New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume II: Applications

Editors

Krassimir T. Atanassov
Władysław Homenda
Olgierd Hryniewicz
Janusz Kacprzyk
Maciej Krawczak
Zbigniew Nahorski
Eulalia Szmidt
Sławomir Zadrożny

SRI PAS IBS PAN

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics Volume II: Applications

Systems Research Institute
Polish Academy of Sciences

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume II: Applications

Editors

Krassimir T. Atanassov
Władysław Homenda
Olgierd Hryniewicz
Janusz Kacprzyk
Maciej Krawczak
Zbigniew Nahorski
Eulalia Szmidt
Sławomir Zadrożny

IBS PAN SRI PAS

Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
www.ibspan.waw.pl

ISBN 83-894-7541-3

Dedicated to Professor Beloslav Riečan on his 75th anniversary

Complex classifiers in image recognition: implementation and application

Wojciech Lesinski[1] **and Wladyslaw Homenda**[1,2]
[1] Faculty of Mathematics and Computer Science
University of Bialystok,
ul. Sosnowa 64, 15-887 Bialystok, Poland

[2] Faculty of Mathematics and Information Science
Warsaw University of Technology
Plac Politechniki 1, 00-661 Warsaw, Poland
wlesinski@ii.uwb.edu.pl

Abstract

In some classification tasks using a single classifier does not give satisfying results. In this situation, you can use classifiers conjunction methods to improve the results. In this paper several known methods of classifiers conjunction, and their application in pattern recognition are presented. There are shown voting methods, bagging and random forest. The work also discusses how to create vector features for these classifiers.

Keywords: pattern recognition, classification, classifiers conjunction, ensemble learning.

1 Introduction

Recognition of images is one of the frequently studied matters of artificial intelligence. The problem of learning a computer to look at objects with human eyes for a long time will keep scientists busy. The computer which knows how to recognize images would be much easier to use. Object recognition facilitates the use of

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications *(K.T. Atanassow, W. Homenda, O. Hryniewicz, J. Kacprzyk, M. Krawczak, Z. Nahorski, E. Szmidt, S. Zadrożny, Eds.), IBS PAN - SRI PAS, Warsaw, 2012.*

modern IT tools for the elderly or disabled people, for whom the use of the mouse and keyboard causes some problems.

In many cases, the recognition process does not have sufficiently high efficiency. This problem is solved in different ways: by searching for more efficient classifiers, by improving features' vector or through the methods of classifiers conjunction. Recognition of graphic characters, such as digits and letters, greatly facilitates the digitization of all kinds of documents. Optical character recognition (OCR) simplifies the process of digitization, because it greatly accelerates the implementation to the computer systems old documents, which often exist only in a paper form.

Programs which recognize printed texts, but also the handwritten digits, are in common use. Their effectiveness is high, but there always could be a better solution. In this chapter, the letters recognition problem will be used to compare the effectiveness of the tested classifiers' conjunction methods.

2 Preliminaries

2.1 Classification problem

Classification [3, 7] is a task of splitting a set of objects into subsets including similar objects. Classification task is performed on observed features of objects rather then on objects directly. We assume that every object is characterized by a vector x with elements representing values of features. For simplicity, a vector of values of features will be simply called a vector of features. The space of vectors of features is denoted by X. A classification algorithm ψ assigns a class number $i \in M$ for every vector of features: In the recognition task the classification algorithm ψ assigns a class number $i \in M$ for every vector of features:

$$\psi \ : \ X \to M$$

Or, equivalently, it divides features' space onto so-called decision regions:

$$D_X^{(i)} = \{x \in X \ : \ \psi(x) = i\} \quad \text{for every} \quad i \in M$$

2.2 Classifiers conjunction

In the case of conjunction methods classifier is created with a number of other classifiers. Classifiers, which we use for connecting, we can call "weak classifiers".In different methods of combining we can use different types of classifiers, or the same. When we use the same classifiers they are differ way of learning. In

