Recognition of handwritten Amazigh characters based on zoning methods and MLP

Nabil AHARRANE, Karim EL MOUTAOUAKIL, Khalid SATORI
Computer sciences, Imaging and Digital analysis Laboratory members
Sciences Faculty of Dhar Mahraz
University Sidi Mohamed Ben AbedAllah
MOROCCO
aharranenabil@gmail.com karimmoutaouakil@yahoo.fr khalidsatori@gmail.com

Abstract: - The main purpose of this work is to develop an optical character recognition system (OCR) of handwritten Amazigh characters employing a feature set of 79 elements based on statistical methods.
The feature set elaborated consists of 37 densities features and 42 shadow features basing on a specific zoning to represent the Amazigh characters; in the recognition phase, we use the multilayer perceptron (MLP) as classifier. The accuracy observed, experimentally, on a database of 24180 characters is 96,47%.
The experimental evaluation performed on a large set of handwritten characters not only verifies that the proposed approach provides a very satisfactory recognition rate but also shows a reasonable time during the test phase.

Key-Words: - OCR, MLP, Handwritten Amazigh charcters, Statistical methods, Features extraction, Neural network.

1 Introduction

In recent years, the recognition of characters handwritten remains one of the most popular problems due to its diverse applications such as address classification system, processing of bank check, indexing archives, documents analysis, etc. Therefore, much work has been achieved for many languages, an overview of the latest works in Optical Character Recognition (OCR) research can be found in [1].

Recently, researchers have begun to give attention to the Amazigh language OCR. In this context, various methods have been used based on: Hidden Markov Models (HMM) [2], Hough transformation [3], neural approaches [4], geometrical and statistical methods [5][6][7], syntactical method rests on finite automata [8], moments features [9] and some hybrid methods [10][11][12].

In this paper, we propose an OCR system based on a statistical approach with a new feature set. This latter creates, for each character, a set of features rests on decomposing the image under study in term of zones, and then we extract a vector of 79 components which are the densities features and the shadow features. After features extraction, in order to its performance and its simple principle, we used the MLP with one hidden layer for recognition phase. The adopted system is as follow:

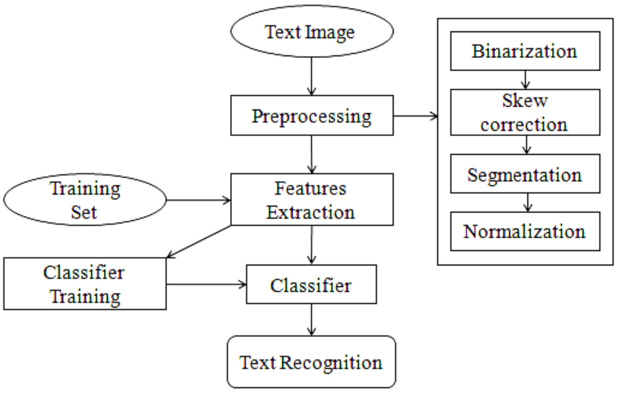


Fig.1: The adopted system

The rest of this paper is organized as follows: In Section 2, we present a description of the Amazigh language. The preprocessing description is given in Section 3 where we delineate all Necessary steps to prepare the image to the next phases. Feature extraction state of art is presented in Section 4. Section 5 exposes the MLP architecture of this work. Section 6 details our procedure to construct the feature set. In Section 7, we present some experimental results to evaluate our work. Finally, we conclude the paper with Section 8.

2 The Amazigh Language

The Amazighs are the indigenous people of North Africa, with their own language, culture and history. They are one of the most ancient peoples of humanity. Their presence in Tamazgha (North Africa) was more than 12000 years. The Amazigh language has existed since the earliest antiquity. It has an original writing system, Tifinagh, used and preserved to this day. In recent decades, all Amazigh groups have reclaimed this ancestral writing. Currently, the Amazigh language is spoken by about 30 million speakers in North Africa (from the oasis of Siwa in Egypt, to Morocco passing through Libya, Tunisia, Algeria, Niger, Mali, Burkina Faso and Mauritania).

