Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 100 (2015) 782 – 790
Procedia Engineering
www.elsevier.com/locate/procedia

25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014

Classification of Test Documents Based on Handwritten Student ID's Characteristics

Stipe Celar[a], Zeljko Stojkic[b], Zeljko Seremet[b,]*, Zeljko Marusic[c], Danijel Zelenika[b]

[a]*Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, R. Boskovica 32, Split, 21000, Croatia*
[b]*Faculty of Mechanical Engineering and Computing, Matice hrvatske bb, Mostar, 88000, Bosnia and Herzegovina*
[c]*Faculty of Science and Education, Matice hrvatske bb, Mostar, 88000, Bosnia and Herzegovina*

Abstract

The bag of words (BoW) model is an efficient image representation technique for image categorization and annotation tasks. Building good feature vocabularies from automatically extracted image feature vectors produces discriminative feature words, which can improve the accuracy of image categorization tasks. In this paper we use feature vocabularies based biometric characteristic for identification on student ID and classification of students' papers and various exam documents used at the University of Mostar. We demonstrated an experiment in which we used OpenCV as an image processing tool and tool for feature extraction. As regards to classification method, we used Neural Network for Recognition of Handwritten Digits (student ID). We tested out proposed method on MNIST test database and achieved recognition rate of 94,76% accuracy. The model is tested on digits which are extracted from the handwritten student exams and the accuracy of 82% is achieved (92% correctly classified digits).

Peer-review under responsibility of DAAAM International Vienna

Keywords: Bag of Words; visual vocabularie; image categorization; Neural Network; Recognition of Handwritten Digits

1. Introduction

Computer vision (CV) technology has rapidly developed with the increase of digitized contents and materials. Generic object recognition and image categorization are especially important themes in CV [1]. Although specific object recognition identifies strictly defined targets, generic object recognition categorizes visual object classes and

* Corresponding author. Tel.: +387 63 811649; fax: +387 36 337012.
E-mail address: seremet.zeljko@gmail.com

Peer-review under responsibility of DAAAM International Vienna
doi:10.1016/j.proeng.2015.01.432

labels them. Thus, generic object recognition needs grouping across semantic categories in general scenes and/or objects, requiring a wide-ranging categorization of images. However, it is hard to extract and define image features in all objects.

Handwritten digit recognition is an active topic in OCR (Optical character recognition) applications and pattern classification/learning research. In OCR applications, digit recognition is dealt with in postal mail sorting, bank check processing, form data entry, etc. For these applications, the performance (accuracy and speed) of digit recognition is crucial to the overall performance. In pattern classification and machine learning communities, the problem of handwritten digit recognition is a good example for testing the classification performance. [2] The performance of character recognition largely depends on the feature extraction approach and the classification/learning scheme. For feature extraction of character recognition, various approaches have been proposed [3].

In the teaching process, teaching staff (professors, assistants) must organize and manage a lot of documentation. Test lists, quizzes, written assignments, tests and attendance lists are just some of them. When we bear in mind that these materials must be archived and stored for some time, it is clear that it is necessary to digitize. This heap of documentation can then be used for various analyses, as well as for the formation of a knowledge base.

Students usually use their signature or index number to identify some of these documents. Handwritten signatures or handwritten index number represent non-invasive and inexpensive identification methods. In this work we have used student ID as a method of identifying the test/exam documents.

The analytical methods to express image features and to categorize visual images have achieved recognizing generic objects with high accuracy. For example, the Speeded-Up Robust Features (SURF) and Scale-Invariant Feature Transform (SIFT) can efficiently explore local information of the object boundary [4]; they have been applied for feature detection in target objects. The bag of features (BoF) scheme has also been demonstrated as an effective style of image feature distribution [1]. To detect individually handwritten digits with high accuracy, the technique for image feature detection may be combined with machines learning method such as neural networks (NN) [5], [6], [7], [8].

In this paper Structural Characteristics with SIFT generate visual vocabulary and it is trained by Multi-Layer Perceptron (MLP) neural networks. Their outputs are then combined to give a more accurate output. Several combination schemes have been trained on the well-known MNIST handwritten digits database [9]. The results show that all combination schemes greatly improved the recognition performance when compared to a single feature extraction-classifier pair alone. A combination module in WEKA using MLP network as combiner is proposed, achieving a recognition rate of 94,76% on the MNIST database, which is the highest recognition rate published for this database to date. The main contribution of this paper is the network accuracy of 82% in classifying of handwritten student exams (92% correctly classified digits).

