

Invariant Descriptors and Classifiers Combination for Recognition of Isolated Printed Tifinagh Characters

M. OUJAOURA, R. EL AYACHI, B. MINAOUI,
M. FAKIR and B. BOUIKHALENE
Computer Science Department
Faculty of Science and Technology
Sultan Moulay Slimane University
Béni Mellal, Morocco

O. BENCHAREF
Computer Science Department
Higher School of Technology
Cadi Ayyad University
Essaouira, Morocco

Abstract—In order to improve the recognition rate, this document proposes an automatic system to recognize isolated printed Tifinagh characters by using a fusion of 3 classifiers and a combination of some features extraction methods. The Legendre moments, Zernike moments and Hu moments are used as descriptors in the features extraction phase due to their invariance to translation, rotation and scaling changes. In the classification phase, the neural network, the multiclass SVM (Support Vector Machine) and the nearest neighbour classifiers are combined together. The experimental results of each single features extraction method and each single classification method are compared with our approach to show its robustness.

Keywords— Recognition system; Legendre moments; Zernike moment; Hu moments; Neural Networks; Multiclass SVM; nearest neighbour classifier

I. INTRODUCTION

The Optical Character Recognition (OCR) is a field of pattern recognition that focuses on character shapes.

In the recent years, some research works of Tifinagh characters are published, such as:

- Es Saady and all [1] are treated in their paper printed characters Amazigh isolated using an automatic character recognition based on the formalism of finite automata.
- Bencharef and all [2] are used the metric descriptors based on the calculation of the Riemannian metric. These descriptors are known for their reliability towards the change of scale, the existence of noise and geometric distortions. For the classification of Tifinagh characters, SVM and neural networks were used.
- El Ayachi and all [3] are developed a system Tifinagh characters recognition using a multilayer neural networks in classification step. In the extraction phase, Walsh transform is adopted.

In this work, the proposed system (Fig.1) contains three phases to recognize isolated printed Tifinagh characters: Pre-processing, Extraction and Classification phases.

In the pre-processing phase, normalization is applied to remove unwanted areas using the method of histogram; In this phase, we first calculate the horizontal and vertical histograms, then the histogram is scanned horizontally in both directions:

respectively from the top to the bottom and from the bottom to the top until finding the first black pixels, thereafter, the vertical histogram is traversed in both directions: respectively from the left to the right and from the right to the left until finding the first black pixels. Finally, after determining the positions of the first black pixels, we eliminate the unwanted areas.

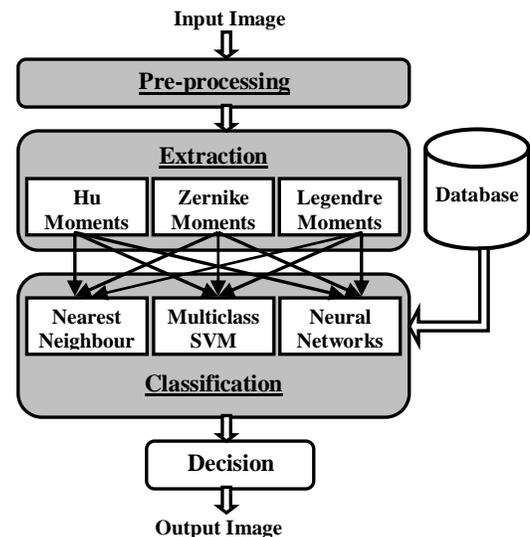


Fig. 1. Bloc diagram of the proposed recognition system

To calculate the feature of image, Legendre moments, Hu moments and Zernike moments are used in the second phase which is the extraction phase.

The classification phase adopts the neural networks, the SVM (Support Vector Machine) and the nearest neighbour classifier to recognize the correct character.

Most of the reported works for recognition systems uses a single method for calculating the parameters of an image and a single approach in the classification phase. The objective of this paper is to combine three methods in the extraction phase and three approaches to classify the characters in order to improve the recognition rate and reduce the error rate.

