Recognition of Gujarati Numerals using Hybrid Approach and Neural Networks

Baheti M. J.
Department of Computer Engineering,
SNJB's College of Engineering
Chandwad, Nashik (M.S.) India

Kale K. V.
Department of CS and IT,
Dr. BAM University, Aurangabad (M.S.) India

ABSTRACT

The handwriting recognition is the scheme of converting text symbolized in the spatial form of graphical symbols into its figurative depiction. Handwritten characters have been the most accredited technique of collecting, storing and transmitting information all the way through the centuries. To give the proper ability to the machine it requires studying the image-form of data which forms a special pattern to be interpreted. Designing and building machines that can recognize patterns remains one of the thrust areas in the field of computer sciences. A lot of work has been done in this field, but still the problem is not answered in its full density. A good text recognizer has many commercial and practical applications, e.g. from finding data in digitized book to computerization of any organization, like post office, which involve manual task of interpreting text. In this paper, we have presented a hybrid approach for recognition of Gujarati handwritten numerals using neural networks as classifier and achieved a good recognition rate for noisy numerals.

Keywords
Gujarati numerals, neural networks, handwriting recognition.

1. INTRODUCTION

The handwriting recognition is the scheme of converting text symbolized in the spatial form of graphical symbols into its figurative depiction [1]. Till today, it remains one of the most challenging and exciting areas of research in computer science. In advanced years it has full-fledged into a grown-up stream of science, bringing into being a massive organization of work. Whilst the computer has hugely simplified the process of producing printed documents, the convenience of a pen and paper still makes it the natural medium for many important tasks.

Since centuries, handwritten script has been the most approved method of collecting, storing and transmitting information. This is useful for making digital copies of handwritten documents, and also in many automated processing tasks, such as automatic mail sorting or cheque processing. In automated mail sorting, letters are directed to the correct location by recognition of the handwritten address. Similarly, cheque processing involves recognizing the words making up the cheque amount. This requires making the machines work like humans [2]. For giving machines the human like abilities, Artificial Intelligence, plays a very vital role. Giving machine the power to see, interpret and the ability to read text is one of the major tasks of Artificial Intelligence.

Across India more or less twenty-four official languages are used from corner to corner within a variety of regions of the nation [3-7]. Each language has the diversity representing its eccentricity as well as carves up some of the resemblance with other languages. Apart from Gujarat, Gujarati is spoken and used as official language across globe within various countries like South Africa. Being derived from Devanagari, Gujarati language is written by using Gujarati script. Fig. (**1**) shows the numerals that belongs to Gujarati script.

Fig. 1Gujarati numerals

From Fig 1 it can be observed that Gujarati numerals show several identically similar numerals in Devanagari. This paper deals with the recognition of Gujarati handwritten numeral by hybrid approach using neural networks as classifier.

This paper is organized in following sections; Section 2 describes brief literature survey done for Gujarati script recognition. Section 3 describes algorithm which we have used to implement the paper. Section 4 elaborates the neural network used for classification. Section 5 details the experimental results and successively section 6 concludes the work done.

2. LITERATURE SURVEY

2.1 Skew Detection and Correction

Patel & Desai [8] had assumed that persons writing style remains uniform in most cases hence to speed up the image processing, they extracted middle part of the text to use as input for the radon transform for skew detection purpose. The radon transform computed projections of an image matrix along specified directions. The extracted middle portion of the text was considered as an input image for radon function. To detect the skew the radon transform for input image was computed at angles from 0° to 179°, in 1° increments of angles. The maximum value of R provided by the function [R,xp] = radon(I, theta); where I was input image and theta varied for 0 to 179, helped to detect the skew angle. The radon transform based techniques of skew detection and correction procured results on the skew angle in the range of -20 to +20 degrees.

2.2 Character Compression

Yajnik [9] had proposed an approach of wavelet descriptors (Daubechies D4 wavelet coefficients) for image compression of printed Gujarati letters. The authors scanned the document at 300dpi. Daubechies D4 wavelet transform was applied to character images 256 low-low coefficients were used to create feature vectors. The image was binarized in to 32x32 matrix. This binarized matrix was given as an input to the algorithm in which convolution of binarized image with father wavelet v and mother wavelet u was done followed by down- sampling by a factor of 2 twice. The image was divided in to four sub bands low-low, low high, high-low, high-high coefficients of the size 16x16. Only low-low (approximation) coefficients were considered which captured the core information of the image. Those coefficients were considered as an input to the recognizer (like nearest neighborhood or Neural Network architectures [10-11]) that reported the results up to 75% in compression.