This paper is divided into four sections. Section 2 describes feature extraction employed for this research. Section 3 discusses the image classification based neural network and section 4 describes the results obtained and puts forward a discussion and analysis. Finally, conclusions and future work are also given at the end of the paper.

2. Feature extraction

Features are key process in achieving high accuracy in identifying signatures. The ideal separation techniques use minimal sets of features that are used to maximize the difference between the digits, while minimizing the difference of the same digits [10].

We used two types of features. Structural Characteristics that we can consider as global features and local descriptors using Scale-Invariant Feature Transform (SIFT) algorithm. The second type of features we gathered with BoW model. The methods are described below.

2.1. Structural Characteristics

This algorithm consists in extracting histograms and profiles, combining them in a single feature vector. The input image is scaled in a 28x28 matrix. Horizontal and vertical histograms are computed by the number of black pixels in each line and column, respectively. Details can be found in [11], [12].

2.2. SIFT features

In content based image classification, the principle basis is the content of the image. Classification results are given based on the similarity of image contents, and image contents are described via image features. The extraction of visual features is the first step to image classification and the basis of image content analysis. It exists in all processing procedures in image analysis and directly influences the ability of describing images. Therefore, it makes a huge difference to the quality of further analysis and the effectiveness of application systems.

SIFT operator is an image local feature descriptor explicated by David G. Lowe in 2004. It is one of the most popular local features, based on scale space and invariant to scaling, rotation and even affine transformation. Firstly, SIFT algorithm detects features in the scale space and confirms the location and scale of key points. Then, it sets the direction of gradient as the direction of the point. Thus the scale and direction invariance of the operator are realized. SIFT is a local feature, which is invariant to rotation, scaling and change of light and stable in a certain extent of changes in visual angle, affine transformation and noise. It ensures specificity and abundance, so it is applicable to fast and accurate matching among mass feature data. Its large quantity ensures that even a few objects can generate a number of SIFT features, high-speed satisfies the requirement of real-time, and extensibility makes it easy to combine with other feature vectors.

For an image, the general algorithm to calculate it's SIFT feature vector has four steps:

(1) The detection of extreme values in scale space to tentatively determine the locations and scales of key points. During this process, the candidate pixel needs to be compared with 26 pixels, namely 8 neighbouring pixels in the same scale and 9×2 neighbouring pixels around the corresponding position of adjacent scales.

(2) Accurately determine the locations and scales of key points via fitting three dimensional quadratic functions, meanwhile deleting the low-contrast key points and unstable skirt response points (for DOG algorithm will generate strong skirt responses).

(3) Set the direction parameters for each key point via the direction of gradient of its neighbouring pixels to ensure the rotation invariance of the operator. Actually, the algorithm samples in the window centred at the key point and calculates the direction of gradient in the neighbouring area via histogram. A key point may be assigned to several directions (one principal and more than one auxiliary), which can increase the robustness of matching. Up to now, the detection of key points is completed. Each key point has three parameters: location, scale and direction. Thus an SIFT feature region can be determined.

(4) Generation of SIFT feature vector. First of all, rotate the axis to the direction of key point to ensure rotation invariance. In actual calculation, Lowe suggests to describe each key point using 4×4 seed points to increase the stability of matching. Thus, 128 data points, i.e. a 128-dimensional SIFT vector are generated for one key point. Now SIFT vector is free from the influence of geometric transformations such as scale changes and rotation. Normalize the length of the feature vector and the influence of light is eliminated. [13]

2.3. Bag of Words model

With the widely application of local features in computer vision, more attention is placed on methods of local feature based image classification. When extracting local features, the number of key point varies in different images, so machine learning is infeasible. To overcome these difficulties, researchers such as Li-Fei Fei from Stanford University were the first to phase Bag of Words model into computer image process as a sort of features [14]. Using Bag of Words model in image classification not only solves the problem brought by the disunity of local features, but also brings the advantages of easy expression. Now the method is extensively used in image classification and retrieval [15].

