FACTA UNIVERSITATIS (NIŠ)
Ser. Math. Inform. Vol. 28, No 2 (2013), 161–178

COMPARISON OF TWO FEATURE EXTRACTION METHODS BASED ON THE RAW FORM AND HIS SKELETON FOR GUJARATI HANDWRITTEN DIGITS

Kamal Moro, Mohammed Fakir, Badr Dine El Kessab, Belaid Bouikhalene and Cherki Daoui

Abstract. This paper deals with an optical character recognition (OCR) system for handwritten Gujarati numbers. One can find much work on Indian languages such as Hindi, Kannada, Tamil, Bangala, Malayalam, Gurumukhi etc., but Gujarati is a language for which hardly any work is traceable, especially for handwritten characters. In this paper we propose a comparison of two methods of feature extraction based on the raw form of the character and its skeleton and we show the advantage of using this method over other approaches mentioned in the paper.

1. Introduction

Gujarati belongs to the Devnagari language family, which originated and flourished in Gujarat, a western state of India, and is spoken by over 50 million people of the state. Though it has inherited a rich culture and literature and is a widely spoken language, hardly any significant work has been done for the identification of Gujarati optical characters. The Gujarati script differs from those of many other Indian languages in that it does not have any *shirolekha* (headlines). Gujarati numerals do not carry *shirolekha* and it applies to almost all Indian languages. The numerals in Indian languages are based on sharp curves and hardly any straight lines are used. Figure 1.1 shows a set of Gujarati numerals.

As it is visible in Figure 1.1, Gujarati digits are very peculiar by nature. Only two Gujarati digits one(1) and five(5) have a straight line, making Gujarati digit identification a little more difficult. Also, Gujarati digits often invite misclassification. These confusing sets of digits areas are shown in Figure 1.2.

Compared to OCR for printed characters, very limited work can be traced for handwritten character recognition for Indian languages. In 1980 Chinnuswami et al. [1] presented their work on recognition of hand-printed Tamil characters. The

Received October 08, 2012.; Accepted May 14, 2013.
2010 *Mathematics Subject Classification.* Primary 68T10; Secondary 68T05

authors used the structure of the Tamil characters. Using the curves and strokes of the characters, the features were identified and a statistical approach was used for the classification. Dutta et al.[2] recognized both printed and handwritten alphanumeric Bengali characters using curvature features in 1993. Here the features such as curvature maxima, curvature minima, and inflexion points were considered. In this work recognition was performed on isolated characters. Thinning and smoothing were also performed prior to classification of characters. In 2007, Banashree et al.[3] attempted an identification of hand written Hindi digits, using diffusion half toning algorithm.16-segment display concept was used here for feature extraction. They proposed a neural classifier for classification of isolated digits. The accuracy level they achieved was up to 98%. In 2008, Rajashekararadhya [4] proposed an offline handwritten OCR technique for four south Indian languages like Kannada, Telugu, Tamil and Malayalam. In this work a feature extraction technique was suggested, based on zone and image centroid. Two different classifiers in the nearest neighborhood and back propagation neural network were used to achieve 99% accuracy for Kannada and Telugu, 96% for Tamil and 95% for Malayalam. In 2009, Shanthi et al. [5] used support vector machine (SVM) for hand written Tamil characters and image subdivision for feature extraction. The recorded accuracy was 82.04%.

The character recognition systems offer potential benefits by providing an interface that facilitates interaction between man and machine. These systems are based on algorithms and essentially consist of three main steps : preprocessing, feature extraction, and classification. Skeletonization is considered an essential part of the preprocessing but not necessary, most researchers have used skeletonization ([6], [7], [8], [9]), while others have chosen to deal directly on the raw form ([10], [11], [12], [13]). We present a method for Gujarati character recognition. The feature extraction method used in this algorithm is applied to the raw form and on the skeleton. The average of these two vectors is the one that is used later in the classification step.

