

Research Article

Recognition of South Indian Language Numerals Using Minimum Distance Classifier

F. Jennifer Alin and K.N. Saravanan

Department of Computer Science, Christ University, Bangalore, 560029, India

Abstract: The recognition of handwritten and machine printed numerals have received extensive attention in pattern recognition. Indian handwritten character identification is slightly demanding because of the presence of character modifiers and also many documents consists of numerals which are both printed and handwritten. There are plenty of recognition systems developed for numeral recognition. Our aim is to develop a single framework or recognition system which could distinguish between numerals of various languages. Mixture of printed and handwritten character numerals in Indian context normally show up in documents such as application forms, postal mail, office letters and so forth. Generally, Humans can distinguish the difference in characters but to make a machine understand is challenging. Hence in order to make the machine understand numerals of different languages, this study has been carried out. The recognition rate of 87.96, 87.93, 86.93 and 89.16%, respectively has been achieved for Kannada, Telugu, Tamil and Malayalam numerals respectively.

Keywords: Character modifiers, handwritten character, pattern recognition, postal mail, solitary document

INTRODUCTION

In this fast moving technological world where machine has gained a certain level of intelligence to recognize and to act accordingly, character recognition system has received extensive attention. Character recognition system is a mechanism by which a text entered into a computer can be easily identified. Rolling over the years it has become evident that online character recognition has gained a major growth than the offline character recognition system. Handwriting recognition is a wide concept which can be categorised into the following based on the way in which it is fed into the computer. Inputs to the computer can be fed either by a touch screen, a pen and stylus or via scanned documents.

When a computer processes handwriting obtained via a stylus or a touch sensitive device, it is referred to as the online handwriting recognition. It is so known because the processing of the handwriting is done simultaneously with the process of entering the input. It basically traces the movement over the touch sensitive device or the movement of the pen/stylus while recognizing the input.

Manually written script identification is the capability of a system to accept and translate comprehensible manually written text from sources such as records, paper reports, photos, touch screens and different gadgets. The road block in handwritten recognition is the huge variability and distortion of

patterns. To overcome these problems and to give an accurate result is the main objective of this study. Handwritten recognition has innumerable applications which comprise of making a path for blind in reading, bank cheques, interpretation of ID number, archive reading, mail sorting, postal address recognition and automatic number plate recognition.

Numerous methods have been proposed by the analysts towards the identification of non-Indian numerals compared to the works done on Indian numerals. India is a multi-lingual country consists of twenty-two official languages. In this research work, four main South Indian languages such as Kannada, Telugu, Tamil and Malayalam have been considered. Almost all South Indian Language numerals are curve in shapes, the demanding part is the distinction between the similar shaped components. Although the structure of a character is similar in all four languages, there is a major difference in the writing style of one individual to another. These distinct features lead to increasing recognition complexity and degree of recognition accuracy. There has been a major elevation in the OCR for the English alphabet system. The OCR that might have a high accuracy rate for writing from an individual may not prove to be efficient for the writing of another individual. It all depends on the training provided to the system. Same numeral might be written in various shapes and on the other hand at least two numerals of a script may take comparable shapes as shown in Fig. 1.

Corresponding Author: F. Jennifer Alin, Department of Computer Science, Christ University, Bangalore, 560029, Karnataka, India

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

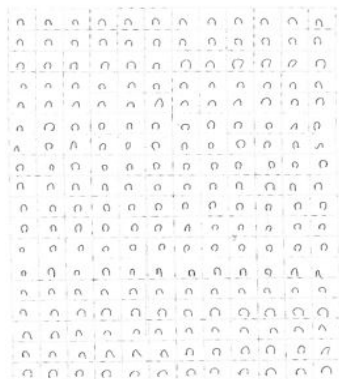


Fig. 1: Number one in Kannada and Telugu

In recent years, researchers have developed many recognition models to identify the hand written texts and have been discussed below.

To extract features in Hindi numerals, Sanossian (1998) has utilized three primitive elements in a segment, namely, boundary distance, pixel density and line distance from centroid.