2.3 Segmentation

For extracting lines from the handwritten Gujarati text Patel & Desai [8] have used horizontal projections. Another effort contributed for Gujarati script was by Shah and Sharma [12] in which they used template matching and Fringe distance classifier as distance measure. Initially the sample images were filtered using low pass filter. Then the binarization was done by considering the optimal threshold method. Skew detection and correction is done within 0.05°. They segmented printed characters in terms of lines, words and connected components. By this effort, for connected component recognition rate was 78.34% for upper modifier recognition rate was 50% where as for lower modifier it was 77.55% and for punctuation marks it was 29.6%, cumulative for overall it was 72.3%.

2.4 Printed Character Recognition

Antani & Agnihotri [13] in 1999 have given the primitive effort to Gujarati printed text. The author has created the data sets from scanned images of printed Gujarati text at 100 dpi and from various sites of internet from 15 font families. For training 5 fonts created 10 samples each. The images were scaled up and then scaled down to a fixed size so that all the samples should be of same size i.e. 30x20. For feature extraction the author computed both invariant moments and raw moments. Also image pixel values are used as features creating 30x20= 600 dimensional binary feature space. For classification the author had used two classifiers, K-NN classifier and minimum hamming distance classifier. The best recognition rate was for 1-NN for 600 dimensional binary features space i.e. 67% 1-NN in regular moment space gave 48% while minimum distance classifier had the recognition rate of 39%. The Euclidean minimum distance classifier recognized only 41.33%.

In 2005, Dholakia [14] have presented an algorithm to identify various zones used for Gujarati printed text. In the algorithm they have proposed the use of horizontal and vertical profiles. They have identified these zones by slope of lines created by upper left corner of rectangle created by the boundaries of connected components from line level and not word level the 3 different document images, 20 lines were extracted where 19 were detected with correct zone boundary. The line where it failed was very much skewed.

Dholakia [15] attempted to use wavelet features, GRNN classifier and KNN classifier on the printed Gujarati text of font sizes 11 to 15 with styles regular, bold and italic by finding the confusing sets of the characters. They collected 4173 samples of middle zone glyphs of initial size 32x32 and 16x16 wavelet coefficients have been extracted creating the feature vector. Two sets of the randomly selected glyphs (2802 symbols) were used for training and 1371 symbols were used for testing. Two classifiers GRNN and KNN with Euclidean distance as similarity measure were used producing 97.59 and 96.71 as their respective recognition rates.

Kayasth & Patel [16] proposed a system for recognition of offline computer generated and printed Gujarati characters using continuous GCRHMM for different font sizes of Gujarati characters. The images of all the segmented characters were rescaled into a common height and widths producing a grid with say 24 x 32 pixel-size (i.e. shaped-zones). The pixel density is calculated as binary patterns and therefore a vector is created. However, due to the varying nature of font family, there was dissimilarity between the feature vectors of the same class. A continuous GCRHMM is used for the recognition, yielding classification accuracy described by the earlier table as the basis for the NILKANTH fonts with 72 point size and 48 point size. In our system we extract features from a binary pattern. This feature serves to build a set of reference prototypes for the different classes of the character shapes. Recognition is then achieved by simple matching of a candidate character shape to the pre-built prototypes of all the Gujarati Character set.

Chaudhari [17] have described a system for recognition of offline multi-font computer generated and machine printed Gujarati numerals. Pursued by the pre-processing techniques, they have used correlation based template matching where a numeral is identified by analyzing its shape and comparing its features that distinguish each numeral. The test set used in this experiment was getting a good set of numerals for classification. The numerals used for the experiment were enclosed in a bounding region of a fixed size. Different font families represent the same numeral differently and the correlation between similar numerals was found to vary from font to font. This segmentation based approach was seen for multi-font, size independent numerals.