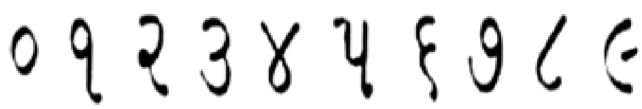

Fig. 1.1 – Gujarati digits

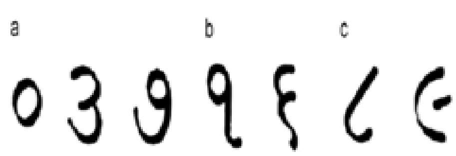

Fig. 1.2 – Confusing Gujarati digits

2. Database

For handwritten English numerals, we have the CEDAR (Centre of Excellence for Document Analysis and Recognition at the University of New York at Buffalo, USA) numeral database. It contains approximately 5000 samples of numerals. The samples were originally collected from US postal ZIP codes found on letters as there is no standard database available at the moment for Gujarati.

For developing a system to identify Gujarati handwritten digits, we collected numerals 0-9 written in Gujarati scripts by a large number of writers. These numbers were scanned in 300 dpi by a flatbed scanner. Initially, they are in separate boxes of 50*30 pixels each. Since our problem is to identify handwritten digits, the first thing required is to bring all the characters in a standard normal form. This is needed because when a writer writes he may use different types of pens and paper, and may even follow different styles of writing.

3. Binarization

Binarization is often the important first step in any process of character recognition. A large number of binarization techniques have been proposed in the literature [14], each being appropriate to a particular type of images. It has as a goal to reduce the amount of information present in the image, and keep only the relevant information.

According to several research papers [15, 16], the techniques of binarization of grayscale images can be classified into two categories : overall threshold, where a single threshold is used in the entire image to the devise in two classes (text and background), and local threshold where the values of the thresholds are determined locally, pixel-by-pixel or well region by region. In this document, we use the method referred to in [17], which is used to calculate the threshold of each pixel locally by following the formula :

$$T = (1 - k) * m + k * m + k * \sigma/(R(m - M)). \tag{3.1}$$

As k is set to 0.5 [10], the difference type and m the average of all the pixels in the window, M is the minimum image grey level and R is the maximum deviation of grayscale on all Windows.

4. Skeletonizatin

A fundamental problem in pattern recognition is a synthetic representation of this. In many cases, to work on the raw form is laborious and unnecessary. In terms of time and quality, it is much more advantageous to work with a refined shape. The notion of skeleton was introduced for this effect. In the ongoing plan, the skeleton of a shape is a set of lines passing through the middle. This is the concept of median axis of a continuous form introduced by Blum [18].

There is currently a wide variety of methods to construct skeletons from shapes, among them the topological thinning which consists of removing the points of the outline of the shape, while preserving its topological characteristics. In this document, we chose to use the Guo_Hall algorithm [19], which uses the parallel approach of thinning and preserves the topology and geometry. It is cited in [7].

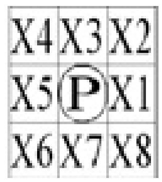

Fig. 4.1 – A point P and its neighborhoods

A point P (Figure 4.1) and its noted neighborhoods X1, X2, X3, X4, X5, X6, X7 and X8, the GUO_HALL algorithm is to remove parallel points of the object P checking the following conditions :
P is 4-adjacent to the supplementary object

$$2 \leqslant N(P) \leqslant 3 \tag{4.1}$$

$$(x_2 \vee x_3 \vee \overline{x_8}) \wedge x_1 = 0 \tag{4.2}$$

$$(x_6 \vee x_7 \vee \overline{x_4}) \wedge x_5 = 0 \tag{4.3}$$

with :

$$N_1(P) = (x_1 \vee x_2) + (x_3 \vee x_4) + (x_5 \vee x_6) + (x_7 \vee x_8) \tag{4.4}$$

$$N_2(P) = (x_2 \vee x_3) + (x_4 \vee x_5) + (x_6 \vee x_7) + (x_8 \vee x_1) \tag{4.5}$$

$$N(P) = Min(N_1(P), N_2(P)) \tag{4.6}$$

5. Feature Extraction

Feature extraction is the crucial step in numeral identification as each numeral is unique in its own way, thus distinguishing itself from other numerals. Hence, it is very important to extract features in such a way that the recognition of different numerals becomes easier on the basis of the individual features of each numeral. In this paper, we propose two extraction methods, each of which is applied on the raw form and the skeleton. The average of the two vectors is that which is subsequently applied to the classification level.