Nagabhushan and Pai (1999) have done a complete analysis and description of OCR Research on Dravidian Scripts. A number of methods of recognition has been carried out in native languages like English, Chinese, Japanese and Arabic. Latin, Chinese, English language recognition methods are excellently reviewed by Pal and Chaudhuri (2004). Extensive work on Indian scripts can be discovered only on few scripts, whereas the work on handwritten South Indian numeral recognition is still in infant stage. Identification of manually written South Indian characters is a tricky task due to the unconstrained shapes, variety in composing style and so forth. In spite of the fact that there are twelve different scripts in India, just a couple explore papers have been published on handwritten numeral recognition and these papers are primarily on Devanagari, Bangla and Oriya scripts.

Candès *et al.* (2006) used feature extraction method like Fast Discrete Curvelet Transform based on Wrapping and Unequipped FFT Transform.

Sharma *et al.* (2006) put forth the directional tie code and zoning methodologies summing up to 100 features for manually written Kannada numeral identification and accomplished sensibly high precision, however the proposed algorithm reported to have high time complexity due to large dimension feature set.

Using Fourier descriptors recognizable proof of separated Devanagari handwritten numerals have been proposed by Rajput and Mali (2010). Hanmandlu *et al.* (2007) concluded that the algorithms and techniques developed for English cannot be connected to Indian dialects on the ground that the vast majority of the Indian scripts are distinguished by the nearness of mastras (or, character modifiers) notwithstanding principle characters.

A proficient method was proposed by Hossain *et al.* (2011) which involves partitioning the image by

figuring the projection of every area. Many papers have been published on Odiya handwritten numerals among which Majhi *et al.* (2011) have built up a classifier for recognition. Features like Curvature and Image inclination have been utilized and the image has been normalized (64*64 pixels of height and width). Utilizing these distinct features, the number of features acquired is 2,592 and then diminished to 64, by means of Principal Component Analysis (PCA) technique.

Another technique for recognising manually written numerals in Hindi in light of auxiliary descriptors was proposed by Elnagar *et al.* (1997). This strategy includes filtering the transcribed numeral and normalizing it to 30-by-30 pixels and then thinning it. Later, elements like strokes and cavity have been extricated and represented syntactically. At last, the obtained representation of the element is coordinated against a stored set of syntactic representation models.

Chaudhuri and Pal (1998) carried out a complete analysis on Bangla OCR system. They extracted shape characteristics, zonal information of each numeral and then applied structural-feature-based tree classifier followed by template matching. Sanjeev Kunte and Sudhaker Samuel (2007) used feature extraction methods like Hu's invariant moments and Zernike moments and have obtained a considerable recognition accuracy.

Rajashekaradhy and Vanaja Ranjan (2008) proposed zone centroid and image centroid based point include extraction framework for isolated Kannada numerals identification. Though the recognition system proposed by Dinesh Acharya *et al.* (2007) have announced the exactness of 90.50%, it additionally requires extra thinning as per outlined in the literatures. Nagabhushan *et al.* (2003) implemented methods that have employed fuzzy features to recognise Kannada vowels based on invariant moments. Saravanan and Anita (2016) put forth the method of extracting features using partial derivatives. Five major steps such as zone based representation, representing in terms of density matrix, obtaining numeric value, calculation of absolute value, final feature value were carried out and obtained a reasonable recognition accuracy.

Purkait and Chanda (2010) have proposed a method where organising component (line) along the major, minor, vertical and horizontal directions have been utilized to get four distinct images and also morphological opening and shutting operation are performed on pre-handled image to obtain more features. Machine learning concepts like Principal Component Analysis has been utilized by other researchers to obtain standard element focuses for decomposing the numeral into its primitives. Mozaffari *et al.* (2005) obtained recognition rate of 94.44% using nearest neighbour classifier. Feature extraction methods involved decomposing the numeral into its primitives.

