

# A Survey on Tamil Handwritten Character Recognition using OCR Techniques

<sup>1</sup>M. Antony Robert Raj, <sup>2</sup>Dr.S.Abirami

Department of Information and Science and Technology  
Anna University, Chennai – 600 025  
Email: antorobert@gmail.com, abirami\_mr@yahoo.com

## ***Abstract***

*In today's fast growing technology, digital recognitions are playing wide role and providing more scope to perform research in OCR techniques. Recognition of Tamil handwritten scripts is complicated compared to other western language scripts. However, many researchers have provided real-time solutions for offline Tamil character recognition also. Offline Tamil handwritten documents recognition still offers many motivating challenges to researchers. Current research offers many solutions on Tamil handwritten documents recognition even then reasonable accuracy and performance has not been achieved. This paper analyses the various approaches and challenges concerning offline Tamil handwritten character recognition.*

## ***Keywords***

*OCR, Pre-processing, Image Extraction and classification*

## **1. INTRODUCTION**

With emergence of the digital content, need for the development of a high performance OCR engine has become essential. OCR research works have been undertaken by several researchers which aim at developing a high performance OCR engine.

The idea behind an OCR is to identify and analyse a document image by dividing the page into line elements, further sub-dividing into words, and then into characters. These characters are compared with image patterns to predict the probable characters. Recognition of characters can be done either from printed documents or from handwritten documents. Handwritten document recognition can be done offline or online. Offline character recognition is more complicated than online. In particular, Tamil handwritten OCR is more complicated than other related works. This is because Tamil letters have more angles and modifiers. Additionally, Tamil script contains large number of character sets. A total of 247 characters; consisting of 216 compound characters, 18 consonants, 12 vowels and one special character. Challenges that researches face during recognition process are due to the curves in the characters, number of strokes and holes, sliding characters, differing writing styles so on.

The steps involved in character recognition comprise pre-processing, segmentation, feature extraction and classification. Three types of features, namely Statistical, Structural and Hybrid can be analyzed there.

Researchers have come up with many approaches for the character recognition, however, many of them have surveyed in the paper. Apart from that, challenges and issues still prevailing in this even for future research has also been surveyed in this paper.

This paper is organized as follows; Pre-processing techniques are surveyed in section 2. Section 3 illustrates the various segmentation techniques available for offline character recognition. Feature extraction methods are explained in section 4. Section 5 explains the various classification approaches are available. Section 6 illustrates a comprehensive study of existing approaches and the final section provides an overall view of these approaches, the scope of our research in this area and conclusion.

## **2. PRE-PROCESSING**

There are numerous tasks to be completed before performing character recognition. A handwritten document must be scanned and converted into a suitable format for processing. Pre-processing consists of a few types of sub processes to clean the document image and make it appropriate to carry the recognition process accurately. The sub processes which get involved in pre-processing are illustrated below:

- Binarization
- Noise reduction
- Normalization
- Skew correction, thinning and slant removal

### **2.1 Binarization**

Binarization is a method of transforming a gray scale image into a black and white image through Thresholding [17][9]. Another approach, Otsu's method may be used to perform histogram based thresholding [16] [14] to get binarized image automatically. Otsu's method has been extended for multi level thresholding, called Multi Otsu method [24]. Normally, most researchers use thresholding concepts to extract the foreground image from background image [18][21][4]. In this method, the threshold value is fixed by taking any value between two foreground gray code images. Histogram based thresholding approach can also be used to convert a gray-scale image into a two tone image. In contrast, Adaptive Binarization method can also be used to identify the local gray value contrast of Image. This will help to extract text information from low quality documents. Another approach named Two-Level Global Binarization Technique represents the output using global thresholding technique [24].

### **2.2 Noise Removal**

Digital images are prone to many types of noises. Noise in a document image is due to poorly photocopied pages. Median Filtering [18], Wiener Filtering method [19] and morphological operations can be performed to remove noise [16]. Median filters are used to replace the intensity of the character image [24], Where as Gaussian filters can be used to smoothing the image [10].

### **2.3 Normalization**

Normalization is the process of converting a random sized image into a standard size. The Roi-Extraction [20] method is used to get the single structural element from the image. Bicubic interpolation [16], linear size normalization [14] and Java Image Class [12] normalization techniques could be used for standard sized images. In many works, input image are normalized

to a size 50 \* 50 after finding the bounding box of each hand written image [25][11][17] for razing processing.

