) $r^{*}=0.1$ has been verified. The best averaged CEP, which is the performance of PDSM6080I-trained TLP at the ratio $r^{*}=0.1$ for

$$\left\{\sigma_{\text{scale}}^{\langle\max\rangle}=0.2,\,\sigma_{\text{PD}}^{\langle\max\rangle}=0.02\right\},$$

has been verified as well for $Q_{\text{pass}}=50$, where $p_{\text{error}}(0.1)<0.681$, though the minimal CEP from a single TLP has been tracked at $r=0.075$, where $p_{\text{error}}(0.075)<0.44$. Amazingly enough that the best averaged CEP over PDSM6080I has been tracked at

$$r=0.125 \text{ by } p_{\text{error}}^{\langle +\text{PD}\rangle}(0.125)<3.34,$$

though the minimal CEP from a single TLP has been tracked at $r^{*}=0.1$, where $p_{\text{error}}^{\langle +\text{PD}\rangle}(0.1)<2.52$. These unstable verification conclusions are generated due to statistical volatility at such smaller values of CEP. They could be probably stabilized if the batch testings' series number of PDSM6080I-trained was much greater.

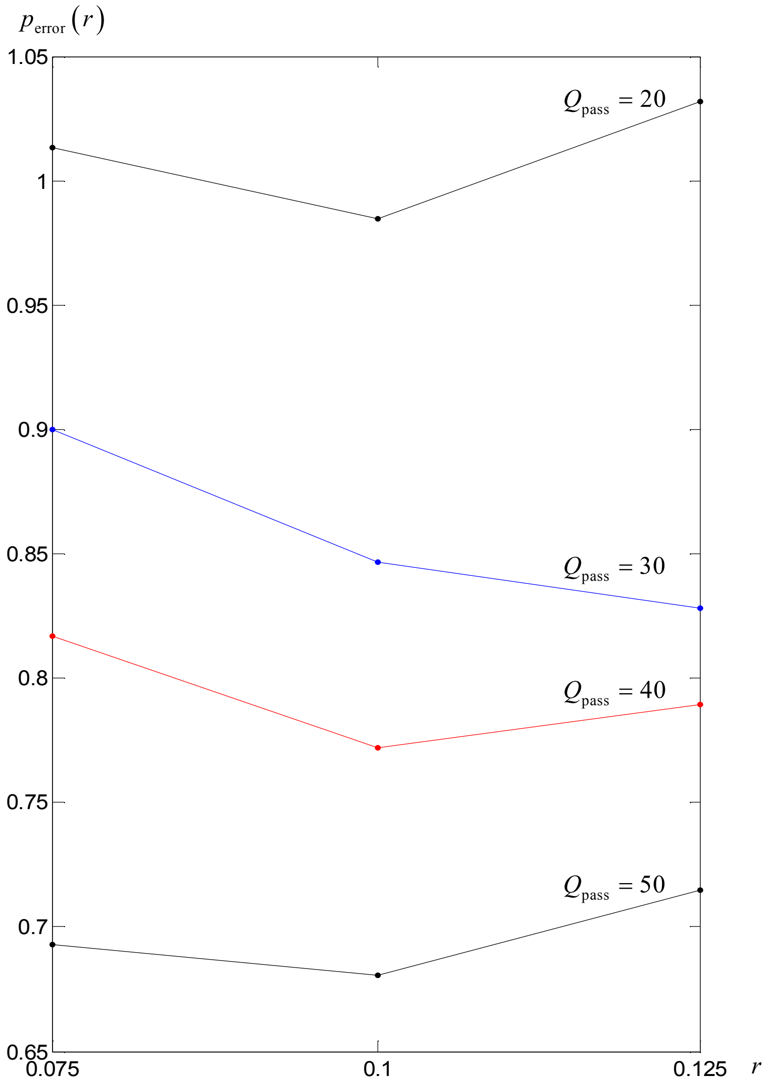


Figure 5 – The localized graph of the function $p_{\text{error}}^{\langle +\text{PD}\rangle}(r)$ about the point $r^* = 0.1$, derived from 200 batch testings of PDSM6080I-trained at $\sigma_{\text{scale}}^{\langle\max\rangle} = 0.2$ TLP for $Q_{\text{pass}} \in \{20, 30, 40, 50\}$, where are 100 TLP at $Q_{\text{pass}} = 20$ and 50 TLP at $Q_{\text{pass}} \in \{30, 40, 50\}$