In Morocco, where nearly 50% of people are amazigh, the Amazigh language is divided into three regional varieties with Tarifite in North, Tamazight in Central Morocco and South-East and Tachelhite in South-West and the High Atlas [13].

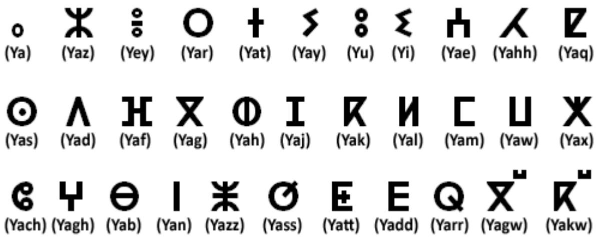


Fig.2: Tifinagh characters adopted by the IRCAM.

The official introduction of the Amazigh language teaching in the Moroccan educational system in 2003 involves the selection of a standard common language to teach. This task was accomplished by “Royal Institute of the Amazigh Culture” (IRCAM) created in July 2001 [14]. Actually, the Tifinagh-IRCAM alphabet is based on 33 characters (Fig 2). In the amazigh OCR field, one works only on 31 characters because ⴳⵯ and ⴽⵯ do not have a Unicode codes.

3 Preprocessing

In this section, we described in details the first phase used which is the preprocessing to prepare the image for the next phases. To this end, as shown in figure 1, we performed a series of operations that are the image binarization, the lines skew correction, the image segmentation into characters and in last the characters normalization.

3.1 Binarization

The output of this operation is a binary image where black pixels represent the text and white pixels indicate the background. In this regard, several algorithms have been proposed in [15].

In this work, we used the nonparametric and unsupervised Otsu’s method [16]. This method performs an automatic thresholding that consists on maximizing the separability of the resultant classes in gray levels. It uses the zeroth and the first cumulative moments of the image histogram. The Otsu method gives good results and it still one of the most used thresholding methods.

3.2 Skew correction

The correction of the line skew consists in rectifying horizontally the oblique writing lines. Several methods are available in [17]. The two most popular are the Hough transform and histograms projection. In this paper, we used the histograms projection, for its simplicity and its rapidity, based on scanning image according to directions D close to the horizontal, and counting the number of black pixels in these directions for each line. The quality of histogram is estimated by its entropy. The most probable direction is the one who maximizes this entropy. The document angle θ is that which corresponds to the histogram of maximum entropy. To correct this inclination, simply apply an image rotation with the angle θ.

3.3 Segmentation

The characters segmentation is one of the most important steps in an Optical Character Recognition system (OCR). The objective is to decompose the image into a sequence of sub-images, each sub-image must contain a single character. For this, a lines segmentation of the image is performed, then each line is segmented into characters. A survey of methods and strategies in character segmentation is presented in [18].

3.3.1 Lines segmentation

In order to segment the image text into lines, we used the horizontal projection histogram. This method can distinguish between high density areas characterizing lines and low density areas indicating the space between the lines.

3.3.2 From lines to characters

Since Amazigh writing, handwritten or printed, is never cursive, character extraction from each line becomes easy. In this context, we used vertical projection histogram. Characters correspond to areas of high density in the histogram.

3.4 Normalization

The segmentation process produces isolated characters in different size, to solve this problem we proceeded to the normalization. This latter consists on resizing all characters to a common size. In this work, due to its zooming quality, we used a spline-based algorithm [19] to resize all characters to a size of 30x30. This optimal spline-based algorithm for the enlargement or reduction of digital images can be realized through a new method of finite differences by calculating the scalar products with analysis functions that are B-splines of any degree. This algorithm achieves a reduction of artifacts such as aliasing and blocking and a significant improvement of the signal-to-noise ratio.

4 Features extraction

The features extraction step is a very important operation for a system of handwriting recognition. Its aim is the selection of the most relevant informations identifying each character to form a features set.