The main steps are as following:

(1) Detect key points though image division or random sampling etc.

(2) Extract the local features (SIFT) of the image and generate the descriptor.

(3) Cluster these feature related descriptor (usually via K-means) and generate visual vocabulary, in which each clustering centre is a visual word.

(4) Summarize the frequency of each visual word in a histogram.

Images are presented only by the frequency of visual words, which avoids complicated calculation during matching of image local features and shows obvious superiority in image classification with a large number of

classes and requiring a lot of training. Despite the effectiveness of image classification based on Bag of Words model, the accuracy of visual vocabulary directly influences the precision of classification and the size of vocabulary (i.e. the number of clusters) can only be adjusted empirically by experiments. In addition, Bag of Words model leaves out spatial relations of local features and loses some important information, which causes the incompleteness of visual vocabulary and poor results. [2], [13]

3. Image classification based on NN

Handwritten digit identification is a pattern recognition problem. Typical pattern recognition system has the following steps: Data Acquisition, Preprocessing, Feature Extraction, Classification and Performance Evaluation. [16]

In our work, in Feature Extraction step, images are presented by Structural Characteristics and Scale-Invariant Feature Transformation combined with Bag of Words model. Here BoW word library is a visual word library constructed on the basis of SIFT. Related library is trained according to each image semantics to get the proper description. Next, the results are given by the classifier in this article, which is generated from the combination of different number node of hidden layer. This method can commendably extract the spatial information contained in the semantics from the features and optimize the parameter selection in NN [17], [18], [19]. The experiments are tested on MNIST handwritten digits dataset and include the comparisons on operational speed, size of Bag of Words and number of hidden layers. The final results show that the classification method based on one hidden layer is effective in image classification and performs better than present algorithms of the same kind. [15]

3.1. Data acquisition

We collected a variety of tests, exam papers of students with whom we done our classification. We scanned documents with flat based scanner with resolution of 300 dpi. After that we extracted the location of student ID segment on document with series of preprocessing operations which are explained in next chapter. Also some steps and operations are dependent on training database that we used (MNIST). Although we could build our own digit database from student information, we decided to use more general approach and therefore we used independent MNIST database of handwritten digits. The MNIST database of handwritten digits has a training set of 60.000 examples and a test set of 10.000 examples.

3.2. Preprocessing

Preprocessing is a set of subsequent operations applied for the improvement of quality of signature image and locating ROI (Region of Interest) of signature [20], [21].

The various sub-processes which can be considered as image preprocessing are: Cropping, Binarization, Noise removal, Morphological Operations, Image normalization, Skeletonization [20] and similar.

3.2.1. Binarization and image denoising

Binarization represents an initial step in most systems for the analysis of document images. It refers to conversion of the gray-scale image into a binary image. It is crucial because it is the basis for successful segmentation and recognition of characters, words or zones. Generally, we distinguish global and local methods. Global methods use one threshold value to classify the pixels in the image object or to classify the background. On the other hand local schemes can use multiple values based on the information in the local area. We used global technique Otzu algorithm to perform binarization. After image acquisition and thresholding there can be significant amount of noise like salt and pepper noise. We used the median filter that resolves the noise of salt and pepper with a very small degree of image degradation.

3.2.2. Data area cropping and rotation

At the top of each test sheet there is a header containing basic information about the exam. This information consists of the name of the exam, the exam type, number of exam period, the date of the exam, the student's signature and the index number or student ID. If we are dealing with documents we must first segment the header (Fig. 1), and then in the header we locate the region of interest, the student ID. (Fig. 2) For this purpose, the

algorithm used is based on the topological structural analysis of the image that has been implemented in OpenCV [22]

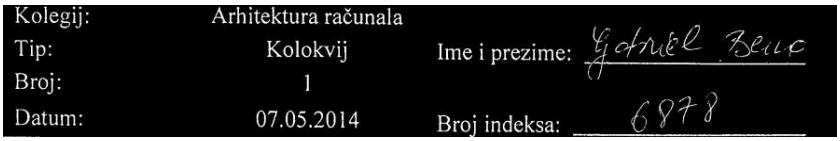

Fig. 1. Header document after median filter in negative.