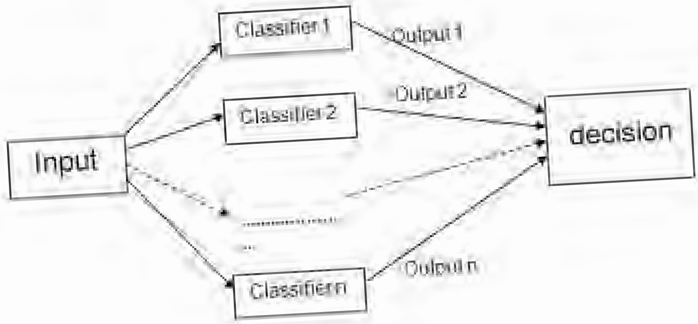

Figure 1: Scheme of classifiers conjunction

ensemble learning tested sample x is recognized by all used classifiers, next the results are compared to adjust only one system response. Usually, the process of joining results is nothing else, but sum of answers with set weights. The main difference is, that classifiers used in this system are trained in different ways.

2.3 Simple voting

Simple voting is one of the most simple conjunction method. We can use any classifiers components in this methods. Classifiers can be already trained, or stage of training can succeed during system creation. Also the way of training algorithm components is not imposed. The only condition of start-up of this algorithm is having trained classifiers, which are statistically independent from each other. The sample $x \in X$ is tested by every weak classifier, then an answer is counted as a sum. The class which is indicated by most of classifiers, is chosen as the right one.

2.4 Combined voting

Combined voting algorithm is a modification of the simple voting method. As in the previous method, the different classifiers are combine together. Conditions required the connected methods are the same as in the simple voting. The difference consists in adding weight to the votes of the weak classifiers. Object x is classified by each of the combined algorithms. The result is the sum of votes multiplied by their weight. The object is classified in this class for which the sum described above will be the highest. Of course, the classifier with better efficiency have higher weights. Application of this method, in contrast to a simple

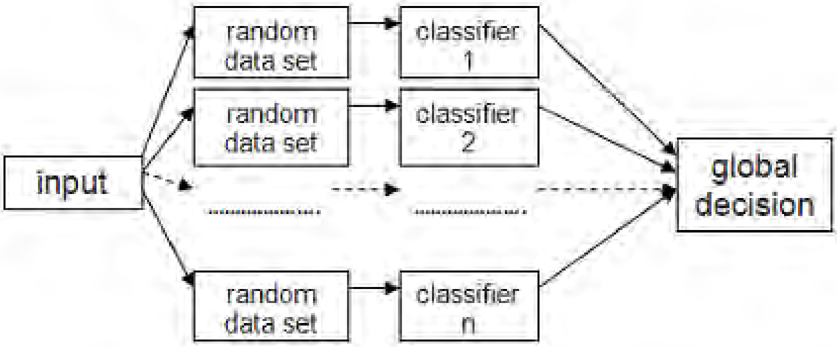

Figure 2: Scheme of bagging classifier

voting, requires knowledge about the effectiveness of the component classifiers. This knowledge we have from previous experiments or from literature.

2.5 Bagging

Bagging, a name derived from ”bootstrap aggregation”, devised by Breiman [1], is one of the most intuitive and simplest ensemble algorithm providing to a good performance. It uses multiple versions of a training set, each created by drawing $n = N$ (where N is a number of elements of original training set) samples from training set D with replacement. Each of these bootstrap data sets is used to train a different component classifier and the final classification decision is based on the vote of each component classifier. Traditionally the component classifiers are of the same general form - for example, all Hidden Markov models, or all neural networks, or all decisions trees - merely the final parameter values differ among them due to their different sets of training patterns. Scheme of bagging is on the figure 2

2.6 Random Forest

Random forest is a relatively new classifier proposed by Breiman in [2]. The method combines Breiman's [1] ”bagging” idea and the random selection of features in order to construct a collection of decision trees with controlled variation.

Random forest is composed of some number of decision trees. Each tree is built as follow:

- Let the number of training objects be N, and the number of features in features vector be M.