In the literature, many features extraction methods have been applied for OCR systems. Arica.N and al [20] categorize them according to their type as follows:

- Global Transformation and Series Expansion;
- Statistical features;
- Geometrical and topological features.

In this paper, after preprocessing phase, we used a feature set based on statistical methods. This latter rests on the decomposition of the character segmented image into several zones according to different directions, then their density and their Length of projections are calculated. A detailed description of our approach is given in Section 6.

5 Character classification

After features extraction of each segmented and normalized image, we used the resulting vector, consists of 79 components, characterizing each character for learning and testing.

For this, several classification approaches are used in the field of handwriting recognition. According to [21], recognition techniques and text classification are grouped into four main categories:

- Pattern matching methods using correlation and distance measure;
- Statistical methods based on discriminant functions;
- Structural and syntactic methods employing rules and grammars;
- Neural networks classifiers.

In this paper, we used the MLP architecture, one of the great families of neural networks, which use a supervised learning method called backpropagation [22]. The MLP used contains three layers:

- The input layer that consists of 79 nodes for the extracted vector ;
- The output layer with 31 nodes to distinguish the 31 classes which represent the number of studied characters in the Amazigh language;
- The hidden layer whose the number of nodes is chosen experimentally.

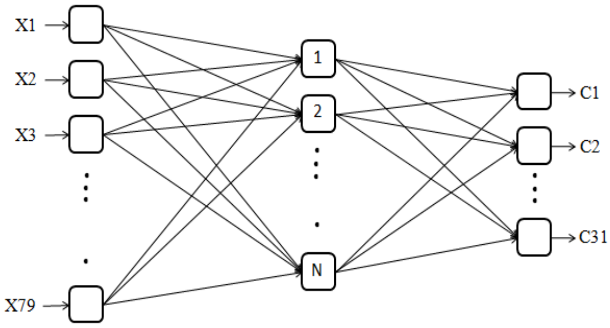


Fig.3: Architecture used of Multilayer perceptrron.

It should be noted that using one hidden layer is sufficient to solve a non linear complex problem and the choice of the number of hidden layer is still a challenging issue [23].

6 Our approach

The choice of relevant features influences largely the performance of the character recognition system. In this context, we developed a feature set to provide a description that can characterize each character. This feature set consists of two subsets: the first one is generated by dividing the image into different overlapped zones, then we compute the densities of black pixels in each zone; concerning the second subset, we calculate the shadow features [24] in different zones, shadow features are the lengths of projections on different sides of the considered zones. The resulting feature set contains 79 components which 37 come from the first subset and the remainder is obtained by the second. To implement this method, the sizes of characters images under study are all resized to 30 × 30. In this section we presented in details the different stages to construct our feature set.

6.1 Density features

To create the first features subset, we carry out different decompositions of the character image and the density of foreground pixels is calculated in each zone to obtain 37 features. We obtained the density of each zone by dividing the number of black pixels by the total number of pixels in this zone.

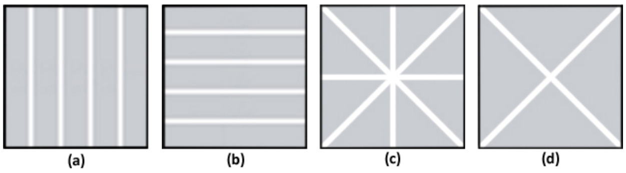


Fig.4: First decomposition of character image
(a) Decomposition to 5 vertical equal zones;
(b) Decomposition to 5 horizontal equal zones;
(c) Decomposition to 8 octants;
(d) Decomposition to 4 quadrants.

The first decomposition, as shown in the Figure 3, consists of dividing the character image vertically (Fig 4.a) and horizontally (Fig 4.b) to five equal zones, then to 8 octants (Fig 4.c) and in last to 4 quadrants (Fig 4.d).

