Scientific Journal of RigaTechnicalUniversity
Computer Science. Information Technology and Management Science
2010
Volume 44

Research of Artificial Neural Networks Abilities in Printed Words Recognition

AndreyBondarenko, *Riga Technical University*, Arkady Borisov, *Riga Technical University*

***Abstract*–This paper provides a brief overview on document analysis and recognition area, highlighting main steps and modules that are used to build recognition systems of the mentioned type. We underline basic workflow of such system down to the problem of single character recognition problem and highlighting possibilities and ways for artificial neural networks usage. Further we are conductinga formal comparison of abilities of printed characters recognition between two well known types of second generation neural networks, namely feed-forward back-propagation multilayer perceptron (MLP) and Kohonen self-organizing features map (SOM).**

***Keywords* –document image processing, optical character recognition, neural networks**

I. Introduction

The last decade has showna growing interest in artificial neural networks (ANN),their functioning background, learning algorithms, architectures and practical application. Good introduction material can be found in [1] and [2].Document analysis and recognition systems are using artificial neural networks on different analysis and recognition steps. Optical character recognition systems are of great interest because the count of paper documents which are part of business processes of every company, institution or even household is constantly increasing. Because of computer capabilities to persist, process, search and mine knowledge on large amounts of electronic documents the mentioned systems capable of automatic translation of paper documents into the digital format are ofgreat interest. Neural networks are successfully applied in the above-mentioned area due to their abilities to give rather low classification errors even when they are applied to disrupted and noisy data. Recently specialized approaches were developed to enable ANN usage in mentioned area. One of the basic components of such system is optical character recognition module which is in charge for recognition of separate characters. According to thelatest studies,while reading textsthe human brain is perceiving information in chunks of letters in defined context using some knowledge about subject area, thus even noisy letters or words can be recognized relatively easy using contextual information. This approach is extensively used by document analysis systems. Nevertheless, basic system module capable to classifying single character should give as high classification rate as possible. Therefore research of algorithms for printed characters recognition is ofgreat interest.

The comparison of performance of multilayer feed-forward perceptron (MLP) and self-organizing features maps (SOM) presents practical and scientific interest. In this paper we are providing a comparative analysis of the effectiveness of the mentioned architectures for pattern recognition of printed letters images. It should be noted that abstract features like relative placement of elements in character or junction patterns and their count are omitted in our experiments due to the fact that raster image was fed into the MLPand SOM [2] networks in the form of one-dimensional array. It was shown that MLP is more preferable for solving character recognition problem with described input data.

This paper is organized as follows: Section II gives an overview of the document analysis and recognition system as a whole. Subsections II.A, II.B, II.C and II.D are focused on neural networks applications on different stages of document image processing. Section III describes experimental setup for neural networks performance measuring and presents the results of test runs. Section IV contains analysis and concludes the paper.

II. Document Analysis and Recognition

Document analysis and recognition problem consists of several sub problems. As per [3] we can name basic modules for such a system along with their respective roles. *Document image preprocessing* is a phase where noise reduction, brightnessbalancing and skew detection and correction take place. *Layout analysis* in this step system is searching for regions of similar content, which usually is either text or graphicsbut other mixed types as well as special types like signature can be present, this step is called *physical layout analysis*. Detected regions are marked with respective recognized type so that later they will be processed with appropriate module.This step is called *logical layout analysis*. As a next step system feeds found region into one of the three modules: *signature segmentation*, *graphic item segmentation* and *word recognition*. Once signature segmentation is done, bounded signature image is fed into the *signature verification module*. Output of graphic segmentation module is fed into *graphic item recognition* module which gives digitized graphic as a result. Finally, segmented words can be recognized in several ways: *character segmentation and recognition*, *word recognition* and *mixed approach* which uses recognized characters data along with recognized words data.The described parts and general workflow of the system areshown in Figure 1.