5.1. First method

In this first method, we use the Box-approach in Refs [6, 20, 21]. This approach requires a special division of the character image. The major advantage of this approach stems from its robustness to small variations, ease of implementation, and relatively high recognition rate. The choice of the box size and the number of boxes is discussed in Section 6. Each character image is divided into 24 boxes so that the portions of a numeral will be in some of these boxes. There could be boxes that are empty, as shown in Figure 5.1. The English numeral 3 is enclosed in the 6*4 grid. However, all boxes are considered for analysis in a sequential order. By considering the bottom left corner as the absolute origin (0,0), the coordinate distance (vector distance) for the kth pixel in the bth box at location (i,j) is computed as :

$$d_{kb} = (i^2 + j^2)^{1/2} \tag{5.1}$$

By dividing the sum of distances of all black pixels present in a box with the total number of pixels in that box, a normalized vector distance (γ_b) for each box is obtained as :

$$\gamma_b = 1/n_b \sum_{i=1}^{n_b} d_{kb} \tag{5.2}$$

where n_b is the total number of pixels in bth box. These vector distances constitute a set of features based on distances. Therefore, 24 γ_b's corresponding to 24 boxes will constitute a feature set. However, for empty boxes, the value will be zero.

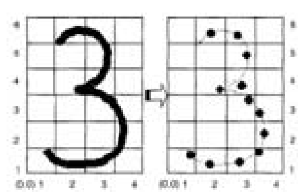

Fig. 5.1 – Portions of the numeral lie within some boxes while others are empty

5.2. Second method

In this part, we have chosen to use a method that is both simple and effective. This method consists of doing the sum of the values of pixels at the following levels : horizontal, vertical, and two diagonal. The cavities show the way to summon the pixels to an image 3 x 3. For example, in considering the form of the Figure 5.3, Table 5.1 shows vector extraction following Figure 5.2 patterns. This method is quoted in [12].

Table 5.1 – Extraction vectors

Figure 3	Extraction vector
A	(2,1,2)
B	(2,1,2)
C	(1,0,3,2,1)
D	(1,0,3,2,1)

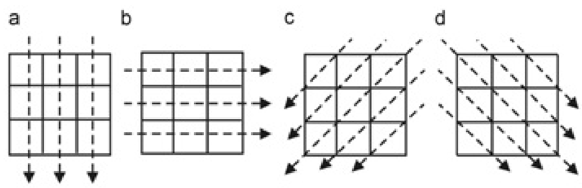

Fig. 5.2 – Pattern profile of 3x3 pattern matrix

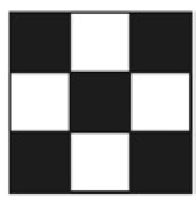

Fig. 5.3 – 3x3 Pattern

Throughout the paper, we will call the first method of extraction *distance method* and the second *sum method.*

6. Neural Networks

As [22, 23] have used neural network for character classification, neural network is suggested in this paper. A feed forward back propagation neural network is used for Gujarati numeral classification. This proposed multi-layered neural network consists of three layers with 50, 30, 10 for distance method and 118, 60 respectively, and 10 neurons for sum method. The input layer is the layer which accepts the profile vector which is of 1*50 and 1*118 in size. As this network is used for classification of 10 digits, it has 10 neurons in the output layer, the function sigmoid as function of activation at the step of the layer entry and hidden, with $\alpha \doteq 0.1$,

$$f(x) = 1/(1 + e^{-\gamma x}) \tag{6.1}$$

logsig at the step of the output layer and we fixed the constant learning to $\gamma \doteq 0.1$.

7. Training of Network

For this experiment, a total of 300 responses were taken into consideration. For training, the features are abstracted first for all of these images of digits. A profile vector for a digit five is shown here for the distance method :

[0 0 0.6672 0.7017 0 0 1.9429 2.8865 0 0 0.5770 2.0683 0 0 0 3.3593 0 0 1.1157 0 2.4256 1.8303 0 2.4254 3.5686 0 4.0410 6.6789 9.9418 0 0 0 0 7.1634 0 0 0 0 8.0269 0 0 9.6223 0 8.6789 0 0 0 10.4075 3.5689 0].