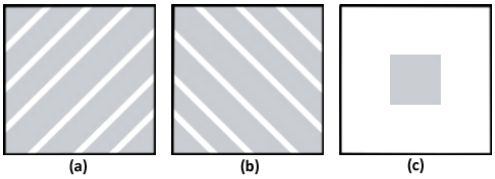


Fig.5: Second decomposition of character image
(a) Left diagonal decomposition;
(b) Right diagonal decomposition;
(c) 10x10 middle zone.

The second decomposition is obtained by dividing the character image to 7 diagonal zones in both left and right directions (Fig 5-a, 5-b), then considering only the middle zone which size is 10x10 (Fig 5-c). Diagonal features increase the recognition accuracy and reduce the misclassification. The middle zone was added to distinguish between some resembling Amazigh characters.

Following these decompositions, we obtained 37 zones and we calculated the density for each zone.

6.2 Shadow features

The 42 shadow features are obtained from the first decomposition of the rectangular boundary enclosing the character image (Fig 4).

We calculated the 10 lengths of horizontal and vertical projections (Fig 6-a, 6-b), 24 lengths of projections on each of three sides of each octant (Fig 6-c, 6-d), and the 8 horizontal and vertical projections of each quadrant (Fig 6-e, 6-f).

Each calculated value must be normalized by dividing it on the maximum possible length of projections on the corresponding side.

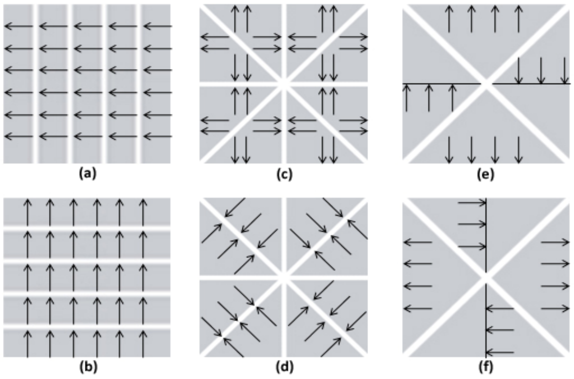


Fig.6: Decomposition for shadow features
(a) Shadow features of 5 vertical zones;
(b) Shadow features of 5 horizontal zones;
(c) Horizontal and vertical shadow features for each octant;
(d) Diagonal shadow features for each octant;
(e) Vertical shadow features for each quadrant;
(f) Horizontal shadow features for each quadrant.

7 Experimental results and discussions

7.1 Database

In order to evaluate the performance of the proposed OCR system, the AMHCD database [25] was used as a source of training and test. The database consists of 25740 isolated Amazigh handwritten characters produced by 60 writers who wrote 13 samples of each 33 classes. As mentioned in section 2, researchers work only on 31 characters excluding characters ⴳⵯ and ⴽⵯ. So, experimentations were carried out on 24180 characters where 70% (16926) were for training and the rest 30% (7254) for test. Noted that the training set is composed of 546 character of each class and the test set is composed of 234 character of each class.

7.2 MLP

The fully connected three-layer perceptron neural network was trained using a sigmoidal activation function, a learning rate of 0.1, a momentum of 0.25 and all weights were randomly initialized in interval [-0.7,0.7]. Several runs of backpropagation algorithm with 16926 epochs were performed for different architectures by varying the number of nodes in the hidden layer. The runs were executed in a compatible HP, Intel (R) Core (TM) Duo CPU 1.4 GHz, and 2 GB of RAM through Java. Table 1 shows the results obtained in tests with the chosen architectures using the feature set developed in this work.