The rest of the paper is organized as follows. The Section 2 discusses the primordial task of any recognition system. It's the classification problems using some classifiers based on a combination of several algorithms like neural network, SVM

and nearest neighbour classifiers. The Section 3 is reserved for another important task which is the features extraction problems in addition to a brief formulation for Hu moments, Legendre moments and Zernike moments as features extraction methods. The Section 4 presents the experimental results for the recognition system. Finally, the conclusion is given in the section 5.

II. CLASSIFICATION

The robustness of the recognition system is based on the decision given by the classification phase. As a result, this document presents some approaches (Neuronal Network, SVM, nearest neighbor classifiers) known in the recognition field [4].

A. Neural network

Neural networks (or artificial neural networks) learn by experience, generalize from previous experiences to new ones, and can make decisions [5, 6].

A multilayer neural network consists of an input layer including a set of input nodes, one or more hidden layers of nodes, and an output layer of nodes. Fig. 2 shows an example of a three layer network used in this paper, having input layer formed by M nodes, one hidden layer formed by L nodes, and output layer formed by N nodes. This neural network is trained to classify inputs according to target classes. The training input data are loaded from the reference database while the target data should consist of vectors of all zero values except for a one element, where its index is the class they are to represent. The transfer function used in this tree layer neural network is hyperbolic tangent sigmoid transfer function defined by:

$$tsig(x) = 2/(1 + \exp(-2x)) - 1 \quad (1)$$

According to authors in [7], the number of neurons in the hidden layer is approximately equal to:

$$L = E\left(1 + \sqrt{M(N+2)}\right) \quad (2)$$

Where:

E(x) denotes the integer part of x.

M and N are respectively the number of neurons in the input and output layers.

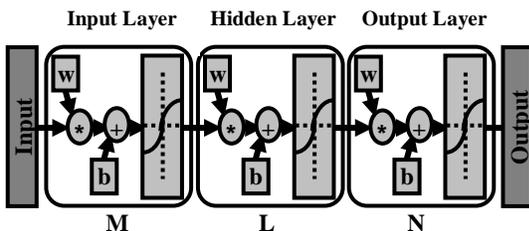


Fig. 2. The three layer neural network

B. Support vector machine (SVM)

Support vector machine (SVM) were originally designed for binary classification. As shown in Fig. 3, SVM is a classification method which is based on finding a hyper-plan that separates data sets into two classes. Several methods have been proposed to construct a multi-class classifier [9] by combining one-against-one binary classifiers or one-against-all binary classifiers. The data sets can be linearly separable or nonlinearly separable. The nonlinearly separable cases require the use of kernel function in order to obtain linearly separable data sets [8] [9].

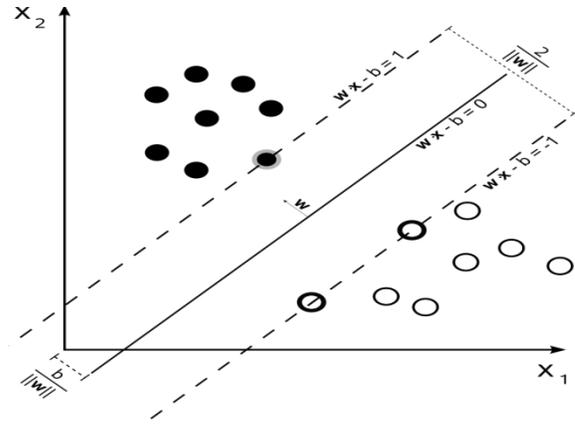


Fig. 3. Binary SVM with samples from two classes

In this paper, the one-against-one and the one-against-all binary classifier are used. Those classifiers are based on the Gaussian kernel function defined by:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (3)$$

Where $\sigma > 0$ equal to 1 in our case.

Many other kernel functions can be used for each binary classifier.

1) One-against-one binary classifier

From N class in data sets, the one-against-one multiclass SVM method constructs N(N-1)/2 binary classifier where each one is trained on data from two classes. The Structure the one-against-one multiclass SVM classifier can be represented by the Fig. 4.