For the sum method :

[0 0 0 0 0 5 2 2 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 2 3 2 2 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 7 3 3 2 3 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 1 1 1 1 0 1 0 1 0 1 0 1 0 1 0 1 1 0 1 1 0 1 1 1 0 1 1 0 0 0 0].

To prevent overlearning, a set of validation characters is used. These characters have to define the best values of synaptic weights on the algorithm. Data validations are neutral in determination of the weight ; they serve only to stop the previous iteration, before the start of overlearning. In our case, we used 100 characters of validation. In Figure 7.1 the complete process of Gujarati numeral optical character recognition is shown.

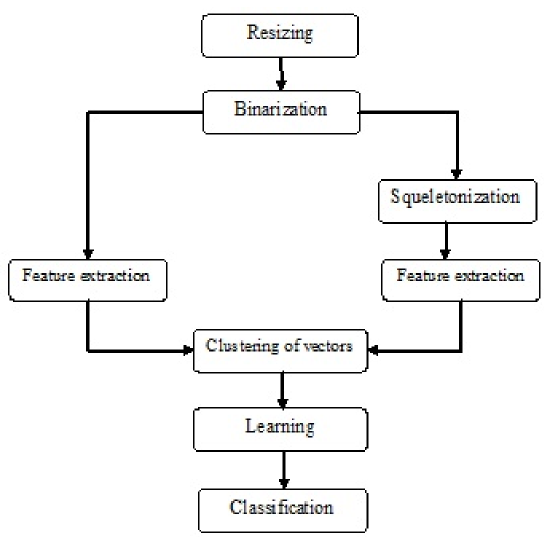

Fig. 7.1 – Recognition process

8. Experimental Results

As mentioned above, this network was trained for a total of 30 sets of digits, and was tested for 60 other new sets of digits. In total the network was trained by 300 digits and tested for 600 digits. Initially, we apply binarization on the character. This operation aims to eliminate the various intensities of gray pixels of the image to make binary. Figure 8.1 shows the result of the use of the Wolf algorithm [10].

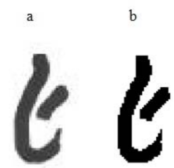

Fig. 8.1 – binarization of a digit, a : before binarization, b : after binarization

After binarization, we begin the skeletonization step. This approach is designed to present the form with a minimum of information ; Figure 8.2 shows the result of the Guo_Hall [12] algorithm that is used in this document.

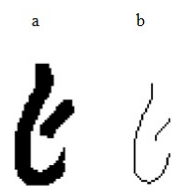

Fig. 8.2 – Skeleton of a digit, a : before sketetonization, b : after skeletonization

Overall, this network gave the following success rates : 83% for the distance method and 81.17% for the sum method using the proposed method ; 82.26% for the distance method and 79.67% for the sum method with a skeleton ; and 79,83% for the distance method and 76.33% for the sum method without a skeleton. The results are summarized in Table 8.1 and Table 8.2.

TABLE 8.1 – Result summary for distance method (%)

Sets	Type of sets	without skeleton	With skeleton	Proposed method
30 sets of digits	Training sets	92,67	98,67	94
60 sets of digits	Testing sets	79,83	82,26	83

Table 8.2 – Result summary for sum method (%)

Sets	Type of sets	without skeleton	With skeleton	Proposed method
30 sets of digits	Training sets	88,67	100	93
60 sets of digits	Testing sets	76,33	79,67	81,17

Let us examine the results obtained for different digits. For the distance method, Tables 8.3, 8.4 and 8.5 show confusion among the identified digits, while testing the proposed network for the training handwritten digits. Tables 8.6, 8.7 and 8.8 show the success rates for testing handwritten digits.