Figure 7 is a try to visualize the performance of PDSM6080I-trained at

$$\left\{\sigma_{\text{scale}}^{\langle\max\rangle} = 0.2,\ \sigma_{\text{PD}}^{\langle\max\rangle} = 0.02\right\}$$

TLP, tested at the highest scale SD. It's a show of how much the image is distorted when it's been scaled at SD $\sigma_{\text{scale}} = 0.2$, and that due to the re-tracked minimal CEP $p_{\text{error}}(0.1) < 0.44$ here only 1 from 45 SM6080I is classified wrong at such SD, and only 1 from 228 SM6080I is classified wrong on average by $\sigma_{\text{scale}} \in [0; 0.2]$.

It ought to be underscored that along with the minimization problem (16) the problem of minimizing the traintime duration might have been stated. However, there are two reasons persuaded not to do that. The first one is that the traintime impacts mainly if the training process recurs periodically. And classification of SM6080I by PDSM6080I-trained classifier doesn't imply re-training. The second reason is that the two-criterion minimization problem wouldn't have had its solution, unless to acknowledge as the solution the equilibrium point [12], being the point of intersection of the locally increasing CEP and the locally decreasing traintime duration against the pixel-to-scale SD ratio.

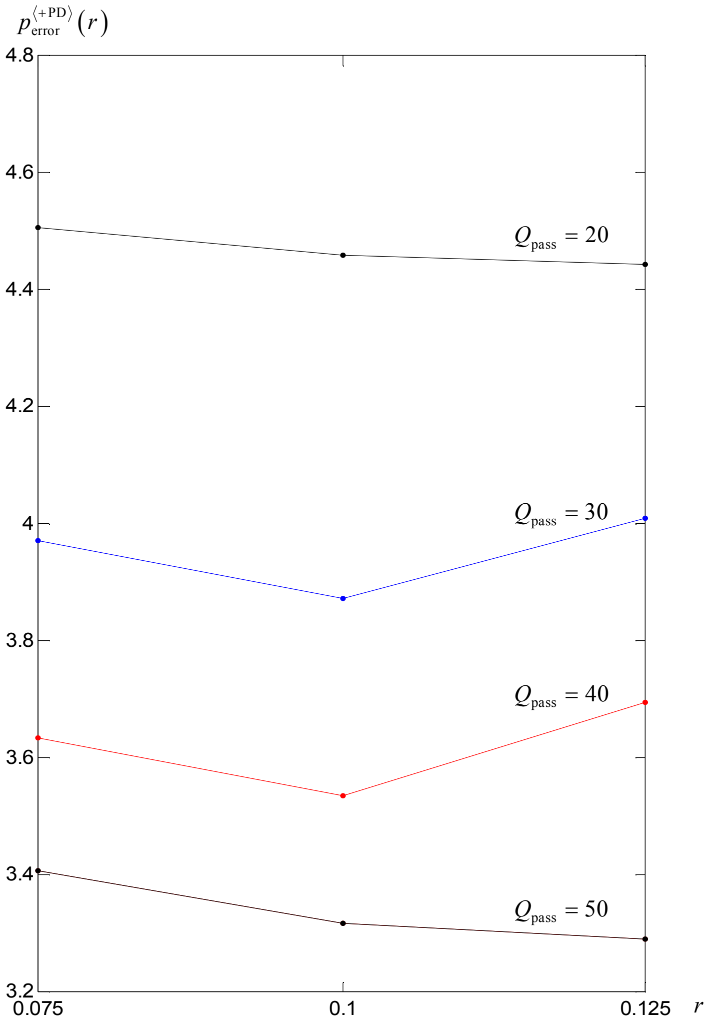


Figure 6 – The localized graph of the function $p_{\text{error}}(r)$ about the point $r^* = 0.1$, derived from 200 batch testings of PDSM6080I-trained at $\sigma_{\text{scale}}^{\langle \max \rangle} = 0.2$ TLP for $Q_{\text{pass}} \in \{20, 30, 40, 50\}$, where are 100 TLP at $Q_{\text{pass}} = 20$, 50 TLP at $Q_{\text{pass}} = 30$, 46 TLP at $Q_{\text{pass}} = 30$, and 43 TLP at $Q_{\text{pass}} = 50$