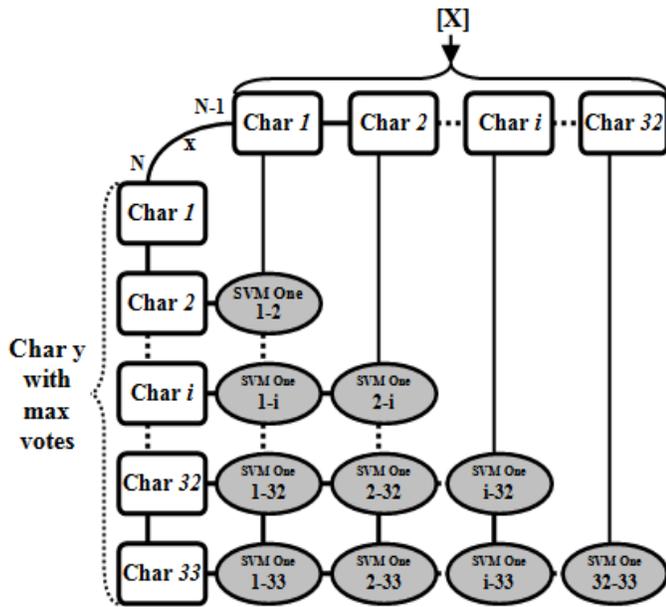


Fig. 4. Structure the one-against-one multiclass SVM classifier

To design and extend SVM binary classifier into a one-against-one multiclass SVM, two groups of data examples are constructed from two classes. The obtained SVM binary classifier is trained to decide if the class is from the first class or it belongs to the second class. This process is repeated for another couple of classes until finishing all the possible couples of the classes from data sets. So, by following this way, multiclass SVM is transformed to a multiple $N(N-1)/2$ SVM binary classifier. Each SVM binary classifier is trained using a matrix of training data, where each row corresponds to the features extracted as an observation from a class. When classifying an object with an input features vector, each binary classifier from the multiclass SVM one-against-one model decides and votes for only one class. The class with the majority votes is the correct class which the object belongs to.

2) One-against-all binary classifier

The one-against-all multiclass SVM classifier contains N binary classifier, where N is the number of class in data sets. The i^{th} binary SVM is trained with all of the data examples in the i^{th} class with positive labels, and all other data examples with negative labels.

To construct a one-against-all multiclass SVM model from binary classifier, the classes are divided into two groups: the first group is formed by one class, and the second group is all the other classes. The obtained SVM binary classifier is trained to decide if the class is from the first group or it belongs to the second groups of classes. This process is repeated for the second group that contains more than two classes until having only one class for each group. The process must stop there. So, by following this way, multiclass SVM is transformed to a multiple SVM binary classifier. Each SVM binary classifier is trained using a matrix of training data, where each row corresponds to features extracted as an observation from a class. After training phase, the multiclass SVM model is able to decide the correct class for an input features vector. To classify an object, its input features vector is presented

iteratively to the i^{th} against all binary classifier from the first to the N^{th} classifier while the result is negative. When the i^{th} binary classifier gives a positive result, the process is stopped. This means that the object belongs to the i^{th} class. The Structure of the one-against-all multiclass SVM classifier is given by the Fig. 5.