2.5 Handwritten Character Recognition

In the work [18-19] the author proposed a simple yet robust structural solution for performing character recognition in Gujarati. Pursued by the preprocessing techniques, they used template matching for identifying a character by analyzing its shape and comparing its features that distinguish each character. The Template matching algorithm comprised of Template Classification, Correlation Analysis and Computation of Cross Correlation Coefficient. For each position a correlation coefficient was computed and the corresponding coordinates (pixel accuracy only) were saved. By virtue of the sizes of the template and search windows, a total of forty-nine correlation coefficients were computed to determine the best pattern match location to pixel accuracy for each sampling location. These forty-nine coefficients were actually in the form of a 7×7 matrix which corresponded to the overlap positions of the template upon the window. The maximum coefficient in this sample was assumed to be the best representation for a true match with average overall recognition rate of 71.66 %.

In the reported work of E.Hassan [20], multiple kernels learning (MKL) was used for Gujarati character recognition. The authors applied three different feature representations for symbols obtained after zone wise segmentation of Gujarati text. The MKL based classification was proposed, where the MKL was used for learning optimal combination of different features for classification. In addition MKL based classification results for different features were also found. The multiclass classification was performed in Decision DAG framework. The comparison of results in 1-Vs-1 framework and using KNN classifier was also reported.

Desai [21] collected 300 samples of 300dpi with initial size 90x90 of each numeral and then adjusted the contrast by CLAHE i.e. contrast limited adaptive histogram equalization algorithm considering 8x8 tiles and 0.01 as contrast enhancement constant. The boundaries were then smoothed out by median filter of 3x3 neighborhoods. Image was then reconstructed to the size of 16x16 pixels using nearest neighbor interpolation. For feature extraction four profile vectors were used as an abstracted feature of identification of digit. Five more patterns for each digit are created in both clockwise and anticlockwise directions with the difference of 2degrees each up to 10°. A feed forward back propagation neural network was used for Gujarati numeral classification with 278 sets of various digits. Out of these 278 sets, 11 sets were created by a standard font. From the 265 sets the author recorded the success rate for standard fonts as 71.82%, for handwritten training sets as 91.0% while for testing sets as a score of 81.5% was recorded.

The extension of this work was carried out by maintaining two aspects. One was the information of subdivision of the skeletonized image and second was the aspect ratio of the image before converting it into skeleton. Here pixels information of sub division of images was a structural approach of feature extraction whereas the aspect ratio was a statistical approach. As seen earlier the last form of the image in pre-process phase was the skeletonized image in size of 16 X 16 pixels. This 16 X 16 image was then sub divided into 16 sub images of 4 X 4 pixels. Each of these sub images was then treated as a separate block to find out the total number of on pixels in that block. Thus a vector of 16 such information was created. The aspect ratio of the digits was also considered as one of the features in addition to sixteen features of subdivided images. k-NN classifier for classification of Gujarati handwritten numerals gave accuracy of 96.99% for training set and 92.783% for unseen data [22].

As per the literature survey done for Gujarati script it has been found that a meager amount of work has been done for recognition of handwritten numerals. This motivated us to apply a hybrid approach for recognition of handwritten Gujarati numerals.

3. HYBRID APPROACH

3.1 Handwritten Gujarati Numerals Database

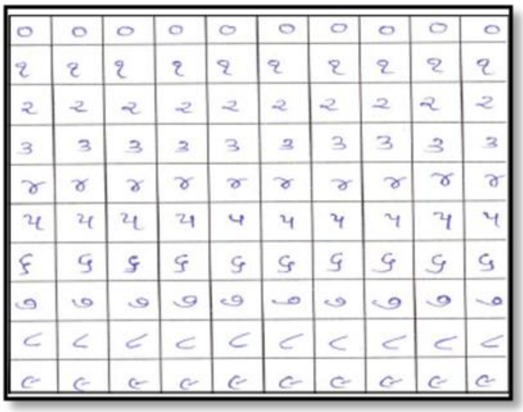

Fig. 2: Specially designed datasheet for Gujarati numerals database collection

It has been reviewed that, at present there is no standardized database available for Gujarati handwritten numerals. For the recognition of Gujarati handwritten numerals the system required to have database, so as the first step the database has been formed. Data was collected from people of different age groups, belonging to different profession, illiterate but knowledge of writing Gujarati irrespective of gender. From such diversity of group, ten samples of each digit from 180 persons was collected on a specially designed datasheet as shown in Fig. (2) for data collection.