A new technique for zoning description was put forth by Impedovo *et al.* (2006) in which best classification results have been accomplished by apportioning the example image into $M = 9$ zones, if

Table 1: Few similar shaped characters

Similar numerals	Handwritten numerals of different languages	
Kannada numeral 6 and 9	೬	೯
Kannada numeral 7 and Tamil numeral 2	೭	೨
Telugu numeral 4 and Kannada 4	౪	೪

handwritten numeral digits were considered on the training sets. It has likewise been expressed that the ideal zoning still beats the customary zoning technique on the testing sets too. Rahiman and Rajasree (2009) carried out a complete survey of Optical Character Recognition on South Indian Scripts. They have discussed about the various recognition methodologies and techniques involved in Pattern recognition field. The comparison of results gives a brief insight of various recognition classifiers used by numerous researchers. Dhandra *et al.* (2010) in his study have extracted spatial elements and considered a component vector of length 13 for manually written solely for Kannada numeral identification.

Various approaches have been proposed by the researchers towards the identification of non-Indian numerals. One of the most commonly utilized methodologies depends on neural network. Couple of different researchers utilized basic approach, where each example class is characterized by structural description and the recognition is carried out according to structural likenesses. Statistical approach is also applied to numeral recognition. Among others, Support vector machines, Fourier and Wavelet description, Fuzzy rules, tolerant rough set, free automatic scheme for unconstrained off-line are reported in the literatures.

From the literature survey, it is apparent that still handwritten character identification is carried out for single dialect and little work is done for bilingual, trilingual and multi-lingual character identification. This has propelled to plan a solitary identification algorithm for handwritten South Indian numerals.

EXISTING SYSTEM

The current system in classifying or recognising handwritten characters involves the recognition of characters pertaining to one dedicated language. But there is no dedicated method to recognize numerals of different languages. Since those languages are constrained to only those regions. There are different types of systems which make use of various techniques and methodologies as discussed above. The essential goal is to accelerate the process of character recognition in archive processing. Thus, the current frameworks can handle enormous archives with-in less time. But not all the existing systems make use of same methodology and hence differs in their own time and space complexities.

Proposed system: Every person varies in their own way of handwriting styles. There are mechanisms like optical character recognition which recognizes the

character using optics. Sometimes it becomes difficult for the machine to understand the same object written by various people. In order to surpass this, different dataset by various people is collected. The collected dataset is checked such a way that if it matches with the same object written by other people. The handwritten documents are scanned and treated as images as it further fed into the machine. Objects which are not matching are excluded and only those objects with highest results of similarity are taken into consideration. These objects are fed to the machine so that it will be able to understand the same object written in various style. Techniques like dimensionality reduction and segmentation are used to segregate the images. In this proposed research work, generic features are selected in such a way that it is common to all the four languages. Since the languages taken into consideration come under Dravidian languages there is considerable similarities between numerals of one language and numerals of the other. The common numeral in all the four languages is Zero. Few similar shaped numerals are shown in Table 1. Steps involved in Training and Testing the system is represented in Fig. 2 and 3 respectively.

Data collection and pre processing: Handwritten numerals will be written in various sizes and shapes due to the variation in writing styles. Composing style in the script is from left to right. The data utilized for the experiment were gathered from various people. There was no restriction or guidance in the writing style thus we noticed that the data sets contain assortments of writing styles. Writers were chosen from schools,