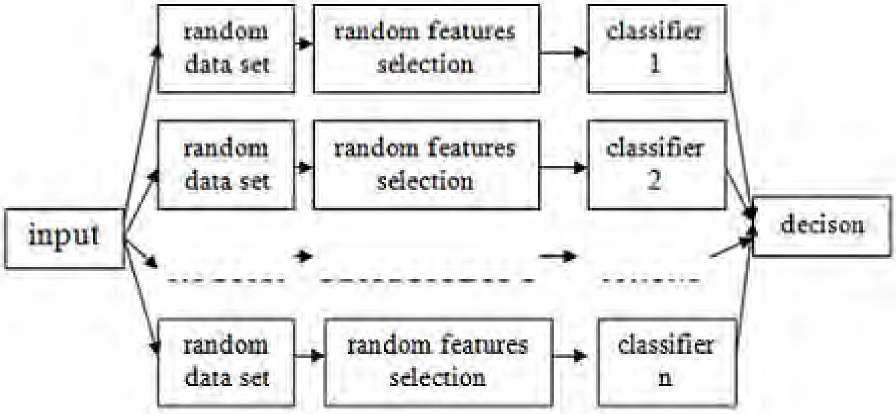

Figure 3: Scheme of random forest

- Training set for each tree is built by choosing N times with replacement from all N available training objects.
- Number $m << M$ is an amount of features on which to base the decision at that node. This features is randomly chosen for each node.
- Each tree is built to the largest extent possible. There is no pruning.

Each tree gives a classification, and we say the tree ”votes” for that class. The forest chooses the classification having the most votes (over all the trees in the forest). A random features selection used in the subsequent divisions of a single tree prevent over-fitting to training data. No pruning allows the use of the ID3 algorithm proposed by Quinlan in [9].

3 Feature extraction

Classification was done on features characterizing every symbol. Vectorized and numerical features characterizing symbols were defined based on experience of authors. The following features were used in the experiment [4, 5, 8]:

Projections Horizontal and vertical projections in a rectangle are taken Horizontal projection is defined for every row of pixels of the rectangle. For a given row, the value of the projection is equal to number of black pixels in this row. By analogy, vertical projection is defined in columns of the rectangle. For both projections, the maximum value, position of the maximum value, the average value and the support (number of nonzero values) are included in the feature's vector.

Transitions Like in case of projections, horizontal and vertical transitions in a rectangle are taken. Horizontal transitions are defined for every row of pixels of the rectangle. For a given row, the value of the transition is equal to number of pairs of consecutive white and black pixels in this row. By analogy, vertical transitions are defined in columns of the rectangle. Transitions reflect shape complexity of the image. For both transitions, the maximum value, position of the maximum value, the average value and the support (number of nonzero values) are included in the feature's vector. Only maximal values of transitions in both horizontal and vertical directions are included in the features' vector.

Margins Left, top, right and bottom margins in a rectangle are taken. Left margin is defined for every row of pixels of the rectangle. For a given row, the value of the left margin is equal to the number of white pixels from the left edge of the rectangle right to the first black one. The value of the right margin is equal to the number of white pixels from the right edge of the rectangle left to the first black one. Top ad bottom margins are defined analogously. These features show the symbol's position in the image. We used maximum value of all margins in features vector. Maximum values of all margins are included in the features' vector.

Moments are used in different fields, e.g. in physics (e.g. mass, center of mass, moment of inertia), in probability (e.g. mean value, variance). In image processing, computer vision and related fields, moment are certain particular weighted averages of the image pixels' intensities. Also, functions of moments are often utilized in order to have some attractive property or interpretation. Image moments are useful to describe objects after segmentation. Simple properties of the image which are found via image moments including area (or total intensity), its centroid and information about its orientation.

Directions for given pixel it is the longest segment of black pixels in given directions (usually horizontal, vertical, left and right diagonal directions are considered) which include given black pixel. Also directions of 22,5, 67,5, 112,5 and 157,5 degrees are considered too.

Derivatives Next group features are derivatives of vector features given above. We defined derivatives as the vector of differences between successive elements of the given vector. Derivatives of projections, transitions and margins are used in our features vector.

4 Experiment

Experimental results are based on the basis of lowercase Latin letters. The learning set included 1252 images in total. These images represent 26 classes corresponding to all lowercase characters from the Latin alphabet. Every class had different number of elements. The biggest class includes 102 images of ”a”. The smallest one has only 5 images of ”q.