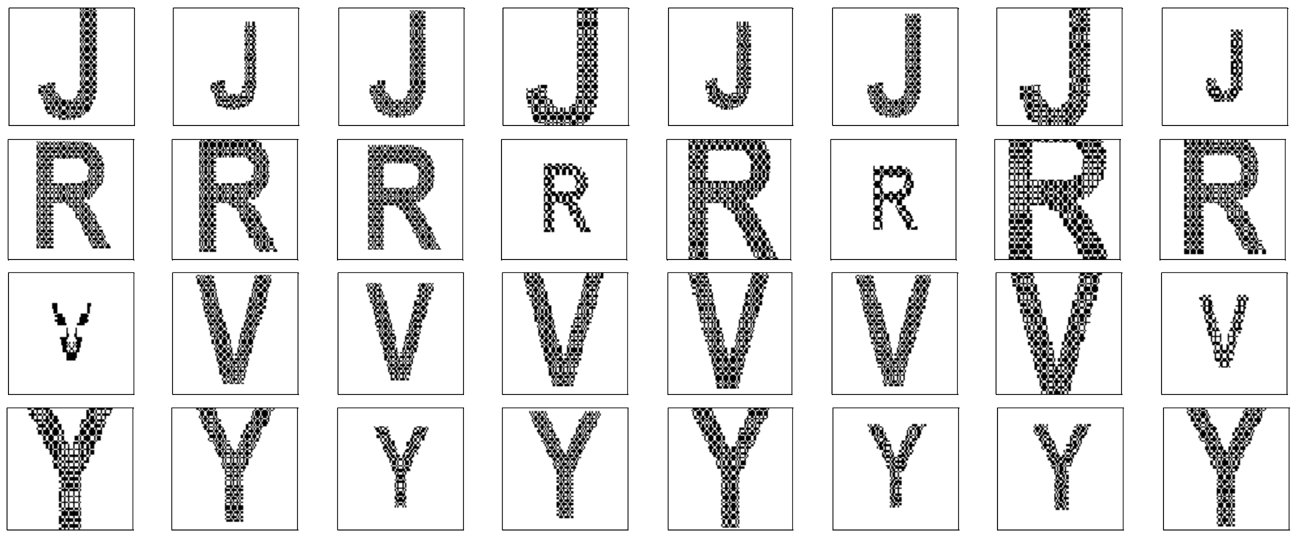

Figure 7 – SM6080I of letters “J”, “R”, “V”, “Y” at scale SD $\sigma_{\text{scale}} = 0.2$, viewed from within MATLAB

CONCLUSIONS. Nonetheless the problem (16) has been solved, for other general totality the ratio $r^*$ may be different. Also the ratio $r^*$ may change for other object model. Nevertheless, the procedure to find the solution of the problem (1) stays the same, and scaled objects, consisting of about 4800 features, will be classified accurately. Besides, there has been considered the averaged CEP, which included all the range of CEP, generated at the scale SD $\sigma_{\text{scale}} \in [0; 0.2]$. If there had been considered the maximal CEP, generated at the highest SD $\sigma_{\text{scale}} = 0.2$ or at the scale SD $\sigma_{\text{scale}} \in [1.5; 0.2]$, the corresponding minimum point of CEP $p_{\text{error}}(r)$ would have been probably different from $r^* = 0.1$, although the difference couldn't be large in regard to the range $[0.025; 10]$. The said difference might appear more significant if a multicriterion problem of minimizing CEP $p_{\text{error}}(r)$ was considered.