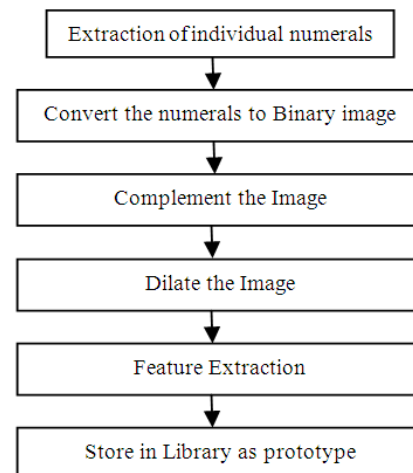


Fig. 2: Flow diagram of training system

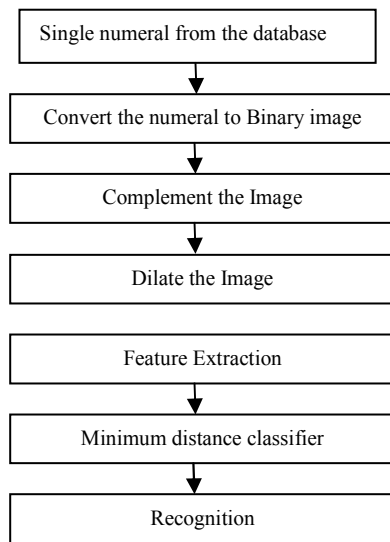


Fig. 3: Flow diagram of testing system



Fig. 4: Scanned image of Kannada numeral four

colleges, professionals, business men, house wives and the reason for numeral gathering was not disclosed to them. Kannada, Tamil, Telugu, Malayalam datasets have been collected from the people who can write those languages. Data of different sizes and slants is also included in the database. Writers were given with an A4 sheet consists of box with dimensions of size 1.3 cm X 1.3 cm and they were asked to write the numerals inside the given box. The gathered documents were scanned using HP flatbed scanner which yield low noise good quality grey scale images. It is guaranteed that the skew introduced amid the archive scanning is irrelevant and consequently disregarded. Figure 4 shows scanned image of numeral four.

A total of 8800 handwritten numerals of four different languages which are Kannada, Telugu, Tamil and Malayalam were collected and stored as data set. Out of those samples of each class we use 70% samples to train the recognition classifier and the remaining 30% to test the recognition classifier. The collected data underwent a certain pre-processing steps.

The step stone in pre-processing is to binarize the numeral images so that the numeral images have pixel values 0 and 1. A thresholding function has been performed on scanned grey scale images. The noise in

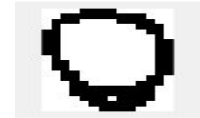


Fig. 5: Numeral zero after pre-processing

the image is expelled using morphological erode and dilate operations. To bring consistency among all the numerals, the numerals are normalized and also the images are complemented. Resizing of images has also been done so that there is a uniformity.

Figure 5 represents pre-processed image of numeral Zero. The scanned image of Kannada numeral four is shown in the Fig. 5.

The scanned A4 size sheet is then segmented to obtain the individual numerals which is in the box of size 1.3 cm×1.3 cm and is shown in Fig. 6.

Table 2 shows the number of Training and testing data for all ten numerals from each language taken for experiment.

Implementation: As mentioned in the earlier sections the primary stages for the implementation like data collection, pre-processing is met. The next stage is to decide upon which and what features has to be extracted. Since we are dealing with four different languages, we want to develop a system which is common to all the four languages. Hence the features decided to extract are Aspect ratio, Horizontal profile, Vertical profile, Left diagonal profile, Right diagonal profile. The samples taken for implementation are categorised into two sections Training and Testing respectively. 70% of the data are treated as Training data and the remaining 30% are taken for Testing. Architecture diagram of the Training and Testing mechanism is shown in Fig. 7.

Algorithm: Training numerals:

- Input: Image of a numeral from the database.
- Output: Store Pattern vector in the Library as a prototype.
- Step 1: Convert the accepted input to binary image.
- Step 2: Perform other pre-processing steps: image dilation, image complement.
- Step 3: Based on the round off mean aspect ratio value of a language, resize the enhanced image.
- Step 4: Calculate the zone division for a language based on round off mean aspect ratio.
- Step 5: Extract the features in each zone of a numeral image.
- Step 6: Repeat steps 1 to 5 and apply minimum distance classifier and store the class prototype in the library.

End of the Algorithm.