In Fig. 2 we can see that base line is extracted with student ID. This line is of great importance to us; we can very easily determine and correct the angle of image acquisition to zero with rotation operation and algorithm in [22]. In this way we can improve accuracy of feature extraction.

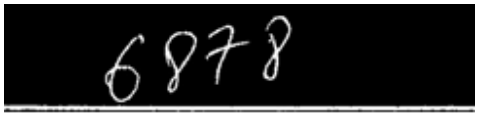

Fig. 2. Student ID with baseline.

3.2.3. Morphological operations

After binarization, cropping and rotation, we continue processing resulting images in order to improve the quality of text regions and preserve connectivity within certain characters, and meet some possible gaps, holes within the character. For this purpose, mathematical morphology operations such as dilation and erosion pictures were used. These operations are usually applied to binary images. Also, in this process we remove the base line with dilatation and bitwise operation.

Then after cropping we found bounding box of student ID like in Fig. 3

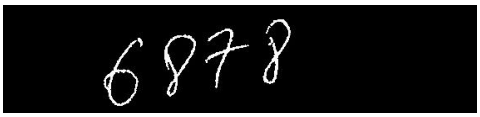

Fig. 3. Student ID without baseline.

3.2.4. Digit segmentation and normalization

The crucial part for recognizing student ID is proper segmentation of digits/characters. Generally handwritten character segmentation is a very difficult problem because there is a large variation in handwriting styles. For example, people may write the same character in different ways, including in different shapes and sizes. As a result, it is usually difficult to ascertain the number of characters in a handwritten text to be segmented. Furthermore, the various ways in which two neighbouring characters could be connected make it very difficult to determine the boundary that separates neighbouring characters by evaluating only local stroke geometries. [23]

We simplified our approach and used exams in which digits are not mutual connected and there is significant gap between digits. We scanned, top-down and left–right, to detect all the connected components of the image. The 8-adjacency was considered in this work. Small components and noise are discarded at this time. The bounding box of the remaining components is then extracted.

Earlier we mentioned that some preprocessed steps are dependent on MNIST database. Every digit in this database is normalized to fit in a 20x20 pixel box while preserving aspect ratio. The resulting images contain grey levels as a result of the anti-aliasing technique used by the normalization algorithm. The images were centred in a 28x28 image by computing the centre of mass of the pixels, and translating the image so as to position this point at the centre of the 28x28 field. Because we used this database as a training set, for the better results of classification, characteristics of our extracted digits are matched with those in database.

Fig. 4. The normalized digits of student ID.

3.3. Feature extraction and organization – image classification based on text classification analogy

In this paper, a novel recognition system by using combination of two feature extraction methods and Bag of Words model is proposed. In traditional SIFT algorithm, massive features will be extracted from each image after key point detection in Gaussian feature space. The organization of these features is critical for the following procedures such as machine learning and classification. Bag of Words model at first appeared in text detection and has achieved great success in text processing. Probabilistic Latent Semantic Analysis model mines the concentrated theme from the text via non-supervise methods, i.e. it can extracts semantic features from bottom. Bag of Words model neglects the connections and relative positions of features. Although this results in the loss of some information, but it makes model construction convenient and fast. Traditional neighbourhood feature extraction techniques in images and videos mainly focuses on the global distributions of colours, textures, etc. from the bottom layer, such as colour histogram and Gabor Filter. For a specific object, always only one feature vector is generated, and Bag of Words model is not necessary in such application. However, recent works have showed that global features alone cannot reflect some detailed features of images or videos. So more and more researchers have proposed kinds of local features, such as SIFT [1], [24], [25]. These feature descriptors of key points are effective in local region matching, but when applied to global classification, the weak coupled features of each key point cannot effectively represent the entire image or video. Therefore researchers have phased Bag of Words model from text classification into image description. The analysis of the relation between text classification and image classification is helpful to adapt all kinds of mature methods from the former to the latter. [26], [27], [28]

Comparing text classification with image classification, we assume that an image contains several visual words, similar to a text containing several text words. The values of key points in an image contain abundant local information. A visual word is similar to a word in text detection. These features are clustered into groups so that the difference between two groups is obvious, and the centre of each group is a visual word. In other images, specific vector images and groups of extracted local features are generated on the basis of certain groups of words. Such descriptive method is suitable to work with linear classifiers such as NN. [15]

3.4. Image classification algorithm in this article

In this section, we will introduce the overall method of image classification system. In this method, we extract feature with Structural Characteristics and Scale-Invariant Feature Transformation from an image, organize it via BoW method and obtain the final blocked histogram descriptor. During the training stage, Multilayer perceptrons are employed for classification.