3.2 Algorithm for Hybrid Approach

Original image scanned from datasheet as shown in Fig 3 was stored in database. This image from the database was directly converted from rgb to grayscale as shown in Fig 4. The gray scale image was then binarized by considering the optimum threshold as seen in Fig 5. This binarized image is then read across all corners i.e. from left, right, top and bottom. After reading the image the blank pixels are found and the image is cropped so that extra background pixels were removed from the image as shown in Fig 6. The process generated the image with a variety in size. Hence to apply our algorithm we needed to normalize the image size-wise. The cropped image was then resized to 70x50 pixels as shown in Fig 7. The resized image was then partitioned into zones of 7x5 matrix, each zone was of size 10x10 pixels. All the ON pixels are summed up. This sum is converted in the ratio of 100 by subtracting the sum from 100 and dividing it by 100. These values (7x5=35) form the feature vector as shown in Fig 8 for classification.

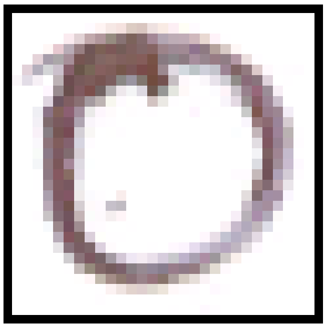

Fig. 3: Original Numeral taken from database

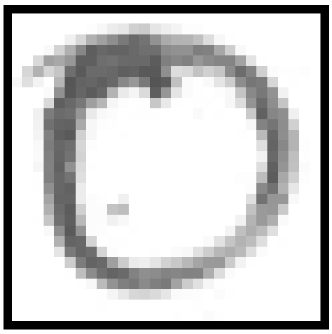

Fig. 4: Numeral from database converted to grayscale

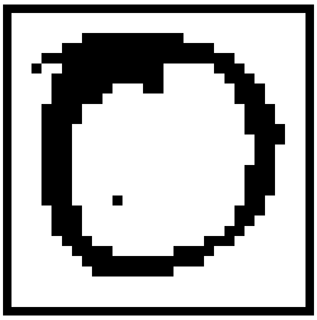

Fig 5 Grayscale image in binarized with optimum threshold

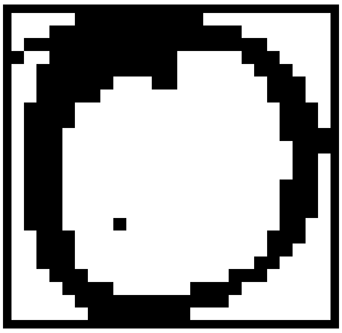

Fig 6 Cropped image after removing extra background pixels

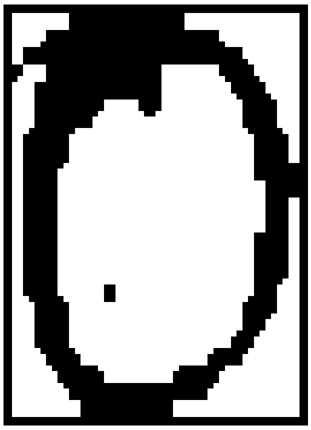

Fig 7 Resized image to size 70x50 pixels

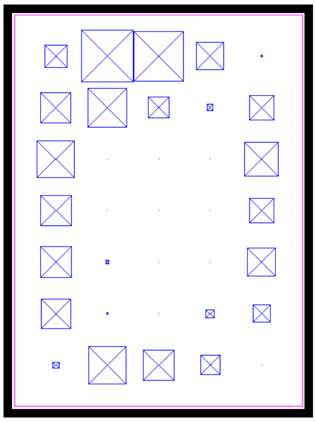

Fig 8 Graphical representation of feature vector created (7x5=35 features)

