ESANN'2005 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 27-29 April 2005, d-side publi., ISBN 2-930307-05-6.

Artificial neural network fusion: Application to Arabic words recognition

Nadir Farah, Mohamed Tarek Khadir and Mokhtar Sellami *

Université Badji Mokhtar of Annaba, Département d'Informatique,
Bp 12, El Hadjar, 23000, Algeria

Abstract. The study of multiple classifier systems has become recently an area of intensive research in pattern recognition in order to improve the results of single classifiers. In this work, two types of features combination for handwritten Arabic literal words amount recognition, using neural network classifiers are discussed. Different parallel combination schemes are presented and their results compared with a single classifier benchmark using a complete feature set.

Key words: Handwritten Arabic recognition, structural and statistical features, MLP combinations.

1 Introduction

Several successful methods have been developed to recognize isolated handwritten characters and numerals. Nowadays the research is carried for handwritten word recognition [7, 9], which present a challenge due to the difficult nature of the unconstrained handwritten words, including the diversity of character patterns, ambiguity of characters, and the overlapping nature of many characters in a single word [1].

Handwriting recognition systems has been studied for decades and many methods have been developed [7]. Some use only the pixel-images as input to a statistical or neural classifier. Others preprocess the data in order to extract structural features that are fed into a classifier. The combination of different types of information has been shown to be promising in many pattern recognition systems [11, 5]. Different type of classifiers, different type of features, different type of combiner, etc may then be considered. In this paper, using two different family features and three neural network classifiers, Arabic words recognition is addressed. Features consists in two families: structural and the statistical.

*emails:farah;khadir;sellami@lri-annaba.net

The remainder of this paper is organized as follows: section 2 presents the characteristics of Arabic writing. In section 3 a brief overview of the system architecture is given. Section 4, describes the features extraction modules. The three individual classification systems are described in section 5 and their results in section 6. Combination approaches of classifiers are introduced in section 7 with their obtained results. The paper concludes with discussion on the obtained results.

2 Arabic writing characteristics

The Arabic language has a very rich vocabulary. More than 300 million people speak the language, and over 1 billion people use it in several religion-related activities. Arabic script is written from left to right.

احد	تسعة	ستون	اربعمائة	آلفا	ملياران
اثنان	عشر	سبعون	خمسمائة	الفان	ملايير
ثلاثة	عشرة	ثمانون	ستمائة	مليون	سنتيم
اربعة	اثنا	تسعون	سبعمائة	ملايين	و
خمسة	عشرون	مائة	ثمانمائة	مليونا	دينار
ستة	ثلاثون	مائتا	تسعمائة	مليونان	دنانير
سبعة	اربعون	مائتان	آلف	مليار	سنتيمات
ثمانية	خمسون	ثلاثمائة	الالف	مليارا	جزائري

Figure 1: *Bank draft lexicon of Arabic literal amounts*

As opposed to Latin which start from left to right. The Arabic alphabet consists of 28 characters. Ten of them have one dot, three have two dots, and two have three dots. Dots can be above or below. The shape of the character is context sensitive, depending on its location within a word. A letter can have up to four different shapes: isolated, beginning connection from the left, middle connection from the left and right, and end connection from the right. Most of the letters can be connected from both sides; the right and the left. However, there are six letters which can be connected from one side only; the right. This characteristic implies that each word may contain from one unit or more (sub-words). An example of Arabic words may be given in the lexicon, Figure 1, used in a literal amount of Arabic bank check. Some ligatures involve vertical stacking of characters, this characteristic complicates the segmentation problem [7], which is not considered in this work.

3 The global system architecture

The recognition system proposed is of modular architecture: feature extraction and word classification. Firstly, a preprocessing module, which binarises, smooth and extract features. These extracted features are transferred toward the MLP classifiers, Figure 2. The shape features are from two sets: statistical

ESANN'2005 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 27-29 April 2005, d-side publi., ISBN 2-930307-05-6.