Table 8.3 – Network performance without skeleton on the learning set for distance method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	30	0	0	0	0	0	0	0	0	0	100
1	0	29	1	0	0	0	0	0	0	0	96,67
2	0	0	29	0	0	1	0	0	0	0	96,67
3	1	0	0	27	0	0	0	2	0	0	90,00
4	0	0	0	0	30	0	0	0	0	0	100
5	0	0	0	0	0	30	0	0	0	0	100
6	0	0	0	1	0	0	28	0	0	1	93,33
7	4	0	0	0	0	0	0	25	0	1	83,33
8	0	0	0	0	1	0	0	0	26	3	86,67
9	3	0	0	0	0	0	0	0	3	24	80,00

Table 8.4 – Network performance with skeleton on the learning set for distance method

Numbers	0	1	2	3	4	5	6	7	8	9	Success (%)
0	29	1	0	0	0	0	0	0	0	0	96,67
1	0	30	0	0	0	0	0	0	0	0	100
2	0	0	30	0	0	0	0	0	0	0	100
3	1	0	0	29	0	0	0	0	0	0	96,67
4	0	0	0	0	29	0	0	0	1	0	96,67
5	0	0	0	0	0	30	0	0	0	0	100
6	0	0	0	0	0	0	30	0	0	0	100
7	0	0	0	0	0	0	0	30	0	0	100
8	0	0	0	0	1	0	0	0	29	0	96,67
9	0	0	0	0	0	0	0	0	0	30	100

TABLE 8.5 – Network performance with proposed method on the learning set for distance method

Numbers	0	1	2	3	4	5	6	7	8	9	Success (%)
0	30	0	0	0	0	0	0	0	0	0	100
1	0	29	1	0	0	0	0	0	0	0	96,67
2	0	0	27	0	0	2	0	0	0	1	90,00
3	1	0	0	27	0	0	0	2	0	0	90,00
4	0	0	0	0	29	0	0	0	1	0	96,67
5	0	0	0	0	0	30	0	0	0	0	100
6	0	0	0	1	0	0	28	0	0	1	93,33
7	0	0	0	0	0	0	0	27	0	3	90,00
8	0	0	0	0	1	0	0	0	26	3	86,67
9	0	0	0	0	0	0	0	0	1	29	96,67

TABLE 8.6 – Network performance without skeleton on the test set for distance method

Numbers	0	1	2	3	4	5	6	7	8	9	Succes(%)
0	54	0	0	0	0	0	0	4	2	0	90,00
1	0	54	2	0	1	3	0	0	0	0	90,00
2	0	11	41	0	5	2	0	0	1	0	68,33
3	0	0	0	55	0	0	2	3	0	0	91,67
4	0	1	0	0	48	2	0	2	7	0	80,00
5	0	10	6	0	0	44	0	0	0	0	73,33
6	9	0	0	3	0	0	45	1	1	1	75,00
7	3	0	0	2	0	0	0	55	0	0	91,67
8	0	0	0	0	2	0	0	0	52	6	86,67
9	0	0	0	0	3	0	0	0	26	31	51,67

TABLE 8.7 – Network performance with skeleton on the test set for distance method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	54	0	0	1	0	0	1	0	3	1	90,00
1	0	53	1	0	2	4	0	0	0	0	88,33
2	0	11	44	1	4	0	0	0	0	0	73,33
3	0	0	0	57	0	0	0	3	0	0	95,00
4	0	3	0	3	47	0	0	0	5	2	78,33
5	0	11	6	1	0	41	1	0	0	0	68,33
6	0	0	0	9	0	2	46	3	0	0	76,67
7	2	0	0	1	1	0	1	54	0	1	90,00
8	0	0	0	0	2	0	0	0	56	2	93,33
9	0	1	0	1	0	0	0	0	14	44	73,33

TABLE 8.8 – Network performance with the proposed method on the test set for distance method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	52	0	0	0	0	0	0	4	3	1	86,67
1	0	53	2	0	2	3	0	0	0	0	88,33
2	0	8	43	0	5	2	0	0	1	1	71,67
3	0	0	0	55	0	0	1	4	0	0	91,67
4	0	1	0	0	49	0	0	1	3	6	81,67
5	0	11	5	0	0	42	2	0	0	0	70,00
6	1	0	0	5	0	0	48	0	0	6	80,00
7	1	0	0	2	0	0	0	57	0	0	95,00
8	0	0	0	0	1	0	0	0	58	1	96,67
9	0	0	0	0	2	0	0	0	17	41	68,33

The Tables 8.9,8.10, 8.11, 8.12, 8.13 and 8.14 show the confusion among for sum method.