Table 1: Recognition rate for different Number of hidden nodes

Number of hidden nodes	Accuracy on test set (%)	Recall (%)	Root mean squared error (%)
55	95,60	95,6	4,95
60	95,79	95,8	4,86
65	95,92	95,9	4,72
70	96,01	96	4,69
75	96,25	96,2	4,57
80	96,19	96,2	4,57
85	96,22	96,2	4,54
90	96,26	96,3	4,54
95	**96,47**	**96,5**	**4,46**
100	96,46	96,5	4,46

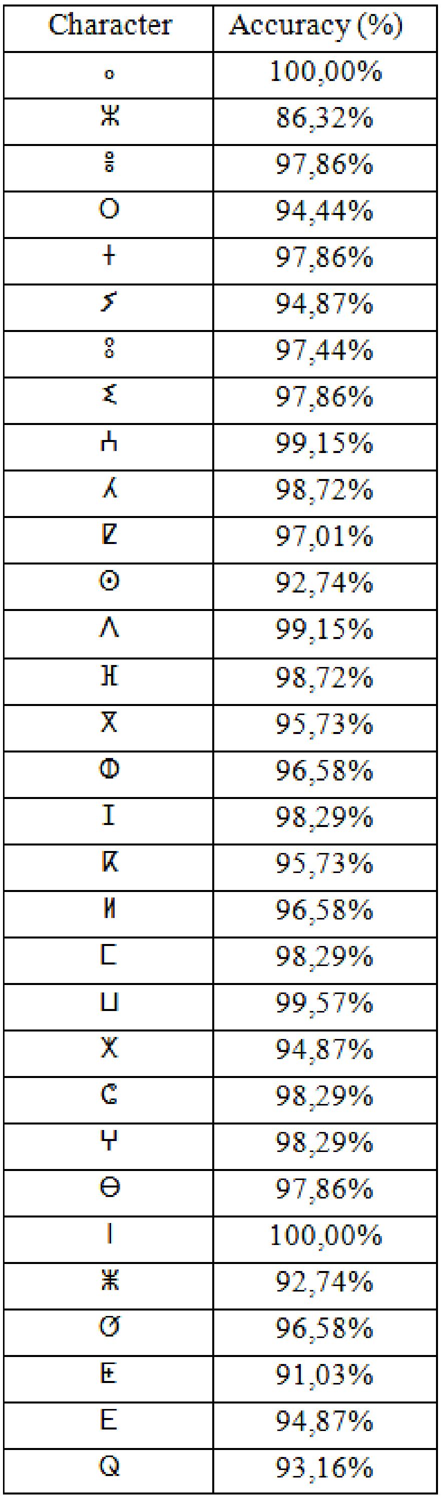
It should be noted that the learning phase is executed only once and weights are stored in a file, simply load the weights before performing recognition.

7.3 Discussions

According to the experiences, the best recognition performance of MLP is obtained when the number of hidden neurons is set to 95. In our experiences, we did not exceeded 100 neurons in the hidden layer to select a number of neurons that provides a compromise between performance and the time taken in the recognition. Thereby, we opted for the 79-95-31 architecture for our MLP which allowed us to achieve a recognition rate of 96.47% and a reasonable time in the recognition phase which is 4 milliseconds for each character.

Basing on this architecture, we computed the individual accuracy for each class of the Amazigh characters in the test data, the results obtained are shown in Table 2.

Table 2: Individual recognition rate for each character



Character	Accuracy (%)
ⴰ	100,00%
ⵣ	86,32%
ⴻ	97,86%
ⵔ	94,44%
ⵜ	97,86%
ⵡ	94,87%
ⵓ	97,44%
ⵉ	97,86%
ⵃ	99,15%
ⵢ	98,72%
ⵇ	97,01%
ⵙ	92,74%
ⴷ	99,15%
ⴼ	98,72%
ⴳ	95,73%
ⵀ	96,58%
ⵊ	98,29%
ⴽ	95,73%
ⵍ	96,58%
ⵎ	98,29%
ⵛ	99,57%
ⵅ	94,87%
ⵄ	98,29%
ⵖ	98,29%
ⴱ	97,86%
ⵏ	100,00%
ⵥ	92,74%
ⵚ	96,58%
ⵟ	91,03%
ⴹ	94,87%
ⵕ	93,16%

The obtained results show that some characters have a relatively low recognition rate compared to others, especially for characters yaz (ⵣ),yatt (ⵟ), yazz (ⵥ) and yas (ⵙ). The misclassifications are due to 2 factors. The first one is the structural similarity of characters. Table 3 shows the confusions matrix between characters.