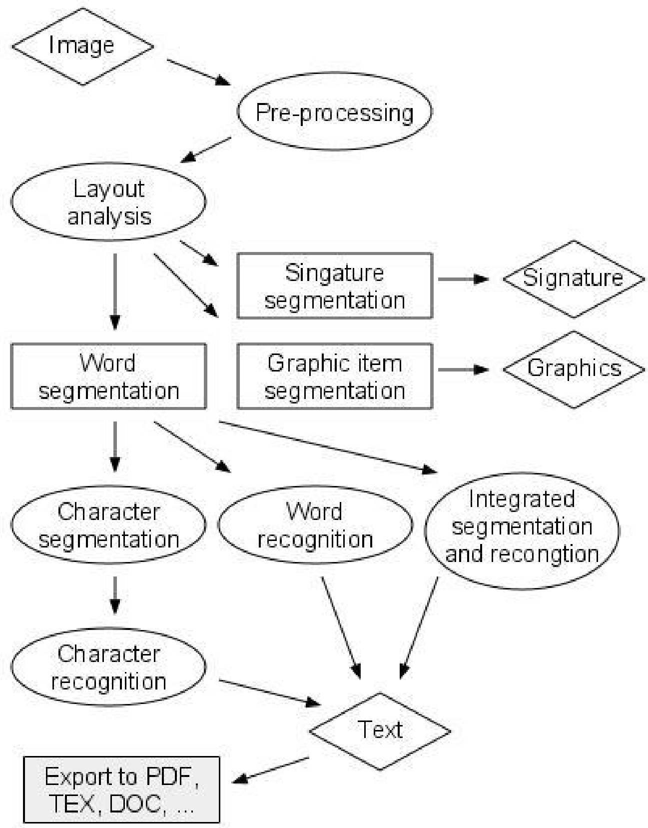

Fig. 1. Document analysis and recognition workflow. Oval figures are denoting workflow stages where artificial neural networks can be successfully applied. Diamonds are representing information which is passed in between stages.

A. Layout Analysis / Zoning

Segmentation algorithms are usually divided into pixel classification and region-based classification, while artificial neural networks are usually applied to problems of next categories: page classification, pixel classification and region classification.

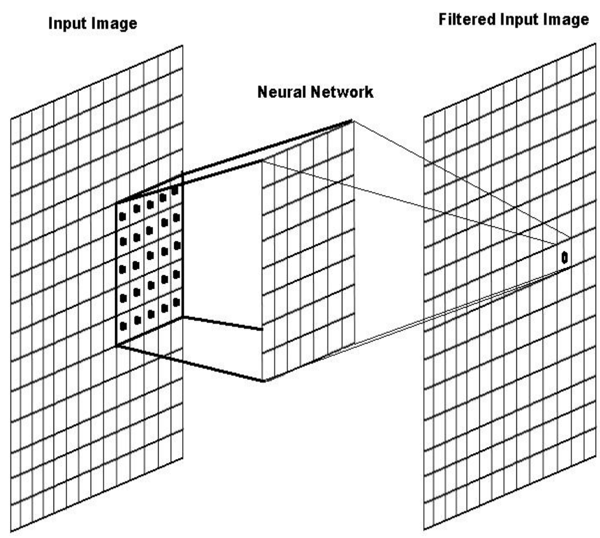

Fig. 2. Filtering process utilizing neural network to get cleared pixel value on the output image.

Skewed imagesare severely limiting the rate of successful image segmentation. Merged and broken graphicsand characters as well as non-constant background are examples of not welcome properties of image which is fed into zone detection module. Therefore detection of such features is essential for successful image analysis [4]. To overcome skew problem, artificial MLP networks are successfully applied. To do that respective properties of the image are extracted and a network is trained to produce skew angle as its output. This angle can be used for de-skewing.As well text restoration can be applied to improve text image quality. Text restoration step is used to repair broken and merging characters.This step is very useful for later phases of document analysis. Typical approaches to text image restoration are line following [5], Kalman filtering[6] and morphological filtering [7]. Although it was reported that MLP network can be used for filtering. To do that, MLP network is fed with pixels from moving window of fixed size.Filtering process is shown in Figure 2. It is impossible to use same trained neural network for the whole diversity of the document images thus it is common to retrain MLP network for each separate page.

In its simplest form pixel classification assigns each pixel to background or foreground [8]. Later pixel classification evolved to handle classification of text, graphics and lines. Several authors used neural networks for pixel classification [9] ,[10] .

For classification of the region its global features are usually extracted and are used as inputs to linear classifier [11]. When neural networks are applied to this problem same region global features along with local ones are used. Local features are computed for each pixel of the region [12]. It was shown [13] that on this problem RBF networks are performing better than MLP, SOM and Probabilistic Neural Network.