## **2.4 Skew correction, Thinning and Slant removal**

Thinning is a pre-process which results in single pixel width image to recognize the handwritten character easily. It is applied repeatedly leaving only pixel-wide linear representations of the image characters. Cumulative scalar product (CSP) [18] of windows text block with the Gabor filters has been used for thinning purpose. Morphology based thinning algorithm [25] [11] and other Thinning algorithms [8] [3] [17] has also been used for better symbol representation and to thin the character images. Skeletonization is the process of shedding off a pattern to as many pixels as possible without affecting the general shape of the pattern. Hilditch's algorithm has also been used for removing unwanted pixels from the images (skeletonization) [3] [5] [10] [1]. Skew is inevitably introduced into the incoming document image during document scanning. Normalization [15], Fourier Spectrum [23] techniques are used for correction of the slant, angle stroke, width and vertical scaling.

## **3. SEGMENTATION**

Segmentation is a process which is used to split the document images into lines, words and characters. Segmentation of hand-written documents is more complex than type-written documents. Histogram profiles and connected component analysis [20] [6] are used for line segmentation. In segmentation process, paragraph space has been checked for identifying paragraphs. Histogram of the image has been used to detect the width of the horizontal lines [2]. Spatial space detection technique has been used for word segmentation. Histogram methods has been used to detect the both width of words [25] and also to convert the image to glyph [21].

The vertical histogram profile method [16] [17] has been used to find spacing within the lines to identify the word boundary. Region probe Algorithm [8] has been used to get individual characters from the image. Segmentation is categorized based on different stages as Projection based, Smearing, Grouping, Hough-based, graph-based and Cut Text Minimization (CTM). Modified cross counting Technique [10], Histogram profile [20] and connected component analysis are also found in the survey to handle the line character segmentation problem.

## **4. FEATURE EXTRACTION**

The feature extraction techniques can be broadly grouped into three class namely statistical features, structural features and the hybrid features. A statistical technique uses quantitative measurements for feature extraction, whereas structural techniques use qualitative measurements for feature extraction. In hybrid approach, these two techniques are combined and used for recognition.

### **4.1 Structural Technique**

Scale Invariant feature transform (SIFE) [20] is used to transfer the character image into a set of local features. Using this approach, 128 dimensions of SIFE features (Interesting points) are identified from the character image. An image is converted to a two tone image, and then converted into frame. The frame point obtained from the frame will possess the vectors. The Normalized Feature Vector (NFV) [22] obtains the prototype from vector, i.e. lines and arcs.

## 4.2 Statistical Technique

In the Zone based method, the normalized characters are divided into non interleaving zones. Pixel density is calculated for each zone. This is then used to represent the features [16] [12] [11] [13]. The height and width of character pixels are counted using the encoding Binary variation approach [4]. Once the top level of row and width is reached, the process halts. A binary flag is set as per the approach and features extracted from that. Images are divided into nine non overlapping blocks of equal size using the Gabor channel method [6]. This gives 24 responses for blocks passing through each channel. Mean and standard derivations are calculated and used as features.

All images are scaled to the same height and width using Bilinear Interpolation Technique [6][5]. Unwanted portions are corrected using the Sobel edge detection algorithm [6]. The Roi-Extraction approach [8] uses morphological closing on image resulting in closed image. This approach casts off the leftmost, rightmost, uppermost and bottommost block out of the limit. The Features Normalized x-y coordinates; first and second derivatives, curvature, aspect, curliness and liness are derived from Time domain features. Features for the frequency domain are computed along with stroke using a sliding hamming window. The real and imaginary part of lowest coefficient is added to the feature vector [9]. Structural features such as end point, fork point, holes, length, shape and curvature of individual stroke are derived using octal graph [7].

The “eight-neighbour” adjustment method is used for tracing the boundaries of the images. The approach scans until it finds the boundary of binary image. Then, the Fourier descriptor [24] is used to find the coefficient and obtain the total number of boundaries. This number of invariant descriptors is given as input to a neural network for further classification.

## 4.3 Hybrid Technique

Hough Transform [1] has been used to detect the horizontal and vertical lines. They have been analyzing branch and position using another simple algorithm. Another technique Bilinear interpolation [5] has been used to extract the features such as slant and strip. A feature extraction technique Zone based hybrid approach [12], which has been used to extract the zone centroid and Image centroid based distance metric features.