Table 3: confusions Matrix between characters

	ⴰ	ⵣ	ⵓ	ⵔ	ⵜ	ⴼ	ⵢ	ⵉ	ⴻ	ⵃ	ⵇ	ⴱ	ⴷ	ⵄ	ⴳ	ⵀ	ⵊ	ⴽ	ⵍ	ⵎ	ⵡ	ⵅ	ⵛ	ⵖ	ⵙ	ⵏ	ⵥ	ⵚ	ⵟ	ⴹ	ⵕ
ⴰ	234	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ⵣ	0	202	0	0	0	0	0	0	0	0	0	0	0	0	3	0	2	0	0	0	0	4	0	0	0	0	23	0	0	0	0
ⵓ	0	0	229	0	1	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ⵔ	0	0	0	221	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	3	0	0	1
ⵜ	0	1	0	0	229	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0
ⴼ	0	0	1	0	0	222	0	8	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ⵢ	0	0	0	4	0	0	228	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
ⵉ	0	0	0	0	0	3	0	229	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ⴻ	0	0	0	0	1	0	0	0	232	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ⵃ	0	0	0	0	0	0	0	0	1	231	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
ⵇ	0	0	0	1	0	0	0	0	0	0	227	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	1	0	0
ⴱ	0	0	0	13	0	0	0	0	0	0	0	217	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0
ⴷ	0	0	0	0	0	0	0	0	0	1	0	0	232	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ⵄ	0	0	0	0	1	0	0	0	0	0	0	0	0	231	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
ⴳ	0	1	0	0	0	1	0	0	0	0	0	0	1	0	224	0	0	3	0	0	0	2	0	0	0	0	2	0	0	0	0
ⵀ	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	226	0	1	0	0	0	0	0	0	0	0	0	2	0	0	2
ⵊ	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	230	0	0	1	0	0	0	0	0	1	0	0	0	0	0
ⴽ	0	2	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	224	1	0	0	0	0	0	0	0	0	0	4	0	0
ⵍ	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0	0	226	0	3	0	0	1	0	0	0	0	0	0	0
ⵎ	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	230	0	0	0	0	0	0	0	0	0	1	0
ⵡ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	233	0	0	0	0	0	0	0	0	0	0
ⵅ	0	6	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	222	0	0	0	0	4	0	0	0	0
ⵛ	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	230	0	0	0	0	1	0	0	0
ⵖ	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	230	0	0	0	0	0	0	0
ⵙ	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	229	0	0	1	0	0	1
ⵏ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	234	0	0	0	0	0
ⵥ	0	12	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	217	0	0	0	0
ⵚ	0	0	0	4	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	226	0	0	0
ⵟ	0	0	0	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	213	16	0
ⴹ	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	4	222	0
ⵕ	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	3	4	0	0	2	0	0	218

The second factor is bad writing of some characters in the database whose classification is difficult even for a human operator, figure 7 illustrates some badly written letters in the database.

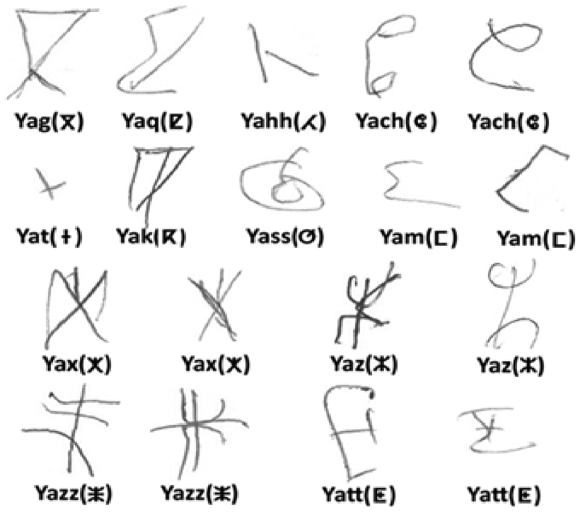


Fig.7: some characters badly written in database

To showcase our method, we compared our obtained results with the other best approaches in term of recognition rate (Table 4). It should be noted that M. Amrouch and al system [2] reports the highest recognition accuracy of about 97.89% in handwritten Amazigh language OCR by using continuous HMMs.