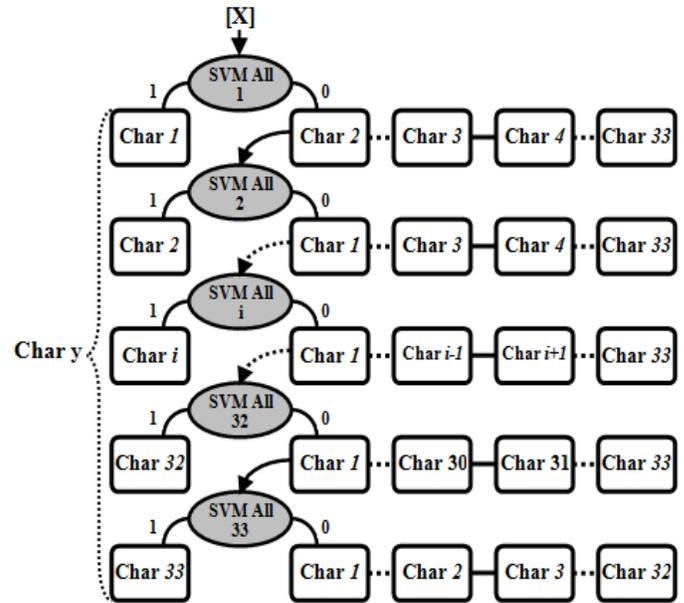


Fig. 5. Structure of the one-against-all multiclass SVM classifier

C. Nearest neighbour

The nearest neighbour classifier is used to compare the feature vector of the input image and the feature vectors stored in the database. It is obtained by finding the distance between the prototype image and the database. The class is found by measuring the distance between a feature vector of input image and feature vectors of images in reference database. The Euclidean distance measurement is used in this paper, but other distance measurement can be also used [10].

Let X_1, X_2, \dots, X_k be the k class features vectors in the database and X_q the feature vector of the query image. The feature vector with the minimum distance is found to be the closest matching vector. It is given by:

$$d(X_q, X_j) = \min_{j \in \{1, 2, \dots, k\}} \left\{ \sqrt{\sum_i (x_q(i) - x_j(i))^2} \right\} \quad (4)$$

The nearest neighbour classifier doesn't need any training phase. But, if the database is very large, it takes a considerable time to calculate all the distances between the query image and database classes.

III. EXTRACTION

Extraction is the second phase to apply in the recognition system. It is an important step, because the computed parameters will be used in the classification stage.

In this paper, the used approaches are: Hu moments [11], Legendre moments [12, 13, 14] and Zernike moments [15, 16].

A. Hu moments

For a discrete image of $M \times N$ pixels with intensity function $f(x, y)$, Hu [11] defined the following seven moments that are invariant to the change of scale, translation and rotation:

$$\phi_1 = \mu_{20} + \mu_{02} \quad (5)$$

$$\phi_2 = (\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2 \quad (6)$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \quad (7)$$

$$\phi_4 = (\mu_{30} - \mu_{12})^2 + (\mu_{21} - \mu_{03})^2 \quad (8)$$

$$\begin{aligned} \phi_5 = & (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \times \\ & [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + \\ & (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \times \\ & [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \end{aligned} \quad (9)$$

$$\begin{aligned} \phi_6 = & (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\ & + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \end{aligned} \quad (10)$$

$$\begin{aligned} \phi_7 = & (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12}) \times \\ & [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - \\ & (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03}) \times \\ & [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \end{aligned} \quad (11)$$

Where

$$\mu_{pq} = \frac{\alpha_{pq}}{\alpha_{00}^{\frac{p+q}{2}}} \quad \text{are the Normalized central moments}$$

$$\alpha_{pq} = \int_{a_1}^{a_2} \int_{b_1}^{b_2} \left(x - \frac{M_{10}}{M_{00}}\right)^p \left(y - \frac{M_{01}}{M_{00}}\right)^q f(x, y) dx dy \quad \text{are the central moments, and}$$

$$M_{pq} = \int_{a_1}^{a_2} \int_{b_1}^{b_2} x^p y^q f(x, y) dx dy \quad \text{are the two-dimensional geometric moment of order } (p + q) \text{ of an image } f(x, y).$$

B. Legendre moments

The Legendre moments were first introduced by Teague [12]. They were used in several pattern recognition applications [13]. The orthogonal property of Legendre polynomials implies no redundancy or overlap of information between the moments with different orders. This property enables the contribution of each moment to be unique and independent of the information in an image [14].