The testing set includes 2550 images in total. They are dispersed in 26 classes of lowercase Latin letters. Like in the case of the learning set, classes have different numbers of images [5].

Symbols underwent preprocessing were cut of from scanned documents and then were subjected preprocessing before classification:

- gray-scale images were converted to monochromatic images using a threshold method: pixels having grayness level below that threshold were turned to black, while pixels with value above that threshold become white,
- symbols were transformed in order to fit the $X \times N$ square, where $N = 16$ or $N = 32$. Transformation was done in three steps:
  - images were turned to bounding boxes of symbols by cutting empty margins,
  - bounding boxes were scaled in order to bring longer size of the bonding box to the value N,
  - non-square bounding boxes were turned to squares by attaching white margins. Final bounding boxes of square shape have symbols centralized.

The experiment tested the effectiveness of recognition following classifiers [3, 6, 7]:

- simple method of voting in which were combined k nearest neighbor classifier, decision tree, naive Bayes classifier, classifier with Mahalanobis minimal distance and k-means algorithm
- combined voting
- bagging with knn and bagging with decision trees
- random forest with 10 and 25 trees

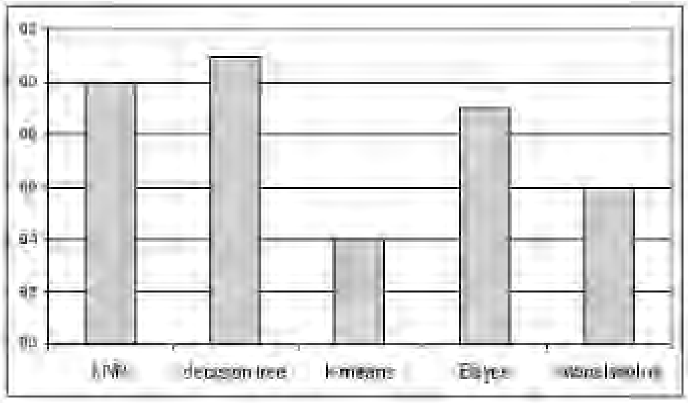

Figure 4: Efficiency of simple classifiers

5 Results

5.1 Simple classifiers

As the first simple classifiers, which were to be used in joining, was tested. These operations aim to determine effectiveness order to determine the weights of the different methods used in combined voting. In this group best result was obtain by k nearest neighbor classifier. It was 91%. Slightly weaker result - 90% - was achieved by decision tree. Naive Bayesian classifier had 89% efficiency. Next result was obtain by classifier with Mahalanobis minimal distance and it was 86%. The worst efficiency has reached k-means classifier. All results shown in Figure 4.

5.2 Classifiers conjunction

Simple voted method was test first. This algorithm uses simple classifiers described above. This algorithm has achieved the worst result among all methods of joining. Its effectiveness has reached 92%. Combined voting was next. Same classifiers as in simple voting were use. For KNN and decision tree was given a weight equal 2, Bayesian classifier 1, k-means and classifier with Mahalanobis minimal distance - 0.75. Weights was chosen on the basis of tests described in Section 5.1. This algorithm received slightly better result, which amounted to about 93%. The parameter k of kNN method used in both votes was 10, and the parameter k of k-means method was equal to 3. Both classifiers used the Euclidean distance. Next method for combining classifiers, which were tested, was

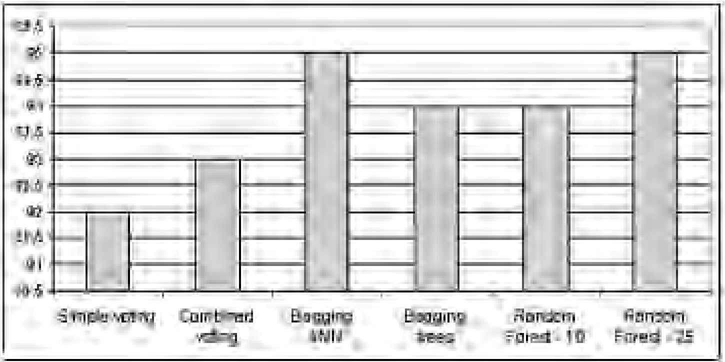

Figure 5: Efficiency of complex classifiers

bagging. For joining two the most effective simple classifiers: kNN and decision tree were use. This algorithm was more effective than voting methods. Bagging built with kNN as weak classifiers obtained equal to 95% efficiency, and created from decision trees with 94%. Unfortunately, bagging with KNN algorithm has shown a significant time-consuming. The last considered classifier was a random forest. Tests were carried out for the forest counting 10 and 25 trees. Forest built from 10 trees had efficiency equal to 94%, while 25 -95%. All results shown in Figure 5.