But if even PDSM6080I traintime duration is desirable to be shortened, accuracies in classifying SM6080I and PDSM6080I still can be perfected with optimizing simultaneously both the pixel-to-scale SD ratio and the hidden layer neurons number. This is a subsequent problem of minimizing the function of two variables. Its solution may correct the ratio $r^*$, and instead both CEP and traintime will probably be lowered and shortened correspondingly.

REFERENCES

1. Wei, W. and Xin, Y. (2010), "Rapid, man-made object morphological segmentation for aerial images using a multi-scaled, geometric image analysis", *Image and Vision Computing*, vol. 28, iss. 4, pp. 626–633.

2. Mokhtarian, F., Khalili, N., and Yuen, P. (2001), "Multi-scale free-form 3D object recognition using 3D models", *Image and Vision Computing*, vol. 19, iss. 5, pp. 271–281.

3. Fukushima, K. (2011), "Increasing robustness against background noise: Visual pattern recognition by a neocognitron", *Neural Networks*, vol. 24, iss. 7, pp. 767–778.

4. Fukushima, K. (2013), "Artificial vision by multilayered neural networks: Neocognitron and its advances", *Neural Networks*, vol. 37, pp. 103–111.

5. Castillo, P. A., Merelo, J. J., Arenas, M. G., and Romero, G. (2007), "Comparing evolutionary hybrid systems for design and optimization of multilayer perceptron structure along training parameters", *Information Sciences*, vol. 177, iss. 14, pp. 2884–2905.

6. Nor Ashidi Mat Isa, and Wan Mohd Fahmi Wan Mamat (2011), "Clustered-Hybrid Multilayer Perceptron network for pattern recognition application", *Applied Soft Computing*, vol. 11, iss. 1, pp. 1457–1466.

7. Stäger, F. and Agarwal, M. (1997), "Three Methods to Speed up the Training of Feedforward and Feedback Perceptrons", *Neural Networks*, vol. 10, iss. 8, pp. 1435–1443.

8. Pham, D. T. and Sagiroglu, S. (2001), "Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms", *International Journal of Machine Tools and Manufacture*, vol. 41, iss. 3, pp. 419–430.

9. González, A. and Dorronsoro, J. R. (2008), "Natural conjugate gradient training of multilayer perceptrons", *Neurocomputing*, vol. 71, iss. 13–15, pp. 2499–2506.

10. Hoi, K. I., Yuen, K. V., and Mok, K. M. (2013), "Improvement of the multilayer perceptron for air quality modelling through an adaptive learning scheme", *Computers & Geosciences*, vol. 59, pp. 148– 155.

11. Romanuke, V. V. (2013), "Dependence of performance of feed-forward neuronet with single hidden layer of neurons against its training smoothness on noised replicas of pattern alphabet", *Bulletin of Khmelnitskiy National University. Technical Sciences*, no. 1, pp. 201–206.

12. Romanuke, V. V. (2013), "Setting the Hidden Layer Neuron Number in Feedforward Neural Network for an Image Recognition Problem under Gaussian Noise of Distortion", *Computer and Information Science*, vol. 6, iss. 2, pp. 38–54.

13. You, J., and Cohen, H. A. (1993), "Classification and segmentation of rotated and scaled textured images using texture "tuned" masks", *Pattern Recognition*, vol. 26, iss. 2, pp. 245–258.

14. Ballabio, D. and Vasighi, M. (2012), "A MATLAB toolbox for Self Organizing Maps and supervised neural network learning strategies", *Chemometrics and Intelligent Laboratory Systems*, vol. 118, pp. 24–32.

15. Kuzmanovski, I. and Novič, M. (2008), "Counter-propagation neural networks in Matlab", *Chemometrics and Intelligent Laboratory Systems*, vol. 90, iss. 1, pp. 84–91.

16. Han, M., and Yin, J. (2008), "The hidden neurons selection of the wavelet networks using support vector machines and ridge regression". *Neurocomputing*, vol. 72, iss. 1–3, pp. 471–479.

17. Yuan, H. C., Xiong, F. L., and Huai, X. Y. (2003), "A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy", *Computers and Electronics in Agriculture*, vol. 40, iss. 1–3, pp. 57–64.