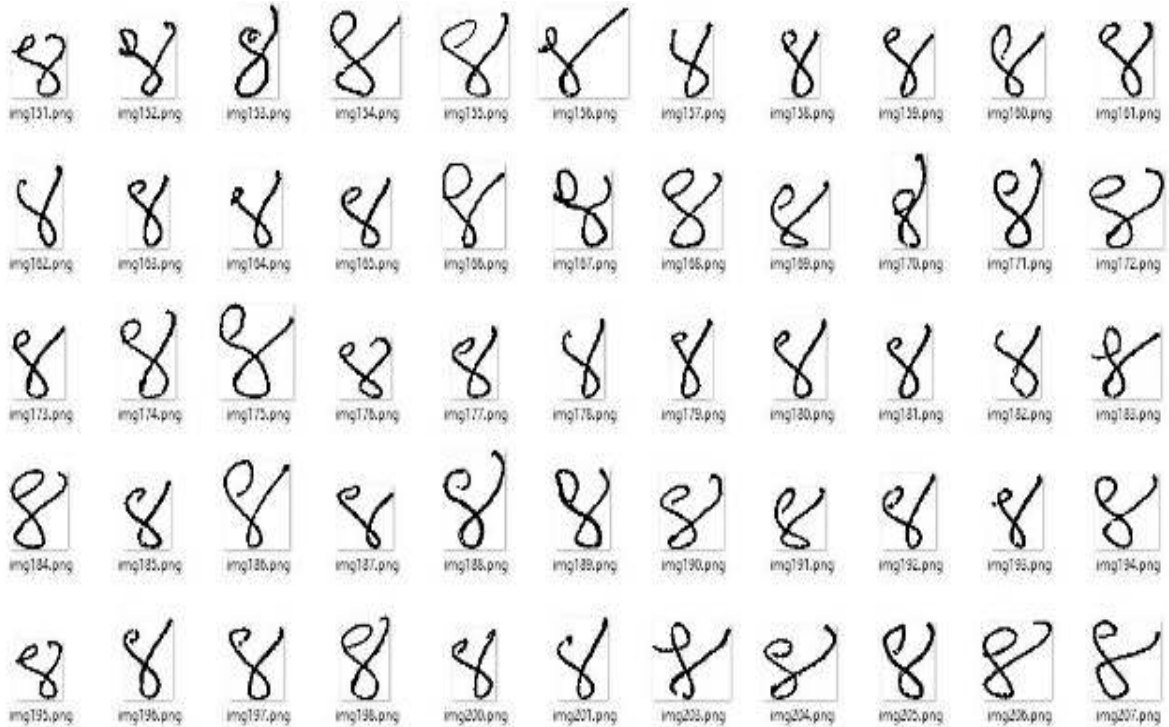


Fig. 6: Segmented image of Kannada numeral four

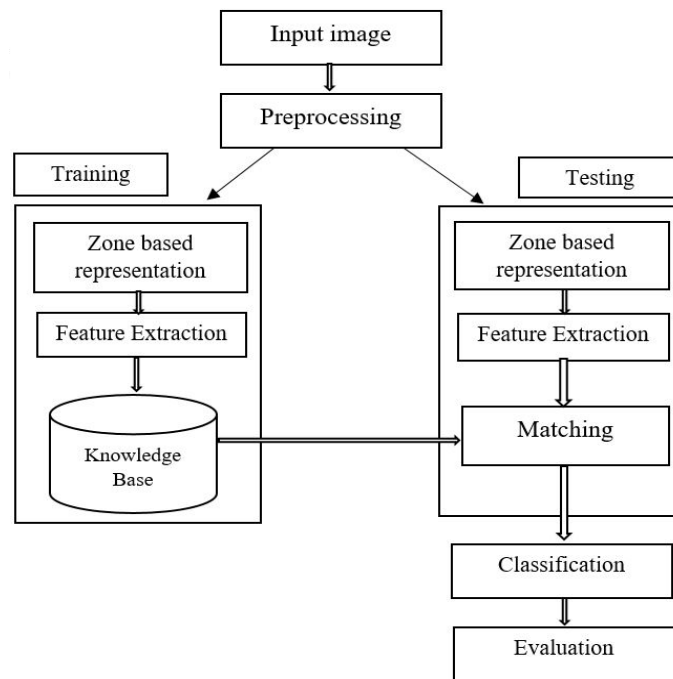


Fig. 7: Training and testing system architecture

Table 2: Training and testing data

Languages	Testing data	Training data
Kannada	1600	600
Telugu	1600	600
Tamil	1600	600
Malayalam	1600	600

Algorithm: Testing Numerals:

Input: Single image of a numeral for a language from the database.
 Output: Classification of a numeral based on the language.