TABLE 8.9 – Network performance without skeleton on the learning set for sum method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	29	0	0	0	0	0	0	0	0	1	96,67
1	0	30	0	0	0	0	0	0	0	0	100
2	0	2	26	2	0	0	0	0	0	0	86,67
3	0	2	0	28	0	0	0	0	0	0	93,33
4	0	0	0	0	30	0	0	0	0	0	100
5	0	0	0	0	1	28	1	0	0	0	93,33
6	0	1	0	0	0	3	26	0	0	0	86,67
7	0	0	0	0	0	9	0	21	0	0	70,00
8	1	0	0	0	1	0	0	0	28	0	93,33
9	0	0	0	0	0	8	0	0	2	20	66,67

TABLE 8.10 – Network performance with skeleton on the learning set for sum method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	30	0	0	0	0	0	0	0	0	0	100
1	0	30	0	0	0	0	0	0	0	0	100
2	0	0	30	0	0	0	0	0	0	0	100
3	1	0	0	30	0	0	0	0	0	0	100
4	0	0	0	0	30	0	0	0	0	0	100
5	0	0	0	0	0	30	0	0	0	0	100
6	0	0	0	0	0	0	30	0	0	0	100
7	0	0	0	0	0	0	0	30	0	0	100
8	0	0	0	0	0	0	0	0	30	0	100
9	0	0	0	0	0	0	0	0	0	30	100

Table 8.11 – Network performance with proposed method on the learning set for sum method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	30	0	0	0	0	0	0	0	0	0	100
1	0	30	0	0	0	0	0	0	0	0	96,67
2	0	2	26	2	0	0	0	0	0	0	90,00
3	0	2	0	28	0	0	0	0	0	0	90,00
4	0	0	0	0	30	0	0	0	0	0	96,67
5	0	0	0	0	1	28	1	0	0	0	100
6	0	0	0	0	0	2	28	0	0	0	93,33
7	0	0	0	0	0	8	0	21	1	0	90,00
8	0	0	0	0	0	0	0	0	30	0	86,67
9	0	0	0	0	0	0	0	0	2	28	96,67

Table 8.12 – Network performance without skeleton on the test set for sum method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	53	0	1	0	0	0	0	1	2	3	88,33
1	0	50	4	0	0	0	1	2	0	3	83,33
2	0	3	47	6	0	0	2	2	0	0	78,33
3	0	10	0	50	0	0	0	0	0	0	83,33
4	0	0	0	0	48	0	0	8	3	1	80,00
5	0	5	0	1	5	30	4	6	0	9	50,00
6	0	10	0	7	1	2	36	0	0	4	60,00
7	0	0	0	0	1	0	0	58	1	0	96,67
8	0	0	0	0	0	0	0	4	56	0	93,33
9	2	0	0	0	1	0	0	16	10	30	50,00

TABLE 8.13 – Network performance with skeleton on the test set for sum method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	39	0	0	0	0	0	0	17	1	3	65,00
1	0	48	6	2	0	3	0	1	0	0	80,00
2	0	2	44	6	0	1	5	2	0	0	73,33
3	0	1	3	53	0	0	2	1	0	0	88,33
4	0	0	1	0	53	0	0	4	1	1	88,33
5	0	7	1	0	5	47	0	0	0	0	78,33
6	0	0	0	11	0	0	45	4	0	0	75,00
7	0	0	0	3	0	0	2	55	0	0	91,67
8	0	0	0	0	0	0	0	0	57	3	95,00
9	0	1	0	0	1	0	0	10	11	37	61,67

TABLE 8.14 – Network performance with the proposed method on the test set for sum method