The procedure of the algorithm is:

1. Extract with Structural Characteristics and Scale-Invariant Feature Transformation.
2. Obtain the vocabulary training.
3. Process the former features via Multi-Layer Perceptrons.
4. Use Multi-Layer Perceptrons as a classifier, optimizme activation function parameters and obtain the final classifier.

In this method, the first step is the features extracting. This algorithm consists in extracting histograms and profiles, combining them in a single feature vector. The input image is scaled in a 28x28 matrix. Horizontal and vertical histograms are computed by the number of black pixels in each line and column respectively.

A 28x28 image is thinned and scaled into a 25x25 matrix. The Sobel operators are used to extract two distinct edge maps: horizontal and vertical. At each point of the image we calculate an approximation of the gradient in that point by combining horizontal and vertical results. The original image is divided into 25 sub-images of 5x5 pixels each. The features are obtained calculating the percentage of black pixels in each sub-image (25 features per image). We extract SIFT features from image and employ BoW model to classify them in 64 classes and summarize the frequency of each features in a histogram (64 features per image). These features form a 96-dimension (28+28+25+15) feature vector.

Compared to traditional SIFT feature, our extract features is more efficient. After the generation of the vocabulary in the experiments we used Multi-Layer Perceptrons. The parameters are undefined, so the calculation needs optimization. Optimizing the selected activation function and other parameters finally results in the optimal solution and corresponding Multi-Layer model. Up to now, the training process is completed. The feature extraction step is the same in testing process, and the same vocabulary is used in BoW feature summarization. [15]

4. Experiments and analysis

4.1. Classification on MNIST database

This section presents the experiments obtained by extended SIFT feature extraction method. In this experiment we used WEKA, the CPU is Intel(R) Core 2 Duo with dual cores of 2,20GHz, the Memory is 4,00GB and the OS is Windows 7 Ultimate. The first part of our experiment was conducted on the well-known MNIST database. This database contains a training set of 60.000 images and a test set of 10.000 images (Fig. 6). All digits are size-normalized and centered in a 28x28 image. The training set was divided in 50.000 patterns for training (5.000 images per digit) and 10.000 (1.000 per digit) for validation. The three layers MLP trained with Backpropagation (BP) [29] algorithm were used as classifier in all feature sets. The best configurations for each feature set is construct 96×24×10 BP network model as number recognition with learning rate 0,3, momentum 0,2 and 0,5 as initialized weight. We refer to abundance of training experiment and choice 0,0228 as error values and maximum 500 iterations.

After the time consumption of training is 3.785,39 seconds, we gave 94,76% correct classification. In relation to some published results [12], this model has improvement potential.

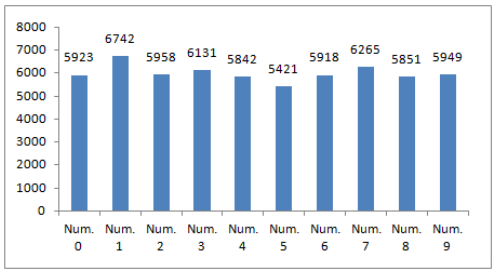

Fig. 5. Overview MNIST dataset for each digit.

4.2. Classification on students exams

In the second part of our experiment, we collected 50 exam documents from various students. Student IDs at our University consists exact of 4 digits. So the total amount of digits was 200. The distribution of digits is not equal (as shown in Fig. 7). The reason of uneven distribution is because we collected student exams of 2 groups of students that were on the second year of undergraduate study. As we can see on Fig. 7 the occurrence of number six (6) and number seven (7) are significantly larger than the other digits.

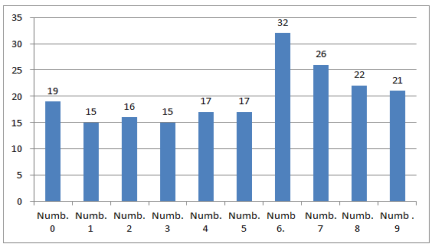

Fig. 6. Number of extracted digits from exam documents.