The Legendre moments for a discrete image of $M \times N$ pixels with intensity function $f(x, y)$ is the following:

$$L_{pq} = \lambda_{pq} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} P_p(x_i) P_q(y_j) f(x, y) \quad (12)$$

Where

$\lambda_{pq} = \frac{(2p+1)(2q+1)}{M \times N}$, x_i and y_j denote the normalized pixel coordinates in the range of $[-1, +1]$, which are given by:

$$\begin{cases} x_i = \frac{2x - (M-1)}{M-1} \\ y_j = \frac{2y - (N-1)}{N-1} \end{cases} \quad (13)$$

$P_p(x)$ is the p^{th} -order Legendre polynomial defined by:

$$P_p(x) = \sum_{k=0}^p \left\{ \frac{(-1)^{\frac{p-k}{2}} (p+k)! x^k}{2^p k! \left(\frac{p-k}{2}\right)! \left(\frac{p+k}{2}\right)!} \right\}_{p-k=\text{even}} \quad (14)$$

In order to increase the computation speed for calculating Legendre polynomials, we used the recurrent formula of the Legendre polynomials defined by:

$$\begin{cases} P_p(x) = \frac{(2p-1)x}{p} P_{p-1}(x) - \frac{(p-1)}{p} P_{p-2}(x) \\ P_1(x) = x, \quad P_0(x) = 1 \end{cases} \quad (15)$$

C. Zernike moments

Zernike moments are the mapping of an image onto a set of complex Zernike polynomials. As these Zernike polynomials are orthogonal to each other, Zernike moments can represent the properties of an image with no redundancy or overlap of information between the moments [15]. Due to these characteristics, Zernike moments have been utilized as feature sets in many applications [16].

The discrete form of the Zernike moments of an image size $M \times N$ represented by $f(x, y)$ is expressed, in the unit disk $x^2 + y^2 = 1$, as follows:

$$Z_{pq} = \frac{p+1}{\lambda} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} R_{p,q}(r_{xy}) e^{-jq\theta_{xy}} f(x, y) \quad (16)$$

Where

$$\left\{ \begin{aligned} R_{p,q}(r) &= \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s (p-s)! r^{p-2s}}{s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} \\ p-|q| &= (\text{even}), |q| \leq p, p \geq 0 \end{aligned} \right. \quad (17)$$

λ is the number of pixels located in the unit circle, and the transformed phase θ_{xy} and the distance r_{xy} at the pixel of coordinates (x, y) are [17]:

$$\left\{ \begin{aligned} \theta_{xy} &= \tan^{-1} \left(\frac{(2y - (N-1))/(N-1)}{(2x - (M-1))/(M-1)} \right) \\ r_{xy} &= \sqrt{\left(\frac{2x - (M-1)}{M-1} \right)^2 + \left(\frac{2y - (N-1)}{N-1} \right)^2} \end{aligned} \right. \quad (18)$$

Most of the time taken for the computation of Zernike moments is due to the computation of radial polynomials. Therefore, researchers have proposed faster methods that reduce the factorial terms by utilizing the recurrence's relations on the radial polynomials [17, 18]. In this paper, we obtained Zernike moments using the direct method and other method based on the recurrence's relations, defined by:

$$\left\{ \begin{aligned} R_{p,q}(r) &= \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s (p-s)! r^{p-2s}}{s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} \\ \text{if } q=0, p-|q| &= (\text{even}), |q| \leq p, p \geq 0 \\ R_{p,q}(r) &= r^p \\ \text{if } p &= q \\ R_{p,q}(r) &= \frac{2rpR_{p-1,q-1}(r) - (p-q)R_{p-2,q}(r)}{p+q} \\ \text{if } p \neq q, q &\neq 0 \end{aligned} \right. \quad (19)$$

The usage of direct method to compute radial polynomials in the case of $q = 0$ will considerably increase the computation time especially when p is very large.

IV. RESULTS

The Tifinagh alphabet adopted by IRCAM [19] is composed from thirty-three characters representing consonants and vowels as shown in Fig. 6.

To test the accuracy of the combined classifiers and descriptors, we calculated the recognition rate of each

classifiers and descriptors separately. In our experiment, we used voting rule classifier combination schemes. Then, each one of the 3 combined classifiers, for each descriptor, votes for the appropriate character. The character with maximum votes is selected and considered to be the suited character.