## 5. CLASSIFICATION

The Extracted features are given as the input to the Classification process. A bag-of-keypoints extracted from the feature extraction approaches are used for classification. There are some approaches are used for classify the character features in the existing systems such as K- Nearest Neighbour approach, Fuzzy system, Neural network, Discriminate classifier, Unsupervised classifier and so on.

K-Nearest Neighbour approach [20] used as classifier to recognize the character sets and better accuracy. The result of the recognition is obtained using the Histogram Equalization. In Self organizing map approach [2][6][25] the weight of each node are calculated using the Euclidean distance method [1]. Then, the input characters are compared with all vectors whose weights were calculated through SOM. If they match, the output will be given as a recognized vector. This is called as Best Matching Unit (BMU). The global features are used if this approach fails to classify [21].

In the Multilayer perception for two hidden layer approach [11], a neural network is designed with all weights mapped to a random number between 1 and -1. The characters are compared

with the network output to find the best recognition result. Two hidden layers used for better performance. If the hidden layers are increased then the performance will be reduced [19][3][24].

In some works, Hidden Markov Model has been used to support handwritten characters recognition model [18][9]. It provides the maximum probable result. In the Nearest Neighbour classifier approach [12][11] features of training sample are computed and stored. These features are used for testing samples. The input samples are compared with the stored samples. In some works, Support Vector machine a binary classifier, where the features are divided into two classes along the hyper plane of minimum margin between two classes has been used. Here margin refers to the distance between hyper plane and closest data points. The outcome is based on the data points, (ie) at the margin; if the margin is maximum, then the hyper plane will be divided into sub plane. If the problem of classifying instances into more than two classes (multi class problem) occurs, use the following algorithms, one-against-all and one-against-one [14][17] [21].

Fuzzy logic approaches have also been used for identifying pattern primitives and labelled of two tone image, then it will be classified as a prototype characters [23][22]. This finds the upper and lower confidential limits of the characters using statistical classifier [5] and they are stored as classified result. In Hierarchical neural network [10] approach two sets are used. The first set is trained to classify multi groups. In second set, each group is further divided into multiple groups and from them, the final result is classified

## 6. COMPREHENSIVE STUDY

The table 1 shows the comprehensive study of which has been made on the different OCR's available for offline character recognition

S. No & Paper No	Pre-processing	Thinning and Skew Detection	Segmentation	Feature Selection	Feature Extraction	Classification
1[20]	<p>Morphological operation are used to achieve Noise removal.</p> <p>Bicubic interpolation is used for standard sized image.</p> <p>Morphological gradient to find character boundaries</p>	-	<p>Histogram profile and connected component analysis for line character segmentation</p>	any interesting character points	<p>Scale Invariant Feature Transform (SIFT) method for Feature Extraction</p> <p>K-Means clustering to create a code book for each characters (two features condition Nearest Neighbour, Centroid condition)</p> <p>CB obtained from LBG algorithm.</p>	<p>Bag - of - keypoints to count the number of patches</p> <p>K-Nearest Neighbour to recognition Scale Invariant Feature Transform(SIFE) transfer character image to set of local frame.</p>
2 [2]	<p>Binarization: (Thresholding)High peaks analogous to white background, smaller peaks corresponding to fore ground</p>	<p>Skew detection checks the angle of orientation between + or - 15 degree</p>	<p>Line (Imaginary) segment and word and characters segment</p>	<p>Character height, width, no. of horizontal and vertical lines and curves, no. of circles and slope lines, image centroid and special dots.</p>	<p>Extract the Features as a vector</p>	<p>Kohonen's Self Organizing Feature Map (SOFM) technique used for Classification</p> <p>Analyse node to find weight and get BMU (Best Matching Unit), analysis of neighbourhood of BMU to recognize characters</p>