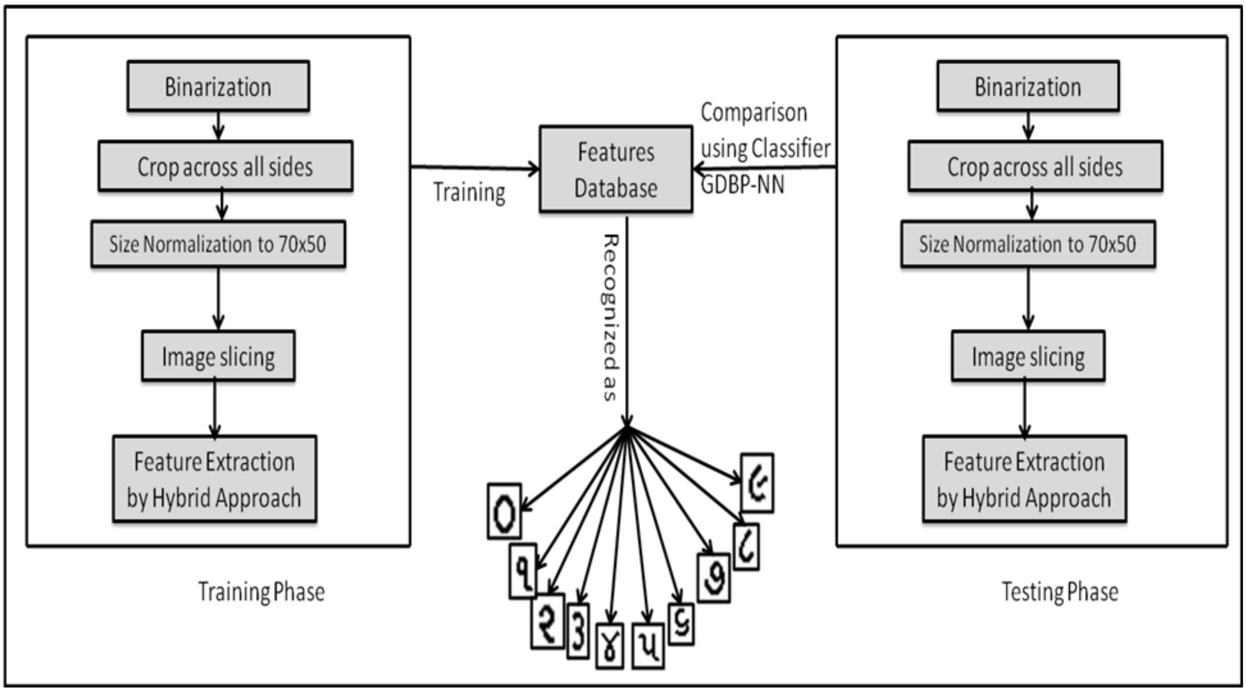

Fig 9: Algorithm employed for volunteer

4. NEURAL NETWORK

Feed-forward neural networks [23], where the data from input to output units is strictly feed forward. The data processing can extend over multiple (layers of) units, but no feedback connections are present, that is, connections extending from outputs of units to inputs of units in the same layer or previous layers. Reinforcement learning type of learning may be considered as an intermediate form of the above two types of learning. Here the learning machine does some action on the environment and gets a feedback response from the environment. The learning system grades its action good (rewarding) or bad (punishable) based on the environmental response and accordingly adjusts its parameters. Generally, parameter adjustment is continued until an equilibrium state occurs, following which there will be no more changes in its parameters. The self organizing neural learning may be categorized under this type of learning. Since our desired outputs must be ranged between 0 to 1, so we have selected log sigmoid as the transfer function for both hidden and output layer. We have used 'Mean Squared Error' (MSE) as performance parameter function. MSE is the average squared error between the network outputs and the target outputs. During training, the weights of the network are iteratively adjusted to minimize the function. We adopt 'Gradient descent back propagation' as a learning algorithm. The algorithm updates weights according to gradient descent momentum and adaptive learning rate. The values we have used to set training parameters are learning rate to 0.01, momentum factor to 0.8 and performance goal to 0.01. The concluding results were commenting either as recognized or misrecognized numeral. The optimum confusion matrix with correctly recognized and misrecognized numerals is created.

5. EXPERIMENTAL RESULTS

The feature set is obtained on the basis of algorithm shown in Fig 9, for 1800 samples from the database we have created. Then feature set shown in table 1 is treated with the Gradient Descent Back propagation with Adaptive leaning rate Neural Network classifier. After the training the network with 600 numerals, the confusion matrix as shown in table 2 is drawn for the 1200 different test samples

Feature Set

Table 1 Feature set (optimum) derived after employing the algorithm shown in Fig 9