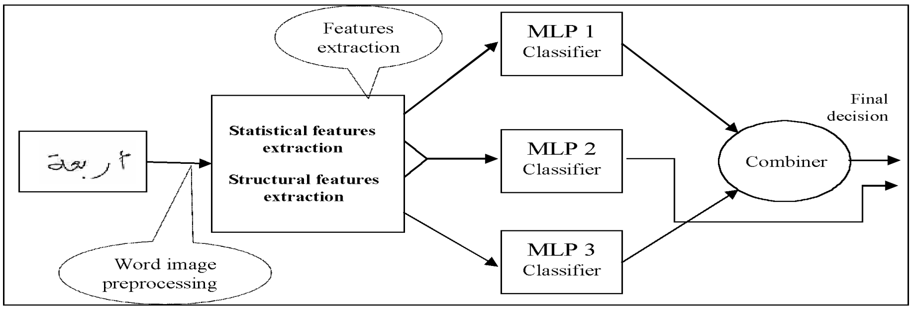


Figure 2: *Global system architecture*

features and structural ones. Each of these feature sets provides different information about the shape of a word. The first classifier receives the structural features, the second one uses structural and statistical features and the third one the statistical features only. The obtained results are then used by the combiner to produce a final decision.

4 Features Extraction

Feature extraction have been highly inspired by the human reading process that considers the global high level words shape [7, 9]. For holistic paradigm there is a wide range of methods to word recognition. They can be basically classified in two categories: statistical and structural. Theses features are automatically extracted, using different algorithms: contour extraction [8] and diacritical dots [10].

- The statistical feature set is pixel based information; it is expressed in terms of partitioning the word feature space as presented in figure 2. The features are the density of the lit pixels in various word image regions. The features are obtained from the zone-pattern regions showed in Figure 3(a).
- The structural feature is expressed as a composition of structural units, and a word is recognized by matching its structural representation with that of a reference words. The main concept in structural features extraction is to calculate the number of ascenders, descenders, loops, etc. Base line detection [8] is one of the most important information that permits us to situate diacritical dot position, and the main part of the word. The features extracted correspond to 9 structural ones, Figure 3(b) according to their possible occurrence numbers in the lexicon, Figure 1: 3 for ascenders, 2 for descenders, 2 for one dot above, 2 for two dots above, 2 for three dots above, 1 for one dot below, 2 for two dots below, 3 for loops, 4 for sub words.

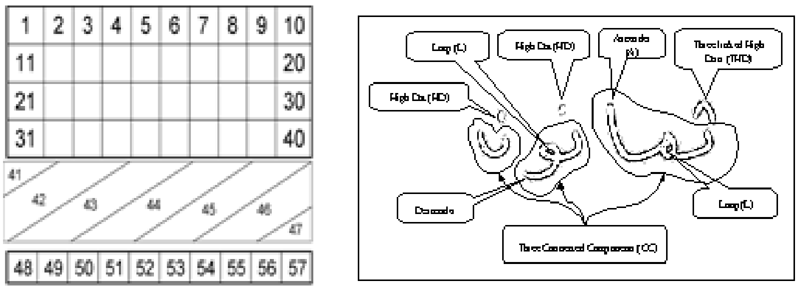
Finally, 57 statistical and 21 structural features are distinguished, Figure 3.

	Recognition rates	**Miss-classification**	**Rejection**	**Reliability**
Structural	87.83% (2108)	1.04% (25)	11.13 % (267)	99.00 %
Statistical	74.38% (1785)	0.79% (19)	24.83 % (596)	99.10 %
Statistical + Structural	89.17% (2140)	1.08% (26)	9.75 % (234)	99.00 %

Table 1: Recognition rates for structural, statistical and both family features

5 Classification stage

A three-layer Multi Layer Perceptron (MLP), with a sigmoid activation function has been used for all three modules Figure 2, which are trained using backpropagation algorithms [4]. The number of neurons contained in the hidden layer are calculated by a heuristic. The characteristics of each classifier are given separately in next sections. The three classifiers have been trained with 2400 words, they should then have the same view of the presented words, and will suggest the same or different word classes.



(a) Statistical features (b) Structural features

Figure 3: Feature extraction for Arabic words

6 Classification results

In this study, MLP classifiers have been used, the obtained results after the classification stage on a test set are summarized in Table 1.

The reliability is defined by: recognition /(100%-rejection) [11].

For recognition needs, 4800 word images were used. This set represents the 48 words of our lexicon, Figure 1, written by 100 different writers. Among these word images, a set 2400 were used to train the classifiers. The rejection criterion is chosen to keep the reliability of at least 99%. The recognition rate using the structural feature set, is superior by 13.40% to the one using the statistical feature set. However, still sensibly lower than the recognition rate of the classifier applied to the complete feature set.

	Recognition	**Misclassification**	**Reject**	**Reliability**
Borda Count	91.70%	0.86%	7.44%	99.10%
Dempster-Shafer	94.87%	0.97%	4.16%	99.00%
Product	93.90%	0.96%	5.14%	99.00%
Sum	94.93%	0.97%	4.10%	99.00%
Average	93.30%	0.96%	5.74%	99.00%
Max	92.12%	0.96%	6.92%	99.00%
Min	93.20%	0.96%	5.84%	99.00%
Naïve Bayes	93.50%	0.96%	5.54%	99.00%

Table 2: Recognition rates using different statistical combination schemes

Structural features set have stronger discrimination and provides better recognition rates than statistical ones Table 2.