Table 3: Representation of round off mean aspect ratio values in terms of zones

Language	Mean aspect ratio	Round off mean aspect ratio	Representation in terms of zones	
			Y axis	X axis
Kannada	0.887895	0.9	9	10
Telugu	1.01016	1	10	10
Tamil	0.823232	0.8	8	10
Malayalam	0.840331	0.8	8	10



Fig. 8: Zone based representation of Kannada numeral one

- Step 1: Convert the accepted input to binary image.
- Step 2: Perform other pre-processing steps: image dilation, image complement.
- Step 3: Based on the round off mean aspect ratio value of a language, resize the enhanced image.
- Step 4: Calculate the zone division for a language based on round off mean aspect ratio.
- Step 5: Extract the features in each zone of a numeral image.
- Step 6: Apply the distance metric between the generated features and the features stored in the library for a particular language and assign the numeral class to the nearest feature vector in the library.

End of the Algorithm.

Representation of round off mean aspect ratio value for binary numeral images: When the primary stages of image pre-processing and segmentation are performed, image representation and feature selection of numerals play a vital role in a recognition system (Table 3). Zone based representation in terms of round off mean aspect ratio value has been proposed in our research work.

Using the proposed representation scheme, the number of zones has been identified along both axes and the zone size would be defined for all the numerals in a language. Once again feature extraction methods are applied separately to all the zones of an image. One of the benefits of the proposed zoning representation scheme is that it could handle various handwritten styles of different writers without influencing the state of the numerals. Zone based division of Kannada numeral one is shown in Fig. 8.

RESULTS AND DISCUSSION

The proposed methodology does not confine only as an idea or theory knowledge but also can go live into the real world. Unlike other systems which is dedicated for a particular language, the proposed algorithm has been implemented using MATLAB where a single framework has been designed and developed to

recognise the numerals of various languages. Character recognition rate of several other methodologies discussed earlier have found a significantly high accuracy. The performance evaluation of proposed method compared with others is different and difficult, because the experimental conditions like number of features, classifier, training and testing dataset size are different. To achieve a higher recognition rate, higher number of training and testing samples should be used.

Sample handwritten numerals has been tested for the languages Kannada, Telugu, Tamil and Malayalam and the results acquired are encouraging and are discussed in the following sub sections.

Kannada language numerals: A total of 220 samples for each numeral was collected. From which 160 were trained and 60 were taken for testing. Amongst the 60 testing samples, results obtained for each numeral are shown in the Table 4. As a whole for language Kannada we obtained the accuracy of 87.96%. This shows that with more testing samples and few other constraints in pre-processing will bring us out the accurate results.

Telugu language numerals: A total of 220 samples for each numeral was collected. From which 160 were trained and 60 were taken for testing. Amongst the 60 testing samples, results obtained for each numeral are shown in the Table 5. As a whole for language Telugu we obtained the accuracy of 87.93%. Kannada and Telugu numerals results vary with minute percentage because most of the numerals are same in both the language.

Tamil language numerals: A total of 220 samples for each numeral was collected. From which 160 were trained and 60 were taken for testing. Amongst the 60 testing samples, results obtained for each numeral are shown in the Table 6. As a whole for language Tamil we obtained the accuracy of 86.93%.

Malayalam language numerals: A total of 220 samples for each numeral was collected. From which 160 were trained and 60 were taken for testing. Amongst the 60 testing samples, results obtained for each numeral are shown in the Table 7. As a whole for language Malayalam we obtained the accuracy of 89.16%. Kannada and Telugu numerals results vary with minute percentage because most of the numerals are same in both the language.

Table 4: Confusion matrix for Kannada Handwritten numerals

Kannada numeral	೦	೧	೨	೩	೪	೫	೬	೭	೮	೯
೦	54	4	0	0	0	0	0	0	2	0
೧	4	50	0	0	3	0	0	2	0	1
೨	0	0	53	0	1	1	0	1	3	1
೩	0	0	3	54	0	0	1	1	0	1
೪	0	0	1	3	51	3	0	0	2	0
೫	0	0	0	0	0	57	0	0	3	0
೬	0	1	0	2	0	2	49	0	0	5
೭	0	0	2	0	0	0	3	53	0	2
೮	0	0	0	2	1	0	1	0	56	0
೯	0	0	1	0	0	0	6	1	0	52