3 [19]	Scanned data encapsulated in Glyphs	-	-	character height, width	-	Multilayer perception(MLP) learning algorithm for two hidden layers with back propagation used for final character identification.  Hyperbolic tangent function used for activation purpose.
4 [18]	Median filter used to replace the intensity filter(noise).  Thresholding method used for Binarization	-	Line segmentation used to segment the characters.	Cover vertical projection profile, word profile, background -to-ink transitions	Profile Features (vertical features and word profile) approach used for feature extraction.	Hidden Markov Model(HMM) method used to recognize character.  recognize the word using Baum-Welch or forward and backward algorithm
5[21]	Thresholding method used for Binarization	Thinning algorithm used to thin the characters  Hilditch's algorithm is used for skeletonization.	Spatial space detection technique.  histogram method used to convert the image to glyph	character height, width, no. of horizontal and vertical lines (long and short),horizontal and vertical curves, circles, slope lines, image centroid and dots	Output of the segmentation part (Image glyph) is subjected to feature extraction procedure.	Support Vector Machine (SVM) used for classification (vapnik's structural). Type 1 SVM- self organizing map used to minimize the errors. neural classification algorithm and Radial- Basis-Function Networks , Hybrid Neuro fuzzy systems used for reduce the recognition problems. RCS Algorithm used for better recognition results.
6 [16]	Gabor Thresholding and Otsu Thresholding methods(global) are used for Binarization	Hilditch algorithm (Thinning algorithms) is used for skeletonization	Horizontal and vertical profile method is used for Segmentation (line and characters). bilinear interpolation technique used for normalization	Zone will capture pixel variation	Zone based approach is used for Feature extraction	Support vector machine (SVM) method is used for classification. Radial basis function (RBF) kernel used for support to SVM to divide the line
7 [4]	Threshold technique used for preprocessing, color image to gray scale. Normalization done using Java Image Class	-	-	row wise, column wise and diagonal wise selection	Encoding binary variation method used for extract the features. Then comparing trained text and tested image for recognize the characters	Comparison technique

8 [14]	Otsu's method used to perform histogram based thresholding automatically. Also this method is extended for Mutil level (Multi ostu method)	Part of the image need to filtered using Gabor filter	-	-	Thinned image passed through the Gabor channels.	Features are extracted and fed in to the trained support vector machines (SVM). Multi class problem solved against one against all algorithm
9 [12]	Input image is normalized to size 50* 50 after find the bounding box of each hand written image	Morphology based thinning algorithm for better symbol representation	-	character height,width, no. of horizontal and vertical lines an curves, no. of circles and slope lines	Zone-base Hybrid approach is proposed for feature extraction.	Nearest Neighbour Classifier (NNC) for classification and recognition - large scale pattern matching. Support vector machine (SVM) - structural risk minimization by Vapnik Chervonenkies (VC) dimensions. SVM- two class- ONO(one against one), ONA(one against all). but ONA used in SVM classification. To separate the problem they used Gaussian kernel algorithms.
10 [6]	-	-	Horizontal projection profile method is used for Segmentation (upper line and lower line)	-	Images scaled in to height and width using bilinear interpolation technique. Unwanted portion corrected using Sobel edge detection (Sobel mask)	1. Preliminary Classification - Crux and Exhaustive character. 2. Self-organizing feature maps (SOP) algorithm (Kohonen SOM ) used for Classification
11 [25]	-	Fourier Spectrum concept is used in Skew Detection.	Spatial space detection technique used for segmentation. Histogram used to detect the width of words	character height, width, no. of horizontal and vertical lines (long and short), horizontal and vertical curves, circles, slope lines, image centroid and dots	Output of the segmentation part (Image glyph) subjected to feature extraction procedure.	SOM(Self-organizing map) - Kohonon's SOM used for classification. Unicode mapping technique used to recognize characters
12 [8]	Histogram Equalization & Gabor Filtering (Contrast value, Convolving Image), Roi-Extraction (Get Single Structural element)	-	Region probe Algorithm (Get Individual characters)	leftmost, rightmost, uppermost and bottommost blocks	Roi Extraction	Algorithm Fusion - Hidden Markov model(HMM) and Support Vector Machine (SVM) used to classify characters.  If both are wrong then RBF(Radial Basis Function Neural Network) used to improve accuracy