The problem of page classification addresses different methods aiming at different goals. Earlier methods required the presence of ruling lines in pre-printed form layout [14]. In business applications typical classes are business letters and technical papers, but lately classification of book pages and journal pages received significant attention [15].Page layout can be described by zoning or by using tree [16] or graph based [17] representation.The problem of fixed size of input window of the neural networks for the classification is usually dealt either by using recurrent neural networks or via formation of feature vector of the fixed size.

B. Signature Segmentation

Signature segmentation is needed for its verification. Usually both dynamic [18] and structure [19] are needed for the elimination of skilled forgery. Dynamics correlates with grey level variations along the line and with the stroke width of the signature. Thus extraction of such variations is essential for successful verification. The majority of the signature verification systems are dealing with the structure of the signature global and local graphometric features.

C. Text / Graphical Segmentation

Text and graphical segmentation can be represented as two class classification problem as assigning region with certainclass. For the classification of the document region as text/non-textmultiple approaches are used. Approaches used to tackle this problem can be divided into: *top-down* and *bottom-up*. The most common *top-down* techniques are run-length smoothing [20]-[21] and *projection profiles* [22]-[25]. *Top-down* approaches are splitting image into regions which are later identified and further spitted into text columns then paragraphs, text lines and finally words [23]. Such approach is showing poor performance on non-rectangular blocks of text.*Bottom-up* methods [26], [27] are typically iteratively grouping elements starting from the pixel level. Later such groups are forming higher order components of words, lines and paragraphs [28]. Neural networks can be applied in this area [29].

D. Word Segmentation and Recognition

The existing strategies for splitting text into separate words and characterscan be divided into three main groups: *character segmentation and recognition*, where image is divided into regions which are matching character properties and later each character image is recognized, *word recognition approach*, which is dealing with recognition of the whole words avoiding segmentation and, finally *integrated approach* which is a mix of previous two approaches, it integrates the results of both words recognition and character segmentation and recognition. Neural networks can be applied to character segmentation (identification of touching characters and location of cutting points). The results of our experiments can be applied in this phase as we are comparing networks performance for the character recognition module.

III. Experiments Setup and Results

For formal comparison of MLP and SOM network we are using scanned image of page with 16 machine printed lines, each containing all Latin alphabet letters printed in capital. Each character line was printed using different font and each one contains 26 characters. Segmentation technique used was based on the assumption that there are no active (dark enough) pixels between rows of text in different lines of text and between characters in the same line. For resizing separate characters, antialiasing was applied. Resizing algorithm used in the experiments is resizing character to the rectangular image with equal height and weight.After resizing with antialiasing some pixels are losing their brightness, i.e. black pixels located on the boundaries of character are thinning and becoming too bright to give high enough input to the neural network. To overcome this immediately after resizing nonlinear contrast correction is applied according to formula 1.

$$Br = \min\begin{cases} 2 * Br^{1.5} \\ 1 \end{cases}. \qquad (1)$$

Where *Br* denotes the brightness of the pixel and belongs to domain [0, 1].Thus semi-black pixels are transformed into ‘pure’ black. Such sequence of transformations resembles character restoration algorithm as it gives as its output resized character with minimal count of artifacts not present in the original image.To examine abilities of neural networks and compare two architectures we setup and run several experiments with different parameters. One of the main questions was to determine how the size of the input image influences neural network ability to correctly recognize previously unseen characters.We used rectangular character images of sizes 18x18, 16x16 and 14x14 pixels.The size of the input image uniquely determines the size of the input layer of the neural network.The second question that was stated is to determine how the size of the train data set influences correct recognition rate. We used data sets of the first five and first eight fonts that are depicted inFigure 3.