Table 5: Confusion matrix for Telugu Handwritten numerals

Telugu numerals	೦	೧	೨	೩	೪	೫	೬	೭	೮	೯
೦	54	4	0	0	0	0	0	0	2	0
೧	4	50	0	0	3	0	0	2	0	1
೨	0	0	53	0	1	1	0	1	3	1
೩	0	0	3	49	0	0	1	5	0	1
೪	0	0	1	3	51	3	0	0	2	0
೫	0	0	0	0	0	57	0	0	3	0
೬	0	1	0	2	0	2	49	0	0	5
೭	0	1	2	0	0	0	3	58	0	2
೮	0	0	0	2	1	0	1	0	56	0
೯	0	0	1	0	0	0	6	1	0	52

Table 6: Confusion matrix for Tamil handwritten numerals

Tamil numerals	೦	೧	೨	೩	೪	೫	೬	೭	೮	೯
೦	54	4	0	0	0	0	0	0	2	0
೧	0	48	0	0	3	0	7	0	0	2
೨	0	2	52	0	2	0	1	0	0	3
೩	0	0	1	54	0	0	5	0	0	0
೪	1	0	4	0	51	0	0	1	3	0
೫	0	2	0	1	0	52	0	4	0	1
೬	0	0	2	0	7	0	49	0	2	0
೭	0	1	5	0	0	0	0	54	0	0
೮	0	2	0	1	0	3	0	0	53	1
೯	0	0	1	0	1	0	2	0	0	56

LIMITATIONS AND FUTURE ENHANCEMENT

The handwritten numerals should not be broken or fragmented, they must be single connected. If the numeral is broken, the algorithm treats the broken numerals as different number of numerals based upon

the number of fragments. This is one of the most important limitations.

The proposed method and the work done so far in this area can be enhanced in future such a way that the system is capable of recognising any numeral of any language in any part of the world. Features are extracted in a manner that it is common to all the four

Table 7: Confusion matrix for Malayalam handwritten numerals

	൦	൧	൨	൩	൪	൫	൬	൭	൮	൯
Malayalam numerals	54	4	0	0	0	0	0	0	2	0
൧	0	50	7	0	1	0	2	0	0	0
൨	0	4	54	0	1	0	1	0	0	0
൩	0	0	0	49	0	0	5	0	0	6
൪	0	0	2	0	55	0	0	2	0	1
൫	1	1	0	0	0	57	0	0	1	0
൬	0	0	0	5	0	0	52	0	0	3
൭	0	2	0	0	3	0	0	53	0	2
൮	0	1	0	0	0	0	0	0	59	0
൯	0	0	0	5	0	0	2	0	0	53

languages considered for experiment. Hence a single model is sufficient to recognise the numerals of various languages.

CONCLUSION

In this study we have presented a zone based feature extraction system using minimum distance classifier for the recognition of single connected component numerals. Using this methodology, we have achieved good results even when certain pre-processing steps like slant removing, smoothing are not considered. Minimum distance classifier has been used for classification. The recognition rate of 87.96, 87.93, 86.93 and 89.16%, respectively has been achieved for Kannada, Telugu, Tamil and Malayalam numerals respectively.

ACKNOWLEDGMENT

We are grateful to Dr Surendra Umesh Kulkarni and Prof Joy Paulose of Christ University, Bangalore, India for their constant support and encouragement throughout this research work.

CONFLICT OF INTEREST

Publishing a research paper in a Scopus indexed journal is part of curriculum and hence started with this study which later became area of interest. This interest strived me up to explore in depth in the field of "Numeral Recognition".

REFERENCES

Candès, E., L. Demanet, D. Donoho and L. Ying, 2006. Fast discrete curvelet transforms. *Multiscale Model. Sim.*, 5(3): 861-899.

Chaudhuri, B.B. and U. Pal, 1998. A complete printed bangla OCR system. *Pattern Recogn.*, 31(5): 531-549.

Dhandra, B.V., M. Hangarge and G. Mukarambi, 2010. Spatial features for handwritten Kannada and English character recognition. *Int. J. Comput. Appl.*, Special Issue on RTIPPR-2010, 3: 146-151.