0	1	2	3	4	5	6	7	8	9
0.85	0.32	0.62	0.15	0.75	0.18	0	0	0	0
0.65	0.68	0.71	0.11	0.16	0.61	0	0	0.23	0.35
0.69	0.9	0.62	0	0	0.66	0.31	0.6	0.56	0.82
0	0.78	0.71	0.05	0.39	0.52	0.83	0.38	0	0.71
0	0.1	0.1	0.78	0	0.05	0.35	0	0	0.29
0.58	0.52	0.31	0.64	0.35	0.69	0	0	0.05	0.31
0.66	0.22	0	0.96	0.19	0.2	0	0.2	0.81	0.47
0.9	0	0	0.55	0	0	0.61	0.78	0.12	0.06
0.36	0.66	0.5	0.66	0.7	0	0.04	0.01	0	0
0	0.56	0.37	0.73	0	0	0.56	0	0.28	0.58
0	0	0	0	0.47	0.77	0.33	0.01	0.45	0.62
0	0	0	0.21	0.1	0.03	0	0.78	0.35	0.01
0.51	0	0.25	0.99	0.09	0	0.6	0.18	0	0
0.5	0.42	0.81	0.94	0.69	0	0.04	0	0.39	0
0	0.85	0.05	0.07	0	0.46	0.53	0	0.56	0.39
0	0	0	0	0.32	0.34	0.54	0.47	0.75	0.53
0	0	0	0.15	0.73	0.75	0	0.43	0.02	0
0.53	0.18	0.45	0.82	0.78	0.4	0.28	0	0.54	0
0.5	0.88	0.6	0.89	0.61	0.51	0.9	0	0.44	0
0	0.49	0.65	0.07	0	0.73	0.88	0	0	0.44
0	0.7	0	0	0	0	0.68	0.92	0.67	0.53
0	0.77	0	0.39	0.09	0.03	0	0.08	0.5	0
0.88	0.79	0	0.5	0.16	0.24	0	0	0.38	0
0.24	0.49	0	0.57	0.53	0.33	0	0	0	0.01
0	0	0.73	0.29	0	0.6	0.52	0	0	0.61
0	0.36	0.35	0	0	0	0.62	0.92	0.88	0.59
0	0.96	0	0.31	0	0	0.2	0.2	0.22	0.12
1	0.72	0	0.73	0	0	0	0	0	0
0	0.3	0.03	0.7	0.69	0.06	0.35	0.08	0.21	0.39
0	0.03	0.7	0.3	0.01	0.84	0.52	0.1	0.2	0.38
0	0	0.38	0	0	0	0.09	0.24	0.3	0.12
0	0.29	0.71	0	0	0.1	0.84	0.72	0.88	0.69
0.85	0.71	0.68	0.47	0	0.7	0.81	0.9	0.8	0.76
0.65	0.87	0.65	0.86	0.5	0.68	0.57	0.95	0.77	0.46
0.55	0.74	0.35	0.08	0.81	0.41	0.04	0.82	0.21	0.01

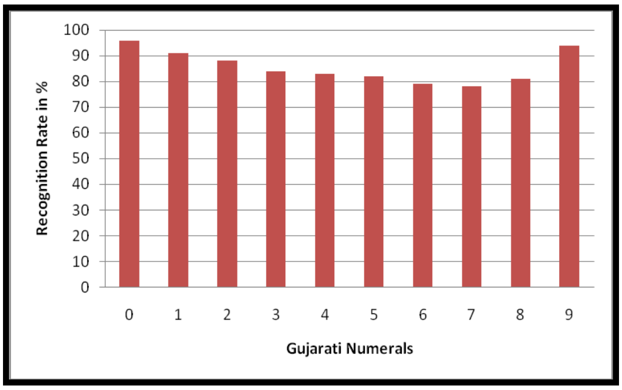

Figure 10 Recognition rates procured by hybrid approach using Neural Network classifier

Table 2 Confusion Matrix for hybrid approach using Neural Network as classifier

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**115**	5	0	0	0	0	0	0	0	0
1	4	**109**	6	1	0	0	0	0	0	0
2	0	3	**106**	10	1	0	0	0	0	0
3	0	1	5	**101**	11	2	0	0	0	0
4	0	0	1	6	**99**	11	3	0	0	0
5	0	0	1	0	6	**98**	13	2	0	0
6	0	0	0	1	1	7	**95**	15	1	0
7	0	1	0	1	0	0	8	**93**	17	0
8	1	2	0	5	1	0	0	8	**97**	6
9	0	1	0	1	0	0	0	0	5	**113**

For numerals 0 and 1 as shown in figure 10 we have reported the recognition rate as 96% and 91% respectively. For numerals 3, 4, 5, 6, 7 and 8 less recognition rate is found to be 84%, 83%, 82%, 79%, 78% and 81%. Numeral 2 is found to procure recognition rate of 88% which is found to be much better as compared to others. Numerals 9 reported to procure very good recognition rate of 94%. The overall recognition rate is **86%.** It has shown good results for numerals 0, 1 and 9.