Confusion matrix or contingency table which visualizes performance of classification is shown in the Table 1. As we can see from the table, the greatest number of misclassifications was for number seven (7), number five (5) and number four (4), totally three errors.

Table 1. Confusions matrix.

Num.0	Num.1	Num.2	Num.3	Num.4	Num.5	Num.6	Num.7	Num.8	Num.9	
19	0	0	0	0	0	0	0	0	0	Num.0
0	**15**	0	0	0	0	0	0	0	0	Num.1
1	0	**14**	**1**	0	0	0	0	0	0	Num.2
0	0	**1**	**14**	0	0	0	0	0	0	Num.3
0	0	0	0	**14**	0	0	0	**1**	**2**	Num.4
0	0	**1**	0	0	**14**	0	0	**2**	0	Num.5
2	0	0	0	0	0	**30**	0	0	0	Num.6
0	0	**1**	0	**2**	0	0	**23**	0	0	Num.7
0	0	0	**2**	0	0	0	0	**20**	0	Num.8
0	0	0	0	0	0	0	0	0	**21**	Num.9

In the Table 2 we can see general performance of our model. Of 200 digits, the algorithm correctly classified 184 instances, and missed 16 instances. Overall performance for every digit was 92%.

Table 2. Result measured.

Evaluation criteria with MLP	Results
Correctly Classified Instances	184
Incorrectly Classified Instances	16
Kappa statistic	0,9105
Mean absolute error	0,0228
Root mean squared error	0,1214
Relative absolute error	1,69 %
Root relative squared error	40,46 %
Total number of instances	200
Accuracy	92 %

Although we achieved good performance on identifying separate digits, our goal was the classification of student exams. Earlier in the paper we mentioned that every exam consists of four digits and that we had 50 test exams. From the Table 2 we can see that 16 digits were misclassified. One document contained 3 misclassified digits, five documents contained for 2 misclassified digits, and three documents contained 1 error digit. Performance of *classification of student exams* was significantly lower with overall accuracy of 82%.

Conclusions

This paper takes the handwritten digits (student IDs) as its study objects. In this paper an image classification algorithm based on Bag of Words model and MLP learning is proposed. We preprocess digit and extract feature with Structural Characteristics and SIFT methods, and then make BP algorithm for the recognition in WEKA software. Our algorithm is a good basis for further research but its efficiency can be improved.

In future work:

1. We can try another kind of neutral network to classify, and modify the structure of network in order to design much better classification.
2. We can study more kinds of digit and combine with them, design multilevel neutral network of classifier in series with different feature and integrated identification system in parallel.
3. We can enlarge sample capacity, especially collecting more kinds of sample in order to perfect the system. Therefore, the system can be adopted in other domains dealing with such documents, such as school, hospital, etc.

References

[1] Csurka, C. R. Dance, L. Fan, J. Williamowski, C. Bray, Visual categorization with bags of keypoints, Proceedings of the IEEE Workshop on Statistical Learning in Computer Vision (SLCV'04), pp. 1-16, 2004.

[2] J. Sivic, A. Zisserman, Video google: a text retrieval approach to object matching in videos, in: International Conference on Computer Vision, 2003, pp. 1470–1477.

[3] Y. LeCun, et al., Comparison of learning algorithms for handwritten digit recognition, in: F. Fogelman Souli_e, P.Gallinari (Eds.), Proceedings of the International Conference on Artificial Neural Networks, Nanterre, France, 1995, pp. 53–60.

[4] O. D. Trier, A. K. Jain, T. Taxt, Feature extraction methods for character recognition—a survey, Pattern Recognition 29 (4) (1996) 641–662.

[5] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding, vol. 110, pp. 346-359, 2008.

[6] N. Cristianini, J. Shawe-Taylor, Introduction to support vector machines and other kernel-based learning methods, Cambridge University Press, Cambridge, 2000.