Table 4: Comparison with other approaches

Approaches	Recognition Rate (%)	Training set size	Test set size
Our Approach	96,47	16926	7254
M. Amrouch and al [2]	97,89	16120	8060
Y. Es Saady and al [7]	96,32	18135	2015
A. Djematene and al [6]	92.30	1000	700
S. Gounane and al [12]	91.05	1940	Not specified

The experimentations gave a satisfactory recognition rate compared to other approaches and showed that our system offers many advantages by using statistical methods.

In fact, our system is easy to understand, fast, minimizes loss of information, and has a low complexity.

In the other hand, and as the statistical methods, the system is sensible to noise and distortions and does not allow the reconstruction of the original image [20].

8 Conclusion

In this work, we had proposed an optical character recognition system of handwritten Amazigh characters employing a statistical approach to develop a new feature set. This one consists on calculating the densities and shadow features of each character by decomposing the image under study in term of zones; basing on this latter, we extract a vector of 79 components to represent each character. Because of its performance, we had used the MLP in the recognition phase. Some experimental results are introduced.

According to the experimental analyses, we conclude that the chosen statistical features are useful features to describe Amazigh characters and recognition rate can be very satisfactory.

In future works, to overcome the statistical methods shortcoming, we will propose a hybrid features set basing on the statistical and structural methods. And to extend our system, we will work on the extraction of Amazigh text from natural scenes before proceeding to its recognition.

References:

[1] X. Peng, H. Cao, S. Setlur, V. Govindarju, and P. Natarajan: “Multilingual OCR research and applications: An Overview”, *Proceedings of the 4th International Workshop on Multilingual OCR,* Article No.1, 2013.

[2] M. Amrouch, Y. Es-saady, A. Rachidi , M. El Yassa, and D. Mammass: “Handwritten Amazigh Character Recognition System Based on Continuous HMMs and Directional Features”. *International Journal of Modern Engineering Research*, Vol. 2, No. 2, 2012, pp. 436-441.

[3] A. Oulamara, and J. Duvernoy: “An application of the Hough transform to automatic recognition of Berber characters”, *Signal Processing*, Vol. 14, Issue.1, 1988, pp. 79-90.

[4] S. Gounane, M. Fakir, and B. Bouikhalene: “Recognition of Tifinagh Characters Using Self Organizing Map And Fuzzy K-Nearest Neighbor”, *Global Journal of Computer Science and Technology*, Vol. 11, No. 15, 2011, pp. 28-34.

[5] O. Bencharef, M. Fakir, and B. Minaoui: “Tifinagh Character Recognition Using Geodesic Distances, Decision Trees & Neural Networks”, *International Journal of Advanced Computer Science and Applications*. Special Issue on Artificial Intelligence, 2011, pp. 51-55.

[6] A. Djematene, B.Taconet, and A. Zahour: “A Geometrical Method for Printing and Handwritten Berber Characters Recognition”, *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, Vol.2, 1997, pp. 564 – 567.

[7] Y. Es Saady, A. Rachidi, M. El Yassa, and D. Mammass: “Amazigh Handwritten Character Recognition based on Horizontal and Vertical Centerline of Character”, *International Journal of Advanced Science and Technology*, Vol. 33, 2011, pp. 33-50.

[8] Y. Es Saady, A. Rachidi, M. El Yassa, and D. Mammass: “Printed Amazigh Character Recognition by a Syntactic Approach using Finite Automata”. *ICGST International Journal on Graphics, Vision and Image Processing*, Vol. 10, No. 2, 2010, pp. 1-8.