6. CONCLUSION

It was found that it was possible to enhance performance of system if a character is divided in a systematic manner and features of each divided part are used in recognition system. our hybrid approach has shown good results for numeral 9 to have maximum recognition at 94%. These results were compared with [18-19, 21-22] and found better for noisy numerals.

7. REFERENCES

[1] L. Harmon. "Automatic recognition of print and script" Proc. IEE, 60:1165-1176, 1972.

[2] Line Eikvil, "Optical Character Recognition", citeseer.ist.psu.edu/142042.html. Accessed on 01 July 2011

[3] http//en.wikipedia.org Accessed 15 June 2007

[4] Lawrance Lo. at http://www.ancientscripts.com/ Accessed 15 June 2007

[5] http//ccat.sas.upenn.edu Accessed 15 June 2007

[6] http//languages.iloveindia.com Accessed 15 June 2007

[7] http//india.mapsofindia.com Accessed 15 June 2007

[8] Patel C N and Desai A A (2009) Skew detection and text line extraction for handwritten Gujarati text , Prajna J Pure Appl Sci 17: 099-103

[9] Yajnik A and Singh D (2010) Feature Extraction (Image Compression) of printed Gujarati and Amharic letters using Discrete Wavelet Transform, S-JPSET, 1(1):5-10.

[10] S. Rama Mohan, Archit Yajnik (2005): "Gujarati Numeral Recognition Using Wavelets and Neural Network" Indian International Conference on Artificial Intelligence, pp. 397-406.

[11] Archit Yajnik, S. Rama Mohan, "Identification of Gujarati characters using wavelets and neural networks" Artificial Intelligence and Soft Computing 2006, ACTA Press, pp. 150- 155.

[12] S K Shah and A Sharma Design and Implementation of Optical Character Recognition System to Recognize Gujarati Script using Template Matching IE(I) Journal–ET Vol.86 pgs. 44-49 2006

[13] Antani S and Agnihotri L (1999) Gujarati Character Recognition. Proc. 5th ICDAR, pp.418-422.

[14] Dholakia J, Negi A, S Rama Mohan (2005) Zone Identification in the Printed Gujarati Text. Proc of 8th ICDAR p272-276.

[15] Dholakia J, Yajnik A, Negi A (2007) Wavelet Feature Based Confusion Character Sets for Gujarati Script. ICCIMA pp366-371.

[16] Kayasth M and Patel B (2009) Offline typed Gujarati Character recognition. NJSIT pp73-82

[17] Chaudhari S and Gulati R M (2010) A font size independent OCR for machine printed Gujarati numerals. 3(1):70-78

[18] Prasad, J.R. ; Kulkarni, U.V. ; Prasad, R.S. (2009) Template Matching Algorithm for Gujarati Character Recognition In Proc. Of 2nd International Conference on Emerging Trends in Engineering and Technology (ICETET), pp 263 - 268

[19] Prasad, J.R.; Kulkarni, U.V.; Prasad, R.S. (2009) Offline Handwritten Character Recognition of Gujarati script using Pattern Matching In Proc. Of 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, pp 611 - 615

[20] E. Hassan, S. Chaudhury, M. Gopal, and J. Dholakia, "Use of MKL as symbol classifier for Gujarati character recognition", in Proc. Document Analysis Systems, 2010, pp.255-262.

[21] Desai A A (2010) Gujarati handwritten numeral optical character reorganization through neural network. Pattern Recognition 43: 2582-2589.

[22] Desai A. A. (2010) Handwritten Gujarati Numeral Optical Character Recognition using Hybrid Feature Extraction Technique. IPCV pp 733-739

[23] http://www.learnartificialneuralnetworks.com/#Mathematical Accessed on July 2011