Numbers	0	1	2	3	4	5	6	7	8	9	Success(%)
0	59	0	0	0	0	0	0	0	1	0	98,33
1	0	47	4	0	0	4	3	2	0	0	78,33
2	0	1	50	5	0	0	1	3	0	0	83,33
3	0	3	0	56	0	0	0	1	0	0	93,33
4	0	0	0	0	50	0	0	8	0	2	83,33
5	0	9	0	0	5	30	7	9	0	0	50,00
6	0	4	0	5	1	3	46	0	0	1	76,67
7	0	0	0	1	4	0	0	54	1	0	90,00
8	0	0	0	0	0	0	0	0	60	0	100
9	0	0	0	0	2	0	0	6	17	35	58,33

What we can immediately notice in the tables above is that digits 9 are confused with 8, and digits 7 are slightly confused with other digits in the three approaches. Tables 8.15 and 8.16 show confusion of characters. We consider a character to be confused with another if the error rate exceeds 10%.

TABLE 8.15 – Confusion of characters with distance method

Character	Without Skeleton	With Skeleton	Proposed Method
0	Any	Any	Any
1	Any	Any	Any
2	1	1	1
3	Any	Any	Any
4	9	Any	8
5	1	1,2	1
6	9	3	0
7	Any	Any	Any
8	Any	Any	9
9	8	8	8

TABLE 8.16 – Confusion of characters with sum method

Character	Without Skeleton	With Skeleton	Proposed Method
0	Any	7	Any
1	Any	2	Any
2	3	3	Any
3	1	Any	Any
4	7	Any	7
5	7,9	1	1,6
6	1,3	3	Any
7	Any	Any	Any
8	Any	Any	Any
9	7,8	7,8	7,8

9. Conclusion

In this paper we proposed a feedforward back propagation neural network for the classification of Gujarati numerals. Various techniques are used in the preprocessing step before implementing classification of numerals. The overall performance of this proposed network is as high as 83% for the distance method and 81.17% for the sum method using the proposed method ; 82.26% for the distance method and 79.67% for the sum method with a skeleton, and 79.83% for the distance method and 76.33% for the sum method without a skeleton.

The performance of any classification model is mainly based on the feature abstraction. To improve the performance of this prototype, the improved features abstraction technique and/or the preprocessing techniques are possibly required. As a whole, this model offers a satisfactory success rate but it is subject to further improvement.

Références

1. P. Chinnuswamy, S.G. Krishnamoorty : *Recognition of hand-printed Tamil characters*. Pattern Recognition 12 (3) 1980, pp : 141-152
2. A. Dutta, S. Chaudhary : *Bungali Alpha-numeric character recognition using curvature features*. Pattern Recognition 26 (12) 1993, pp : 1757-1770.
3. N. P. Banasheree, D. Andhre, R. Vasanta, P.S. Satyanarayana : *OCR for script identification oh Hindi (Devnagari) numerals using error diffusion Halftoning Algorithm with neural classifier*. In Proceeding of World Academy of Science Engineering and Technology 20 (2007), pp : 46-50.
4. S.V. Rajashekararadhya, P. V. Ranjan : *Efficient zone based feature extraction algorithm for handwritten numeral recognition of popular south Indian*. Journal of Theoretical and Applied Information Technology 7 (1) (2009) pp : 1171-1180.
5. N. Shanthi, K. Duraiswamy : *Anovel SVM based hand written Tamil character recognition*. Pattern Analysis and Application, 2009, Publised Online : /http ://www.springerlink.com/content/k708558648652q12/S.
6. M. Hanmandly, O.V. Ramana Murthy : *Fuzzy model based recognition of handwrittern numerals*. In Pattern recognition 40, 2007, pp : 1840-1854
7. K. Moro, M. Fakir, B. Bouikhalene, S. Safi : *Skeletonization Methods Evaluation for the Recognition of Printed Tifinaghe Characters*. In Sitacam '09 Agadir Morocco, 12-13 December 2009, pp : 33-47
8. Elbaati A., Kherallah M., Ennaji A., Alimi A. M. : *Temporal Order Recovery of the Scanned Handwriting*. ICDAR'2009, Barcelone, Espagne, pp. 1116-1120
9. Kherallah M., Hadded L., Mitiche A., Alimi A. M. : *On-Line Recognition Of Handwritten Digits Based On Trajectory And Velocity Modelling*. International journal of Pattern Recognition Letter. Vol. 29. pp. 580-594. 2008
10. M.Amrouch et al : *Apprentissage Markovien et Neuronal : cas des caractres amazighes imprims*. In Sitacam'09 Agadir Morocco, 12-13 December 2009, pp : 58-67.
11. B. El Kessab, B. Bouikhalene, M. Fakir, S. Safi : *Reconnaissance des caractres Tifinaghe par l'utilisation des rseaux de neurons multicouches*. In Sitacam'09 Agadir Morocco, 12-13 December 2009, pp : 68-83
12. A. Desai : *Gujarati handwritten numeral optical character reorganization through neural network*. In Pattern Recognition 43-2010, pp : 2582-2589
13. R. El Yachi, K. Moro, M. Fakir, B. Bouikhalene : *Utilisation des moments invariants et la programmation dynamique pour la reconnaissance des caractres Tifinagh*. In Journal of Theoretical and Applied Information Technology, 2010, pp : 61-66
14. A. Kefali, T Sari et M. Sellami : *Evaluation de plusieurs techniques de seuillage d'images de documents arabes anciens*. In 5ieme Symposium Internationnal-IMAGE'2009
15. N. Arica, F. T. Yarman-Vural : *An overview of character recognition focused on off-line handwriting*. In IEEE transactions on systems, man and cybernetics - part C : Applications and reviews, 31(2), 2001, pp : 216-233