Character number	Character	Character number	Character	Character number	Character
1	ⵝ	12	ⵏ	23	ⵓ
2	ⵉ	13	ⵏ	24	ⵏ
3	ⵏ	14	ⵏ	25	ⵓ
4	ⵏ	15	ⵏ	26	ⵓ
5	ⵏ	16	ⵏ	27	ⵓ
6	ⵏ	17	ⵏ	28	ⵓ
7	ⵓ	18	ⵏ	29	ⵓ
8	ⵏ	19	ⵏ	30	ⵓ
9	ⵏ	20	ⵏ	31	ⵓ
10	ⵏ	21	ⵓ	32	ⵓ
11	ⵓ	22	ⵓ	33	ⵓ

Fig. 6. Tifinagh characters adopted by IRCAM

Finally, the efficacy of the recognition system is based on the decision given in the classification phase. The computation of recognition rate and error rate is presented in the following table (table 1):

TABLE I. RECOGNITION RATE AND ERROR RATE OF SINGLE CLASSIFIER AND DESCRIPTOR

	Descriptors	Classification Approaches			
		Neural Network	Nearest Neighbour	SVM One	SVM All
Recognition rate (%)	Hu Moments	72.73	83.64	56.97	41.82
	Zernike Moments	80.61	95.76	69.70	87.88
	Legendre Moments	81.21	97.58	49.70	78.79
Error rate (%)	Hu Moments	17.86	16.36	43.03	58.18
	Zernike Moments	19.39	4.24	30.30	12.12
	Legendre Moments	18.79	2.42	50.30	21.21
Execution time (s)	Hu Moments	8.34	1.42	94.01	3.36
	Zernike Moments	859.92	943.90	1152.36	754.65
	Legendre Moments	853.16	805.33	892.41	806.00

The recognition rate and error rate of the proposed method are presented in the following table (table 2):

TABLE II. RECOGNITION RATE AND ERROR RATE OF THE PROPOSED METHOD

	Recognition rate (%)	Error rate (%)	Execution time (s)
Proposed method	98.79	1.21	2248.09

From the results in Table 1 and Table 2, the recognition rate of the proposed system is improved in the recognition field of the isolated printed Tifinagh characters as shown in Fig. 7, despite of the processing time which is significantly increased.

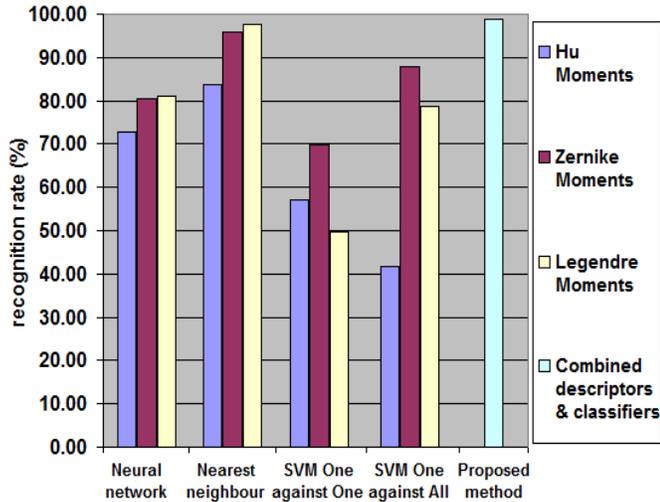


Fig. 7. Comparison of recognition rates

In order to show the robustness of the proposed approach, we give more details by calculating the confusion matrix of some method as presented in Fig. 8 and Fig. 9.