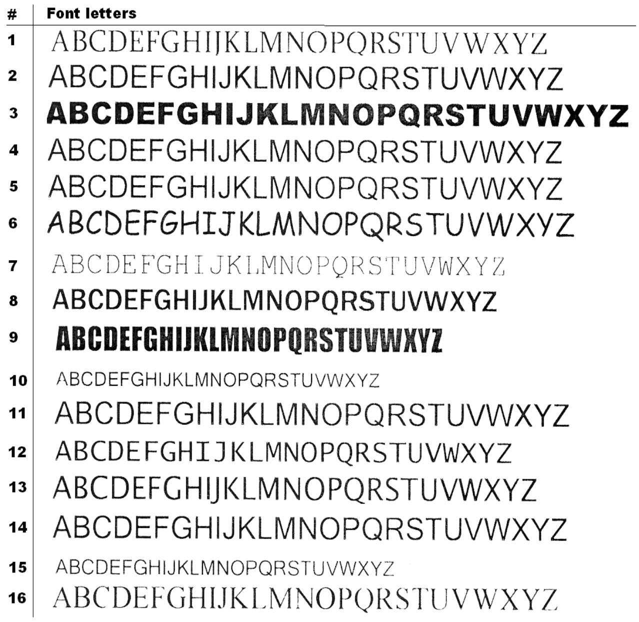

Fig. 3. Whole characters data set. Two font sets were used for the training of the neural networks. First sets contains font #1, while second set contains fonts #1, #2 and #6.

It should be noted that we could determine the best set of five and eight fonts respectively that are giving best possible training for the network, by training and testing nets with different combinations of fonts fed to the network in different sequence. But in our case, the aim was to find how the size of the data set is influencing recognition rate, not determination of the optimal combination of defined size of the fonts that are describing whole data set. These experiment factors were common for both MLP and SOM neural networks experiments. The third factor specific for MLP network wasthe size of the hidden layer. We choose three different sizes: 50, 75 and 100 neurons. We used only one hidden layer as it is known that MLP with arbitrary count of hidden layers can be reduced to network with single hidden layer. At last,the fourth factor, which is also specific to MLP, was maximum number of training epochs; we choose 1500 and 3000 respectively. Thus for the MLP network we have 36 unique experiments. For the SOM network the third parameter was the size of the Kohonen layer it was chosen to be 30 and 60 elements respectively thus for the SOM network we get 12 unique experiments. Each experiment was conducted 20 times to get statistically meaningful results.

Looking from the technical point of view,experiments were carried out by Encog library [30]. We used single hidden layer for the MLP network along with *tanh*activation function. Learning rate was constant and was set to 0.8, momentum was remained to be 0.3. SOM network was run with *NeibourhoodBuble* neighborhood function. Neighborhood value for mentioned function was set to 5. We linearly decreasedlearning rate and neighborhood radius to 0.01 and 1 respectively. Epochs count was set to be 50 as already starting from 10th or 20th epoch network error was constant making further training pointless.The fact that SOM network utilizes learning without teacher makes it a little bit difficult for use as a classifier. We used a simple approach. First of all at the end of training interrogate network using all training data. During that process save information about best matching unit (BMU) neurons for each letter respectively. Filter out BMU's that are winners for more than one letter. Afterwards test the network. In case the network gives BMU not known to persisted ones, then determine closest BMU to the new winning neuron using Euclidean distance. Here we are utilizing clustering properties of the SOM network assuming that nearby (in sense of Euclidean distance) neurons are classifying same character.

The results of the experiments can be found in Table 1 and Table 2 for MLP and SOM networks respectively. Looking at MLP network recognition performance we can clearly see thatthe increase in epochs count gives stable growth of recognition rate. By saying recognition rate we are referencing to mean correctly recognized letters count across all test fonts.For both tables, the size of the number denoting recognition rate (lowest two rows for Table 1 and single lowest row for Table 2) is denoting results performance in comparison to other experiments results. Bold underlined number denotes best recognition rate, 'small'underlined number printed in small font size denotes the worst result. It can be seen that training fonts count gives better recognition results along with lower hidden neurons count. On the other hand, the usage of largest input image - 18x18 pixels in combination with only 5 training fonts gives worst results.

Looking at SOM results in Table 2 we can notice reversed input image dimensions size influence on recognition rate. The larger input image is, the better results the network can produce. SOM experiments show interesting correlation between recognition rate and Kohonen neurons layer size. The larger the layer is, the better recognition rate network can show. Same is true for input image size. Both of these correlations are topical due to cauterization nature of the SOM networks. Larger count of input neurons, i.e. dimensions can be mapped into lower dimension of Kohonen network in more diverseway.