7 Combination

Statistical combination methods are built around two MLP classifiers, performing classification separately on structural and statistical feature. Combination is done on the first and the third MLPs, Figure 2 described in Section 3. The MLP using both feature sets, is used as a comparison benchmark.

The first six combination schemes: product, average, maximum, minimum, sum [2] and Dempster-Shafer's evidence theory [11] of the corresponding pairs of the classifier outputs are used to make final decision.

The naive Bayes scheme uses the confusion matrices of member classifiers to estimate the certainty of the classifier decision [11]. The Borda count combination is a generalization of the majority vote [3]. The combination methods except Borda count, assume a unique interpretation of the confidence values, for instance as a posteriori probabilities. This is not the case, due to the specific characteristics of the individual classifiers and their different training sets. For this reason a normalization given in [6] is used, which permits to have a normalized scale of the output neurons activation.

For neural network, each node in the output layer is associated to one class and its output O_i, with [zero to one] range, reflects the response of the network to the corresponding class w_i. To facilitate the combination the responses are normalized and used as estimates of the a posteriori probability of each class [6]:

$$P(w_i|x) = \frac{O_i}{\sum_k O_k} \tag{1}$$

In our experiments we used different combination schemes, each classifier yields as output the 48 words of our lexicon with their confidence values $P(w_i|x)$.

Combination results on the 2400 test words are shown in Table 2. The normalized outputs of the two MLP were used as output confidences. The best recognition rates (superior to 93.00%) are obtained by six combination schemes. These results are about 4% better than the recognition rate of the MLP using both feature sets.

8 Conclusion

The combination of two different feature types has been presented in this paper, producing excellent results. The main contribution of this paper is the use of different statistical combination schemes for Arabic word recognition. In our paper, our experimental results show that the combination scheme of single classifiers outperforms classifier using both family features. For this particular study and using this particular set of words, we showed that it is more reliable to investigate simpler classifiers and combine them, instead of using a complex one. Therefore, the problem of curse dimensionality [5] may be avoided. Further investigations may rely on features extraction process and other types of combining schemes using larger word images number.

References

[1] Blumenstein M., Verma, B.K. *Neural-based solutions for the segmentation and the recognition of diffucult handwritten words from a benchmark database*, In the 5th International Conference on Document Analysis and Recognition, Bengalore, India, pp.281-284, 1999.

[2] Duin R.P.W., *The combining classifier: To train or Not to train?*, Proceedings 16th International Conference on Pattern Recognition, Quebec City, Canada, August 11-15, 2002.

[3] Ho T. K., Hull J.J., Srihari S.N., Decision combination in multiple classifier systems, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 16, No.1, pp.66-75, January 1994.

[4] Jain A. K., Mao J., Mohiuddin K., Artificial Neural Networks: A Tutorial', *IEEE Computer, Special Issue on Neural computing*, March 1996.

[5] Jain A. K., Duin R.P.W, Mao J., Statistical pattern recognition: A review, *IEEE transactions on pattern analysis and machine intelligence*, vol 22, no 1, January 2000

[6] Kittler J., Hatef M., Duin R.P.W., Matas J., On combining classifiers, *IEEE transactions on pattern analysis and machine intelligence*, vol 20, no 3, March 1998.

[7] Madhvanath S., Govindaraju V., The Role of Holistic Paradigms in Handwritten word Recognition, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 23 no.2, February 2001.

[8] Pavlidis T., *Algorithms for Graphic and Image Processing*, Rockville, MD: Computer science press, 1982.

[9] Steinherz T., Rivlin E., Intrator N., Off-line cursive script word recognition: A survey, *International Journal on Document analysis and Recognition, IJDAR*, Vol 2, pp: 90-110, 1999.

[10] Ameur A, Romeo-Pakker K, Miled H, Cheriet M., *Approche globale pour la reconnaissance de mots manuscrits Arabes*. In Proc. 3 ème Colloque National sur lEcrit et le Document, pp: 151-156, 1994.

[11] Xu L., Krzyzak A., Suen C. Y., Methods of combining multiple classifiers and their applications to handwriting recognition, *IEEE Transactions on systems. Man, and cybernetics*, vol. 22, No 3, May/June 1992.