16. K. Khurshid, I. Siddiqi, C. Faure, N. Vincent : *Comparison of Niblack inspired Binarization methods for ancient documents*. In 16th International conference on Document Recognition and Retrieval, USA, 2009

17. C. Wolf, J. M. Jolion, and F. Chassaing : *Extraction de texte dans des vidos : le cas de la binarisation*. In 13me Congrs Francophone de Reconnaissance des Formes et Intelligence Artificielle, 2002, pp : 145-152

18. H. BLUM : *A transformation for extracting new descriptions of shape*. In Models for the Perception of Speech and Visual Form, MIT Press, 1967, pp : 362-380

19. Z. Guo and R.W. Hall : *Parallel thinning with two subiteration algorithms. Comm*. ACM, 32(3), March 1989 pp : 359-373

20. M. Hanmandlu, K.R. Murali Mohan, S. Chakraborty, S. Goyal, D. Roy Choudhury : *Unconstrained handwritten character recognition based on fuzzy logic*. In Pattern Recognition 36 (3), 2003, pp : 603-623

21. M. Hanmandlu, M.H.M Yusof, M. Vamsi Krishna : *Off-line signature verification and forgery detection using fuzzy modelling*. In Pattern Recognition 38 (3), 2005, pp : 341-356

22. C. Luh Tan, A. Juntan : *Digit recognition using neural networks*. In Malaysian Journal of Computer Science, 17 (2), 2004, pp : 40-54.

23. M.B. Sukhswami, P. Seetharamulu, A. Pujari : *Recognition of Telugu characters using neural networks*. In International Journal of Neural Systems, 6 (3) ,1995, pp : 317-357.

Kamal Moro
Information processing and telecommunication teams
Faculty of Science and Technology,
Sultan Moulay Slimane University
523, Beni Mellal, Morocco
kamalmoro@hotmail.com

Mohammed Fakir
Information processing and telecommunication teams
Faculty of Science and Technology,
Sultan Moulay Slimane University
523, Beni Mellal, Morocco
fakfad@yahoo.fr

Badr Dine El Kessab
Information processing and telecommunication teams
Faculty of Science and Technology,
Sultan Moulay Slimane University
523, Beni Mellal, Morocco
bade10@hotmail.fr

Belaid Bouikhalene

Information processing and telecommunication teams
Faculty of Science and Technology,
Sultan Moulay Slimane University
523, Beni Mellal, Morocco
bbouikhalene@yahoo.fr

Cherki Daoui
Information processing and telecommunication teams
Faculty of Science and Technology,
Sultan Moulay Slimane University
523, Beni Mellal, Morocco
daouic@hotmail.fr