TABLE I
MULTILAYERED PERCEPTRON: TEST FONTSMEAN RECOGNITION RATES (%)

MLP		**Image Dimensions**																	
		18x18						**16x16**						**14x14**					
		Training Fonts Count						**Training Fonts Count**						**Training Fonts Count**					
		8			**5**			**8**			**5**			**8**			**5**		
		Hidden Neurons			Hidden Neurons			Hidden Neurons			Hidden Neurons			Hidden Neurons			Hidden Neurons		
		50	75	100	50	75	100	50	75	100	50	75	100	50	75	100	50	75	100
Recognition Rate %	**1500 Train Cycles**	**59**	**54**	**49**	43	38	27	**67**	**59**	**54**	**47**	41	28	**73**	**68**	**62**	**50**	**47**	35
	3000 Train Cycles	**61**	**59**	**52**	43	40	34	**67**	**63**	**56**	**50**	42	39	**75**	**69**	**67**	**53**	**48**	43

TABLE II
SELF-ORGANIZING FEATURES MAP:TEST FONTS MEAN RECOGNITION RATES (%)

SOM	**Image Dimensions**											
	18x18				**16x16**				**14x14**			
	Training Fonts Count				**Training Fonts Count**				**Training Fonts Count**			
	8		**5**		**8**		**5**		**8**		**5**	
	Output Neurons		Output Neurons		Output Neurons		Output Neurons		Output Neurons		Output Neurons	
	30	60	30	60	30	60	30	60	30	60	30	60
Recognition Rate %	15	**37**	**18**	17	13	**30**	**21**	**21**	17	18	**19**	**20**

IV. CONCLUSION

Experimental results for both networks show that recognition rate iscorrelated with network functioning parameters. For MLP network input image with lower dimensions show better results, while for SOM network we see an opposite trend. The main reason for this is networks functioning background. For SOM network to do better classification it is crucial to have as large input image as possible. We can foresee that an uncontrolled increase in input field size starting from one point will cause smaller recognition rate. Obviously same is true for MLP network, reduction of input layer size at the end will cause loss of critical features (due to resizing) and network will be unable to recognize characters correctly. All such problems are caused by fact that we are working with optical information of lowest possible level. As well our networks are not using receptive fields like, for example, Cognitron and Neocognitron neural networks do [31]. These networks are able to tackle visual information by extracting abstract featureslike joints and their orientation. For OCR systems utilizing regular MLP and SOM networks this is usually done by preprocessing module which is extracting such abstract information from character image. Thus we clearly see two ways of development, either introduction of more advanced neural architectures that are mimicking human visual cortex and processing information in hierarchical way – by extracting abstract features, or by utilizing preprocessing modules that can extract abstract information regardless of character size, position and orientation and feed such information to regular classificatory like MLP or SOM. It should be noted that usually MLP and SOM are set up in document analysis systems in a way that output of one network is in fact input of another network. Both approaches have pros and cons, Cognitron and Neocognitron are computationally more intensive than opposite approaches (which is an issue for mobile computing), on the other hand they can substitute a large amount of simpler interconnected modules. Thus with development of such systems we can come to a situation when only several (1-3) modules containing sophisticated hierarchical networks will be present. Further research can concentrate on either of two approaches to the problem of optical character recognition.