[9] M. Oujaoura, R. El Ayachi, B. Minaoui, M. Fakir, and B. Bouikhalene: “Invariant Descriptors and Classifiers Combination for Recognition of Isolated Printed Tifinagh Characters”, *International Journal of Advanced Computer Science and Application*, Special Issue SITACAM, 2013, pp. 22-28.

[10] M. Amrouch, Y. Es-Saady, A. Rachidi, M. El Yassa, and D. Mammass: “Printed Amazigh Character Recognition by a Hybrid Approach Based on Hidden Markov Models and the Hough Transform”, *Multimedia Computing and Systems*, 2009, pp. 356-360.

[11] H. Moudni, M. Er-rouidi, M. Oujaoura, and O. Bencharef: “Recognition of Amazigh characters using SURF & GIST descriptors*”, International Journal of Advanced Computer Science and Application*. Special Issue SITACAM, 2013, pp. 41-44.

[12] S. Gounane M. Fakir, and B. Bouikhalene: “Handwritten Tifinagh Text Recognition Using Fuzzy K-NN and Bi-gram Language Model”, *International Journal of Advanced Computer*

Science and Applications, Special Issue SITACAM, 2013, pp. 29-32.

[13] M. Ameur, A. Bouhjar, F. Boukhris, A. Boukouss, A. Boumalk, M. Elmedlaoui, E. Iazzi, and H. Souifi: “Initiation à la langue amazighe”, *Publications de l'Institut Royal de la Culture Amazighe*, Manuels N.1, 2004, pp. 9.

[14] http://www.ircam.ma/doc/divers/presentation_of_ircam.pdf, presentation of The Royal Institute of the Amazigh Culture, 2014.

[15] M. Sezgin, and B. Sankur: “Survey over image thresholding techniques and quantitative performance evaluation”, *Journal of Electronic Imaging*, Vol. 13, No. 1, 2004, pp. 146-168.

[16] N. Otsu: “A Threshold Selection Method from Gray-Level Histograms”, *IEEE Trans. Syst. Man Cybern*, Vol. SMC-9, No. 1, 1979, pp. 62-66.

[17] W. Chin, A. Harvey, and A. Jennings: “Skew Detection in Handwritten Scripts”, *TENCON '97, IEEE Region 10 Annual Conference, Speech and Image Technologies for Computing and Telecommunications. Proceedings of IEEE*, Vol. 1, 1997, pp. 319-322.

[18] G. Casey, and E. Lecolinet: “A Survey of Methods and Strategies in Character Segmentation”, *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, Vol. 18, No. 7, 1996, pp. 690-706.

[19] A. Muñoz Barrutia, T. Blu, and M. Unser: “Least-Squares Image Resizing Using Finite Differences”, *IEEE Transactions on Image Processing*, Vol. 10, No. 9, 2001, pp. 1365-1378.

[20] N. Arica, and F.T. Yarman-Vural: “An overview of character recognition focused on off-line handwriting”, *IEEE Trans. Syst. ManCybern. C Appl*, Vol. 31, No. 2, 2001, pp. 216-232.

[21] A.K. Jain and Mao Jianchang: “Statistical Pattern Recognition: A Review”. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, Issue. 1, 2000, pp. 4-37.

[22] J. De Villiers, and E. Barnard: “Backpropagation Neural Nets with One and Two Hidden Layers”, *IEEE Transactions on Neural Networks*, Vol. 4, No. 1, 1992, pp. 136 – 141.

[23] S. Karsoliya: “Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture”, *International Journal of Engineering Trends and Technology*, Vol. 31, No. 6, 2012, pp. 714-717.

[24] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri, and D.K. Basu: “Handwritten Bangla alphabet recognition using MLP based classifier”, *Proc. of the 2nd National Conf. on Computer Processing of Bangla*, 2005, pp. 285-291.

[25] Y. Es Saady, A. Rachidi, M. El Yassa, and D. Mammass: “AMHCD: A Database for Amazigh Handwritten Character Recognition Research”, *International Journal of Computer Applications*, Vol. 27, No. 4, 2011, pp. 44-48.