Dinesh Acharya, U., N.V. Subba Reddy and Krishnamoorthi, 2007. Isolated handwritten Kannada numeral recognition using structural feature and K-means cluster. *Proceeding of the IISN*, pp: 125-129.

Elmagar, A., F. Al-Kharousi and S. Harous, 1997. Recognition of handwritten Hindi numerals using structural descriptors. *Proceeding of the IEEE International Conference on Systems, Man and Cybernetics, Computational Cybernetics and Simulation*. Orlando, FL, 2: 983 988.

Hanmandlu, M., A.V. Nath, A.C. Mishra and V.K. Madasu, 2007. Fuzzy model based recognition of handwritten Hindi numerals using bacterial foraging. *Proceeding of the 6th IEEE/ACIS International Conference on Computer and Information Science*. Melbourne, Australia.

Hossain, M.Z., M.A. Amin and H. Yan, 2011. Rapid feature extraction for Bangla handwritten digit recognition. *Proceeding of the International Conference on Machine Learning and Cybernetics (ICMLC, 2011)*. Guilin, 4:1832-1837.

Impedovo, S., R. Modugno and G. Pirlo, 2006. A feedback-based framework for continuous evaluation of e-learning activities. *Proceeding of the 3rd WSEAS/IASME International Conference on Engineering Education*. Vouliagmeni, Greece, pp: 258-262.

Majhi, B., J. Satpathy and M. Rout, 2011. Efficient recognition of odiya numerals using low complexity neural classifier. *Proceeding of the International Conference on Energy, Automation and Signal (ICEAS, 2011)*, pp: 140-143.

Mozaffari, S., K. Faez and M. Ziaratban, 2005. Structural decomposition and statistical description of Farsi/Arabic handwritten numeric characters. *Proceeding of the 8th International Conference on Document Analysis and Recognition*, 1: 237-241.

Nagabhushan, P. and R.M. Pai, 1999. Modified region decomposition method and optimal depth decision tree in the recognition of non-uniform sized characters – An experimentation with Kannada characters. *Pattern Recogn. Lett.*, 20(14): 1467-1475.

- Nagabhushan, P., S.A. Angadi and B.S. Anami, 2003. A fuzzy statistical approach to kannada vowel recognition based on invariant moments. Proceeding of the 2nd National Conference on Document Analysis and Recognition (NCDAR), pp: 275-285.
- Pal, U. and B.B. Chaudhuri, 2004. Indian script character recognition: A survey. *Pattern Recogn.*, 37(9): 1887-1899.
- Purkait, P. and B. Chanda, 2010. Off-line recognition of hand-written Bengali numerals using morphological features. Proceeding of the 12th International Conference on Frontiers in Handwriting Recognition. Kolkata, India, pp: 363-368.
- Rahiman, M.A. and M.S. Rajasree, 2009. A detailed study and analysis of OCR research in south Indian scripts. Proceeding of the International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom'09), pp: 31-38.
- Rajashekaradhy, S.V. and P. Vanaja Ranjan, 2008. Neural network based handwritten numeral recognition of Kannada and Telugu scripts. Proceeding of the IEEE Region 10 Conference on TENCN 2008-2008, pp: 1-5.
- Rajput, G.G. and S.M. Mali, 2010. Fourier descriptor based isolated Marathi Handwritten numeral recognition. *Int. J. Comput. Appl.*, 3(4): 9-13.
- Sanjeev Kunte, R. and R.D. Sudhaker Samuel, 2007. A simple and efficient optical character recognition system for basic symbols in printed Kannada text. *Sadhana*, 32(5): 521-533.
- Sanossian, H., 1998. Feature extraction technique for Hindi numerals. Proceeding of the IEEE Signal Processing Society Workshop Neural Networks for Signal Processing VIII. Cambridge, pp: 524-530.
- Saravanan, K.N. and R. Anitha, 2016. Recognition of multi-lingual handwritten numerals using partial derivatives. *Res. J. Appl. Sci. Eng. Technol.*, 12(1): 27-36.
- Sharma, N., U. Pal and F. Kimura, 2006. Recognition of handwritten Kannada numerals. Proceeding of the 9th International Conference on Information Technology (ICIT'06), pp: 133-136.