6 Conclusions

The most efficient and statistically independent in recognition process classifiers, were used in conjunction method. The best results from all methods mentioned in this paper, had obtained the random forest with 25 trees algorithm, presenting 97% efficiency.

Classifiers conjunction gives much better results than singular methods. When constructing those algorithms you have to pay attention to component classifiers. The component classifiers can not use complicated calculations and have to be high efficient.

The study confirmed the hypothesis about the combining classifiers effectiveness. The further works will focus on testing other combining methods and building a better features vector. The next stage will be the implementation of the described algorithms to recognize other object.

Acknowledgments

This work is supported by The National Center for Research and Development, Grant no N R02 0019 06/2009.

References

[1] Breiman L., (1996), Bagging predictors, in Machine Learning, 26(2), pp. 123-140

[2] Breiman L., (2001), Random Forests, in Machine Learning 45, pp. 5-32

[3] Duda R. O., Hart P. E., Stork D. G., (2001), Pattern Classification, John Wiley & Sons, Inc.,New York

[4] Homenda W., Lesinski W., (2009), Optical Music Recognition: Case of Pattern recognition with Undesirable and Garbage Symbols, in: Image Processing and Communications Challenges - Choras R. et al (Eds.), pp. 120-127, Exit, Warsaw

[5] Homenda W., Lesinski W., (2011), Features selection in character recogntion in random forest classifier, Lecture Notes In Artificial Intelligence, LNAI 6922, pp. 93-102, Springer

[6] Koronacki J., Cwik J., (2008), Statystyczne systemy uczace sie (in Polish), Exit, Warszawa

[7] Kurzynski M., (1997), Rozpoznawania obrazw metody statystyczne (in Polish), Oficyna Wydawnicza Politechniki Wroclawskiej, Wroclaw, 1997

[8] Malina W., Smiatacz M., (2005), Metody cyfrowego przetwarzania obrazow (in Polish), Exit, Warszawa

[9] Quinlan, J. R., (1986), Induction of Decision Trees. Machine Learning 1, 1 (Mar. 1986), pp. 81-106.

[10] http://www.stat.berkeley.edu/users/breiman/RandomForests/

The papers presented in this Volume 2 constitute a collection of contributions, both of a foundational and applied type, by both well-known experts and young researchers in various fields of broadly perceived intelligent systems.
It may be viewed as a result of fruitful discussions held during the Tenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2011) organized in Warsaw on September 30, 2011 by the Systems Research Institute, Polish Academy of Sciences, in Warsaw, Poland, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences in Sofia, Bulgaria, and WIT - Warsaw School of Information Technology in Warsaw, Poland, and co-organized by: the Matej Bel University, Banska Bystrica, Slovakia, Universidad Publica de Navarra, Pamplona, Spain, Universidade de Tras-Os-Montes e Alto Douro, Vila Real, Portugal, and the University of Westminster, Harrow, UK:

Http://www.ibspan.waw.pl/ifs2011

The consecutive International Workshops on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGNs) have been meant to provide a forum for the presentation of new results and for scientific discussion on new developments in foundations and applications of intuitionistic fuzzy sets and generalized nets pioneered by Professor Krassimir T. Atanassov. Other topics related to broadly perceived representation and processing of uncertain and imprecise information and intelligent systems have also been included. The Tenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2011) is a continuation of this undertaking, and provides many new ideas and results in the areas concerned.

We hope that a collection of main contributions presented at the Workshop, completed with many papers by leading experts who have not been able to participate, will provide a source of much needed information on recent trends in the topics considered.

ISBN-13 9788389475411