[7] T. Saric, G. Simunovic, K. Simunovic, Use of neural networks in prediction and simulation of steel surface roughness, International Jounal for Simulation Modelling, vol. 12(4), 2013, pp. 225-236

[8] T. Irgolic, F. Cus, M. Paulic, J. Balic, Prediction of Cutting Forces with Neural Network by Milling Functionally Graded Material, 24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013.

[9] MNIST database, http://yann.lecun.com/exdb/mnist/ (retreived at: 14.7.14.).

[10] L. Batista, D. Rivard, R. Sabourin, E. Granger, P. Maupin, State of the art in off-line signature verification, in: Verma B., Blumenstein M. (eds.), Pattern Recognition Technologies and Applications: Recent Advances, (1e). IGI Global, Hershey (2007).

[11] E. Kavallieratou, K. Sgarbas, N. Fakotakis, G. Kokkinakis, Handwritten word recognition based on structural characteristics and lexical support, International Conference on Document Analysis and Recognition, pp. 562–567, 2003.

[12] R. M. O. Cruz, G. D. C. Cavalcanti, I. R. Tsang, Handwritten Digit Recognition Using Multiple Feature Extraction Techniques and Classifier Ensemble, IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing.

[13] M. Rosenblum, Y. Yacoob, L. S. Davis, Human expression recognition from motion using a radial basis function network architecture, IEEE Transactions on Neural Networks, vol. 7, pp. 1121–1138, 1996.

[14] D. Lowe, Distinctive image features from scale-invariant key points, Proceedings of International journal of computer vision, vol. 60, no. 2, 2004., pp. 91-110.

[15] X. Lou, D. Huang, L. Fan, A. Xu, An Image Classification Algorithm Based on Bag of Visual Words and Multi-kernel Learning, Journal of Multimedia, Vol. 9, No. 2, February 2014, pp. 269-277.

[16] R. O. Duda and P. E. Hart, Pattern Classification (2e), Wiley India Private Limited. 2006.

[17] Simon Haykin, Neural networks a comprehensive foundation, Second Edition, Tsinghua University Press, 2001.10, pp. 159.

[18] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, 2nd Edition. Prentice Hall, 2002.

[19] C. Kaensar, Analysis on the Parameter of Back Propagation Algorithm with Three Weight Adjustment Structure for Hand Written Digit Recognition, 10th International Conference on Service Systems and Service Management (ICSSSM), pp. 18-22. 2013.

[20] A. McCabe, J. Trevathan, W. Read, Neural Network-based Handwritten Signature Verification, Journal of Computers, vol. 3, no. 8, August 2008.

[21] E. J. R. Justino, F. Bortolozzi, R. Sabourin, A comparison of SVM and HMM classifiers in the off-line signature verification, Pattern Recognition Letters, vol. 26, no. 9, pp. 1377–1385, 2005.

[22] S. Suzuki, K. Abe, Topological Structural Analysis of Digitized Binary Images by Border Following, CVGIP 30 1, 1985, pp .32-46

[23] D. Salvi, J. Zhou, J. Waggoner, S. Wang, Handwritten text segmentation using average longest path algorithm, Applications of Computer Vision (WACV), 2013 IEEE Workshop on, 15-17 Jan. 2013, pp. 505 – 512.

[24] L. Rothacker, S. Vajda, G. A. Fink, Bag-of-features representations for offline handwriting recognition applied to Arabic script, in: 2012 International Conference on Frontiers in Handwriting Recognition (ICFHR), IEEE, 2012, pp. 149–154.

[25] R. Shekhar, C. Jawahar, Word image retrieval using bag of visual words, in: 2012 10th IAPR International Workshop on Document Analysis Systems (DAS), 2012, pp. 297–301.

[26] A. Bolovinou, I. Pratikakis, S. Perantonis, Bag of spatio - visual words for context inference in scene classification, Pattern Recognition 46 (3) (2013) 1039 – 1053.

[27] M. Nilsback, A. Zisserman, A visual vocabulary for flower classification, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.2, 2006, pp. 1447 – 1454.

[28] T. Deselaers, L. Pimenidis, H. Ney, Bag-of-visual-words models for adult image classification and filtering, in: ICPR, 2008, pp.1–4.

[29] M. Riedmiller and H. Braun, A direct adaptive method for faster backpropagation learning: The RPROP algorithm, International Conference on Neural Networks, 1993, pp. 586–591.