REFERENCES

[1] **S. Haykin**, *Neural Networks: A comprehensive foundation(2nd edition).*IEEE, 1999.
[2] **Y. H. Hu**, **J.-N. Hwang**, *Handbook of neural network signal processing.*CRC Press, 2002.
[3] **M. Gori**, **S. Marinai**, **G.Soda**, *Artificial Neural Networks for Document Analysis and Recognition.* Technical Report N.1 University of Florence, 2003.
[4] **L. O'Gorman**, **R.Kasturi**, *Document Image Analysis.*Los Alamitos, California: IEEE, Computer Society Press 1995.
[5] **A. P. Whichello**, **H. Yan**, *Linking broken character borders with variable sized masks to improve recognition.* Pattern Recognition, vol. 29, no. 8, pp. 1429-1435, 1996.
[6] **M.-Y. Yoon**, **S.-W. Lee**, **J. Kim**, *Faxed image restoration using Kalman filtering.* In Proc.3rd Int'l Conf. Doc. Anal. Rec, pp. 677-680, 1995.
[7] **J. Liang**, **R. Haralick**, **I. T. Phillips**, *Document image restoration using binary morphological filters.* In Proc. SPIE – Doc. Rec. III, pp.274-285, 1996.
[8] **T. Taxt**, **P. J. Flynn**, **A. K. Jain**, *Segmentation of document images.* IEEE Trans. PAMI vol. 11, no. 12, pp. 1322-1329. 1989.
[9] **K. Etemad**, **D. S. Doermann**, **R. Chellappa**, *Multiscale segmentation of unstructured document pages using soft decision integration.* IEEE Trans. PAMI, vol. 19, no. 1, pp. 92-96, 1997.
[10] **K. Nakamura**, **S. Yamamoto**, **T. Itoh**, *Document image segmentation into text, continuous tone and screened-halftone region by the neural networks.* In ISET / SPIE, pp.358-361, 1998.
[11] **D. Wang**, **S. N. Srihari**, *Classification of newspaper image blocks using texture analysis.* Computer Vision, Graphics, and Image Processing, vol. 47, no. 3, pp. 327-352, 1989.
[12] **C. Strouthopoulos**, **N. Papamarkos**, *Text Identification for document image analysis using a neural network.* Image and Vision Computing, vol. 16, no. 12/13, pp. 879-896, 1998.
[13] **D. X. Le**, **G. R. Thoma**, **H. Wechsler**, *Classification of binary document images into textual or nontextual data blocks using neural network models.* Machine Vision and Applications, vol. 8, no. 5, pp. 289-504, 1995.
[14] **S. L. Taylor**, **R. Fritzson**, **J. A. Pastor**, *Extraction of data from preprinted forms.* Machine Vision and Applications, vol. 5, no. 5, pp. 211-222, 1992.
[15] **C. K. Shin**, **D. S. Doermann**, *Classification of document page images based on visual similarity of layout structures.* In Proc. SPIE – Doc. Rec. Retr. VII, 182-190, 2000.
[16] **G. Nagy**, **S. Seth**, *Hierarchical representation of optically scanned documents.* 7th Int. Conf. Pattern Recognition, pp. 347-349, 1984.
[17] J. Yuan, Y. Y. Tang, C. Y. Suen, *Four directional adjency graphs (FDAG) and their application in locating fields in forms.* In Proc. 3rd Int'l Conf. Doc. Anal. Rec., pp. 752-755, 1995.
[18] **M. Ammar**, **Y. Yoshida**, **T. Y. Fukurama**, *A New Effective Approach for Off-Line Verification of Signature by Using Pressure Features.* 8th ICPR, pp. 566-569, Paris, France, October 1986.
[19] **R. Sabourin**, **R. Plamondon**, *On the implementation of some graphometric techniques for interactive signature verification: A feasability study.* Proc. 3rd Int. Symp. on Handwriting and Computer Applications, pp. 160-162, Montreal Canada, July 20-23, 1987.
[20] **P. Chauvet**, **J. Lopez-Krahe**, **E. Taflin**, **H. Maitre**, *System for an intelligent office document analysis, recognition and description.* Signal Processing, vol. 32, no. 1-2, pp. 161-190, 1993.
[21] **F. Shih**, **S.-S. Chen**, **D. Hung**, **P. Ng**, *A document image segmentation, classification and recognition system.* In Proc. Int'l. Conf. Systems Integration, pp. 258-267, 1992.
[22] **O. Iwaki**, **H. Kida**, **H. Arakawa**, *A segmentation method based on office document hierarchical structure.* In Proc. Int'l Conf. Systems, Man and Cybernetics, pp. 375-390, 1987.
[23] **D. Wang**, **S. N. Srihari**, *Classification of newspaper image blocks using texture analysis.* Computer Vision, Graphical Image Processing, vol. 47, pp. 327-325, 1989.
[24] **G. Nagy**, **S. Seth**, *A prototype document image analysis for technical journals*. Computer, vol. 25, no. 7, pp. 10–22, 1992.
[25] **M. Krishnamoorthy**, **G. Nagy**, **S. Seth**, **M. Viswanathan**, *Syntactic segmentation and labeling of digitized pages from technical journals.* IEEE Trans. Patter. Anal. Machine Intell. Vol, 15, no. 7, pp. 737-747, 1993.
[26] **L. A Fletcher**, **R. Kasturi**, *A robust algorithm for text string separation from mixed text/graphics images.*IEEE Trans. Patter Anal. Machine Intell., vol. 10, pp. 910-918, Nov, 1988.
[27] **T. Saitoh**, **T. Pavlidis**, *Page segmentation without rectangle assumption.*in Proc.Int'l Conf. Patter Recognition, pp. 277-280, 1992.
[28] **C. L. Tan**, **B. Yuan**, **W. Hung**, **Z. Zang**, *Text/graphics separation using pyramid operations.* In Proc. Int'l Conf. Document Analysis and Recognition, pp. 169-172, 1999.
[29] **M. Ignacio**, **C. Marguia**, *A Fuzzy Neural Network Approach for Document Region Classification Using Human Visual Perception Features.* Computacion y Sistemas, vol. 6, no. 2, pp. 83-93, Mexico, 2002.
[30] Encog – advanced neural network and bot programming library [Online]. Available: http://code.google.com/p/encog-java/ [Accessed: Sept. 08, 2010].
[31] **K. Fukushima**. *Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unnafected by shift in position.* Biological Cybernetics, vol. 36, no. 4, pp. 93-202, 1980.