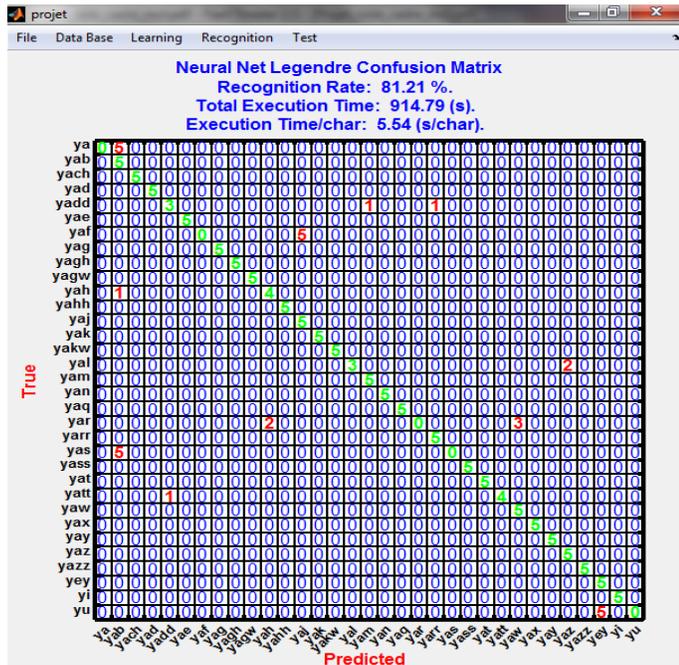


Fig. 8. Confusion matrix of Legendre Moments descriptor and Neural network classifier.

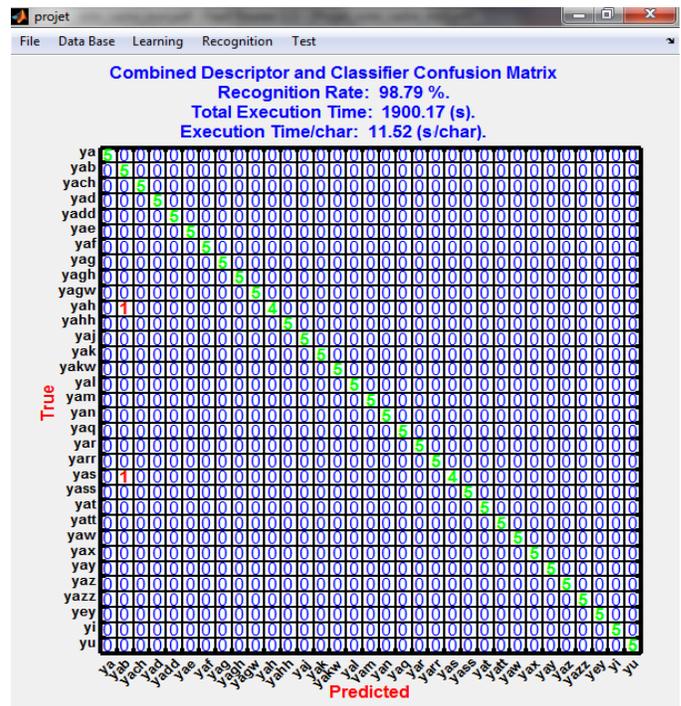


Fig. 9. Confusion matrix of the proposed method.

We can see from the 2 Figures of confusion matrix (Fig. 8 and Fig.9) that the misclassified characters (indicated by red color in the confusion matrix) in the case of using Legendre moment as single descriptor and neural network as single classifier are reduced in the case of using our proposed approach based on combination of many descriptors and classifiers.

All the tests are performed using a database containing a set of 33 reference characters and 165 image characters [20]. The proposed system has been implemented and tested on a core 2 Duo personnel computer using Matlab software.

V. CONCLUSION

The objective of the recognition system is to choose one image of the references images which is similar to the input image. The recognition system adopted in this paper contains three phases: Preprocessing, Extraction and Classification.

In the extraction phase, Hu moments, Legendre moments and Zernike moments are used to compute the attributes of the input image. Neural network, Multiclass SVM and nearest neighbour are three methods used in the classification phase.

Generally, the recognition system uses one method to calculate the parameters of the input image and one approach in the classification. In order to increase the recognition rate, three methods are combined in the extraction phase and three approaches are combined to make classification.

The obtained results show the robustness of the system based on the proposed approach in the recognition field of the isolated printed Tifinagh characters. The use of other fast descriptors and classifiers may increase the recognition rate of Tifinagh characters.