18. Hagan, M. T. and Menhaj, M. B. (1994), "Training feedforward networks with the Marquardt algorithm", *IEEE Transactions on Neural Networks*, vol. 5, iss. 6, pp. 989–993.

19. Yu, C., Manry, M. T., Li, J., and Narasimha, P. L. (2006), "An efficient hidden layer training method for the multilayer perceptron", *Neurocomputing*, vol. 70, iss. 1–3, pp. 525–535.

20. Bershad, N. J., Shynk, J. J., Vaughn, J. L., and Cowan, C. F. N. (1995), "Stochastic convergence analysis of a partially adaptive two-layer perceptron using a system identification model", *Signal Processing*, vol. 43, iss. 3, pp. 279–292.

21. Bhattacharya, U. and Parui, S. K. (1995), "On the rate of convergence of perceptron learning", *Pattern Recognition Letters*, vol. 16, iss. 5, pp. 491–497.

22. Haykin, S. (1999), *Neural Networks: A Comprehensive Foundation*, New Jersey Prentice Hall, Inc., USA.

23. Nied, A., Seleme, S. I., Jr., Parma, G. G., and Menezes, B. R. (2007), "On-line neural training algorithm with sliding mode control and adaptive learning rate", *Neurocomputing*, vol. 70, iss. 16–18, pp. 2687–2691.

24. Yoo, S. J., Park, J. B., and Choi, Y. H. (2007), "Indirect adaptive control of nonlinear dynamic systems using self recurrent wavelet neural networks via adaptive learning rates", *Information Sciences*, vol. 177, iss. 15, pp. 3074–3098.

## ОПТИМИЗАЦИЯ СООТНОШЕНИЯ СРЕДНЕКВАДРАТИЧНЫХ ОТКЛОНЕНИЙ ПИКСЕЛЬНЫХ ИСКАЖЕНИЙ И МАСШТАБИРОВАНИЯ ДЛЯ ОБУЧЕНИЯ ДВУХСЛОЙНОГО ПЕРСЕПТРОНА НА МАСШТАБИРУЕМЫХ ИЗОБРАЖЕНИЯХ ФОРМАТА 60-НА-80 С ПИКСЕЛЬНЫМИ ИСКАЖЕНИЯМИ В ЗАДАЧЕ КЛАССИФИКАЦИИ МАСШТАБИРУЕМЫХ ОБЪЕКТОВ

**Романюк В. В.**

Хмельницкий национальный университет

ул. Институтская, 11, 29000, г. Хмельницкий, Украина. E-mail: romanukevadimv@mail.ru

Решается оптимизационная задача для идентификации классификатора масштабируемых объектов на основе двухслойного персептрона. Генеральная совокупность формируется из монохромных изображений формата 60-на-80 букв алфавита, и она состоит из 26 классов. Двухслойный персептрон моделируется, обучается и тестируется при 250 нейронах в его скрытом слое и MATLAB-функции для обучения "traingda". Оптимизационным параметром является соотношение среднеквадратичных отклонений пиксельных искажений и масштабирования для надлежащего формирования обучающих множеств. Оптимизационным критерием является минимизация процента ошибок классификации. Оптимальное соотношение статистически отслеживается при прохождении обучающих множеств через персептрон на протяжении 20 раз. Со временем, увеличивая число прохождений до 50 при оптимальном соотношении, статистически отслежено минимум процента ошибок классификации. Для того, чтобы избавиться нечёткости в выводе относительно минимизации, применяется добавочный критерий, где классификатор тестируется как на исключительно масштабируемых изображениях, так и на масштабируемых изображениях с пиксельными искажениями. В процедуре верификации, когда применяется обученный при оптимальном соотношении классификатор, выявлено, что процент ошибок классификации меньше 0.44 в среднем по диапазону масштабирования.

**Ключевые слова:** классификация масштабируемых объектов, двухслойный персептрон, обучающее множество, процент ошибок классификации, монохромное изображение, оптимизация соотношения среднеквадратичных отклонений пиксельных искажений и масштабирования.

Стаття надійшла 11.03.2014.