Andrey Bondarenko received the B.Sc. and Ms.Sc. degrees in 2004 and 2006 in The University of Latvia and Transport and Communication Institute in Riga, Latvia. Since 2010 he is continuing studies at The Riga Technical University to get the Ph.D degree in Computer Science. Currently major field of study is:Extraction of concise rules from learned artificial neural network.
Currently holds position of Lead Java Developer at CTCo – Riga based IT company, previously was employed at Latvian Cancer Center. Previous publication: *Decompositional Rules Extraction Methods from Neural Networks*, Proceedingsof Conference, MENDEL 2010, Brno, Check Republic, 2010.Research interests include: concise rules extraction from neural networks, neural networks, support vector machines, computer vision and cognition.E-mail: andrejs.bondarenko@gmail.com

Arkady Borisov is Professor of Computer Science in the Faculty of Computer Science and Information Technology at Riga Technical University. He holds a Doctor of Technical Sciences degree in Control in Technical Systems and the Dr. habil. sc. comp. degree. His research interests include fuzzy sets, fuzzy logic, computational intelligence and bioinformatics. He has 210 publications in the area.
He has supervised a number of national research grants and participated in the European research project ECLIPS.
Contact information: Institute of Information Technology, Riga Technical University,1 Kalku Street, Riga, LV-1658, phone: +371 67089530.E-mail:arkadijs.borisovs@cs.rtu.lv

Andrejs Bondarenko, Arkādijs Borisovs. Mākslīgo neironu tīklu iespēju pētījums drukāto vārdu atpazīšanā
Dokumentu analīzes un atpazīšanas sistēmas nozīmīgumu ir grūti pārvērtet, jo šādas sistēmas atrod sev plašu pielietojumu dažādās cilvēku darbības jomās. Šajā darbā ir dots pārskats par dokumentu analīzes un atpazīšanas sistēmas iekšējo struktūru un tās nozīmīgākajiem moduliem. Dotajā pārskatā ir aplūkotas mākslīgo neironu tīklu pielietojuma iespējas minētajā sistēmā. Darba ietvaros bija uzstādīts uzdevums salīdzināt daudzslāņu perceptronu (MLP) ar kļūdas atpakaļizplatīšanu, kurš ir apmācams ar apmācītāju, unpašorganizējošo pazīmju Kohonena karti(SOM),kura ir apmācama bez apmācītāja, lielo latīņu alfabēta drukāto burtu klasifikācijai un atpazīšanai, kuri ir drukāti, izmantojot sešpadsmit dažādus fontus. Eksperimenti bija sastādīti tā, lai atrastu korelāciju starp dažādiem neironu tīklu funkcionēšanas parametriemun tīklaiespēju klasificēt drukātus burtus no testu datiem.MLP tīkls tika testēts ar sekojošiem mainīgiem parametriem: ieejas attēla izmēri – 18x18, 16x16 un 14x14 pikseli, apmācības datu kopas lielums – 8 un 5 fonti, slēptā slāņa apjoms 50, 75 un 100 neironu, ka arī apmācības periodu daudzums – 1500 un 3000. Katrs no 36 eksperimentiem tika uzstādīts 20 reizes, lai saņemtu statistiski nozīmīgus rezultātus. Labākus rezultātus parādīja tīkls, kura slēpta slāņa izmērs bija 50 neironi, tas tika apmācīts izmantojot 8 fontus 3000 apmācības periodu laikā ar ieejas slāņa izmēriem 14x14 neironi. Vidējais veiksmīgais atpazīšanas procents šim MLP tīklam,pārbaudot to uz testu fontiem, bija 75%. SOM tīklam izņemot ieejas slāņa izmērus, apmācības un testu fontu daudzumu startp eksperimentiem tika mainīts Kohonena slāņa neironu skaits – 30 un 60 neironi. Šeit labāko vidējo atpazīšanas rezultātu - 37% parādīja tīkls, kurš tika apmācīts izmantojot 8 apmācības fontus, Kohonena slāņa 60 neironus un ieejas attēla izmērus – 18x18 pikseļi.Tādējādi, parādīts lielu ieejas un pazīmju slāņu izmērunozīmīgumslabai atpazīšanai (klasterizācijai) SOM tīklam,ka arī mazas ieejas un slēpta slāņa izmēri MLP tīklam. Salīdzināšanas gaitā tika parādīta MLP tīkla priekšrocība par SOM tīklu uzstādīta uzdevuma robežās.Nepieciešams atzīmēt, ka izskatītie neironu tīkli neizmanto receptora laukuma principus un tie nav hierarhiski tīkli, tādējādi tie nav spējīgi patstāvīgi iegūt abstraktas ieejas attēla pazīmes. Atpazīšanas iespēju palielinājums ir iespējams tikaiar SOM un MLP tīkla kombinēto izmantošanu, ieejas attēla priekšapstrādes algoritmu izmantošanu, lai atrastu invariantas simbolu pazīmes, kā arī izmantojot kompleksas hierarhiskas neironu tīklu arhitektūras.