REFERENCES

- [1] Y. Es Saady, A. Rachidi, M. El Yassa, D. Mammass, « Printed Amazigh Character Recognition by a Syntactic Approach using Finite Automata », ICGST-GVIP Journal, Volume 10, Issue 2, June 2010.
- [2] O. Bencharef, M. Fakir, N. Idrissi, B. Bouikhalen et B. Minaoui, « Application de la géométrie riemannienne à la reconnaissance des caractères Tifinaghe », Agadir-Maroc, 06-07 Mai 2011. pp : 179 :188.
- [3] R. El Ayachi, K. Moro, M. Fakir et B. Bouikhalene, « On The Recognition Of Tifinaghe Scripts ». JATIT, vol. 20, No. 2, pp: 61-66, 2010.
- [4] Mustapha Oujaoura, Brahim Minaoui and Mohammed Fakir, Multilayer Neural Networks and Nearest Neighbor Classifier Performances for Image Annotation, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No. 11, pp.165–171, 2012. Published by The Science and Information Organization, New York, USA.
- [5] Yue Cao, Xiabi Liu, Jie Bing and Li Song, Using Neural Network to Combine Measures of Word Semantic Similarity for Image Annotation, IEEE International Conference on Information and Automation (ICIA), pp. 833 – 837, 2011.
- [6] P.Simard, D. Steinkraus, J. C. Platt, Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis, ICDAR, 2003, pp. 958-962.
- [7] R. Lepage, & B. Solaiman. Les réseaux de neurones artificiels et leurs applications en imagerie et en vision par ordinateur, Ecole de technologie supérieure, 2003.
- [8] R. Rifkin, A. Klautau. In defence of one-versus-all classification. Journal of Machine Learning Research, Vol. 5, pp. 101–141, 2004.
- [9] K.-B. Duan, S.S Keerthi, Which is the best multiclass SVM method? An empirical study. Technical Report CD-03-12, Control Division, Department of Mechanical Engineering, National University of Singapore, 2003.
- [10] Oren Boiman, Eli Shechtman and Michal Irani, In Defense of Nearest-Neighbor Based Image Classification, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2008.
- [11] R. El Ayachi, B. Bouikhalene et M. Fakir, « Transformation de Fourier et Moments Invariants Appliqués à la Reconnaissance des Caractères Tifinaghe », SITACAM'11, Agadir-Maroc, 06-07 Mai 2011. pp : 41-54.
- [12] F. L. Alt, Digital Pattern Recognition by Moments, J. Assoc. Computing Machinery, Vol. 9, pp. 240-258, 1962.
- [13] M.R. Teague, Image analysis via the general theory of moments, J. Opt. Soc. Amer. 70, pp. 920–930, 1980.
- [14] Chee-Way Chonga, P. Raveendranb and R. Mukundan, Translation and scale invariants of Legendre moments, Pattern Recognition 37, pp. 119 – 129, 2004.
- [15] Sun-Kyoo Hwang, Whoi-Yul Kim, A novel approach to the fast computation of Zernike moments, Pattern Recognition 39, pp. 2065 – 2076, 2006.
- [16] A. Prata, W.V.T. Rusche, Algorithm for computation of Zernike polynomials expansion coefficients, Appl. Opt. 28, pp. 749–754, 1989.
- [17] Mustapha Oujaoura, Brahim Minaoui and Mohammed Fakir. Article: Image Annotation using Moments and Multilayer Neural Networks. IJCA Special Issue on Software Engineering, Databases and Expert Systems SEDEX (1): pp. 46-55, September 2012. Published by Foundation of Computer Science, New York, USA.
- [18] C.W. Chong, P. Raveendran, R. Mukundan, A comparative analysis of algorithms for fast computation of Zernike moments, Pattern Recognition 36 (3) , pp. 731–742, 2003.
- [19] R. El Ayachi, M. Fakir and B. Bouikhalene « Recognition of Tifinaghe Characters Using Dynamic Programming & Neural Network », chapter in the book "Document Recognition and Understanding", ISBN 978-953-307-995-0, INTECH, 2011.
- [20] Y. Ait ouguengay , M. Taalabi,. Elaboration d'un réseau de neurones artificiels pour la reconnaissance optique de la graphie amazighe : Phase d'apprentissage". 5e Conférence internationale sur les "Systèmes Intelligents : Théories et Applications", Paris : Europa, cop. 2009 (impr. au Maroc), ISBN 978-2-909285-55-3. (2009)