Андрей Бондаренко, Аркадий Борисов.Исследование возможностей искусственных нейронных сетей в распознавании печатных слов
Значение систем анализа и распознавания документов сложно переоценить;подобные системы находят применение в различных областях человеческой деятельности. Вданнойстатьеприведенобзорпостандартной структуресистемы анализаираспознаваниядокументов, а также-возможностейиспользованияискусственныхнейронныхсистемвних.Была поставлена задача по сравнениювозможностейискусственныхнейронныхсетейтипамногослойныйперсептронсобратнымраспространениемошибки(MLP), обучаемой с учителем, икартысамоорганизующихсяпризнаковКохонена (SOM) с обучением без учителя для решения задачи распознавания и классификации заглавных печатных символов латинского алфавита, напечатанных различными шрифтами. Экспериментыбылисоставленытакимобразом,чтобывыявитьвлияние различных параметров работы сетей на способность классификации печатных букв из тестового набора данных.СетьMLPтестироваласьсоследующимиизменяемымипараметрами: размервходногоизображения 18x18, 16x16 и 14x14 пикселей, размер обучающей выборки - 8 и 5 шрифтов, размер скрытого слоя – 50, 75 и 100 нейронов, а также количество эпох обучения – 1500 и 3000. Каждыйиз 36 экспериментовпроводился 20 раз, наилучшиерезультатыпоказаласетьс50 нейронами в скрытом слое, обучаемая на 8 шрифтах в течении 3000 эпох, и размером входного слоя 14x14 нейронов – средний процент распознавания символов по тестовой выборке составил 75%. ДлясетиSOMкроме размера входного слоя и количества шрифтов в обучающей выборке менялось количество нейронов в слое Кохонена – 30 и 60 нейронов. Здесьнаилучшийсреднийрезультатраспознаванияв37% был достигнут при размере слоя Кохонена 60 нейронов, 8 обучающих шрифтах и размере входного изображения 18x18 пикселей. Таким образом, показана важность больших входного слоя и слоя Кохонена для успешного распознавания (кластеризации) символов, а также использования меньшего входного слоя для сети MLP. В ходе сравнения показано преимущество сети MLPнад сетью SOM в рамках поставленной задачи. Необходимо отметить, что рассмотренные нейронные сети не используют принципов рецепторного поля, а также не являются иерархическими сетями, таким образом, они не в состоянии самостоятельно извлекать абстрактные параметры входного изображения. Повышение возможностей распознавания возможно с использованиемкомбинации сетей SOMи MLP, использованием алгоритмов предобработки изображения для выявления инвариантных признаков символов, а также с применением более продвинутых иерархических архитектур нейронных сетей.