13 [9]	<p>Binarization: (Thresholding) High peak analogous to white background, smaller peak corresponding to the foreground</p> <p>Median Filtering &amp; Wiener Filtering method used for noise removal</p>	Cumulative scalar product (CSP) of windows text block with the Gabor filters	Line (Imaginary) segment and word and characters segment	Curve, Line, Aspect(height and width) curliness (deviation), Stroke	<p>Time Domain Features - (Normalize the x-y Coordinates, Normalize the first derivatives, Normalize the second derivatives, Curvature, Aspect, Curliness and liness ) and Frequency Domin Features - (Sliding Hamming Window - Stroke and Discrete Cosine Transform (DCT) - evaluate on windows sequence)</p>	<p>Hidden Markov model(HMM) - Training done using the Baum-Welch re-estimation procedure. Character computed using HMM forward Algorithms)</p>
14 [11]	<p>Input image is normalized to size 50* 50 after finding bounding box of each hand written image</p>	morphology based thinning algorithm for better symbol representation	-	shape	<p>Zone based hybrid approach - Zone centroid and Image centroid based Distance metric feature extraction system, hybrid feature extraction system technique are used for feature extraction</p>	<p>Nearest neighbour and Feed forward back propagation neural network classifiers are used for subsequent classification and recognition purpose.</p>
15 [23]	<p>Histogram based threshold approach is used for converting gray to tone image. Polygonal approximation approach used to obtain Skeleton from image. Fuzzy functions for identify the pattern.</p>	-	-	edge, distances		<p>NFA(Normalized Feature vector) used to identify characters using Fuzzy logic</p>

Table 2 list outs the accuracy obtained by the various OCR and their appreciation and limitation obtained in character recognition.

S.No	Appreciations	Accuracy	Limitation	Font size, Input format and output format
1[20]	Achieved in experiments using six thousand training and two thousand testing images of 20 selected characters	87%	Achieved result for 20 selected characters. problem to recognize the abnormal writing and similar shaped characters	Image is an input, Identified SIFE Features, the generate the Code book (CB)
2 [2]	Sentence sample from 5 persons but first 8 letters only used. This also reduced the amount of time and processing power needed to run the experiment.	Not Specified (100% was obtainable)	Recognised done only for 8 characters, Accuracy is not obtained properly.	TIF, JPG or GIF image is segmented into paragraph, paragraph to line, line to words , word to image glyph
3 [19]	Taken 5 different samples from different handwritten for five different script (Tamil, Urdu, Telugu, Devnagri and English) - 10 characters have taken from each languages for recognize.	Single hidden layer - 96.53, Double hidden layer - 94%	Considered 10 characters for recognise, Five samples of each letters were collected for input.	-
4 [18]	100 people (age group 14 to 40) are made to write four pages of text in Tamil languages. Lists of ten words have taken from each of the four documents for testing. 40 search words written by each writers form a text set. This set have take for recognition	Test - 90% Sample - 80.7%	Tested only 40 words.	Document
5[21]	25 users selected from 40 for sample and another 15 for Test Set. Skew detection check for the angle orientation between +or- 15 degree.	Sample set - 100%, Test Set - 97%, Speed .1 Sec	Minimal set of samples (top three letters), problems to recognize the abnormal writing, similar shaped characters and joined letters.	Tif, Jpg or BMP

6 [16]	Result shows that the algorithm works well for all the characters	Avg. 82.04% (62.8 % for 3 characters, 98.9 % for 12 characters, all 34 characters - 82.04 )	34 characters 6048 samples, Recognition errors based on abnormal errors, similar shaped characters	32X32, 48X48 and 64X64 size image
7 [4]	10 characters are involved in the recognition	Not given	Consider only 10 characters, Not specified about accuracy(not given 100%)	Image is an input, Text is an output
8 [14]	Six frequently used Tamil font characters with different style printed are tested. Initially 216 instances were taken for training and gradually incremented	SVM - 92.5%, Six font variation(86% - 100%)	No Handwritten character are in the process. Six frequently used Tamil fonts	regular, bold, bold italic, italic. Printed text as input
9 [12]	Collected 2000(1000 testing, 1000 training) Tamil numeral samples from 200 people for recognize, 10 sample characters are considered. Give good accuracy rate for Kannada	SVM - 93.9%, NN - 94.9%,	10 characters only considered for recognition	Input format is BMP
10 [6]	Considered 67 symbols, 200 training and 800 testing samples and 100 text lines are used. SOM model is to capture the invariant features of the Tamil Scripts	98.50%	Recognised done only for 67 characters. Problems to recognize the abnormal writing, similar shaped characters and joined letters.	250 X 250 pixels INPUT FONT
11 [25]	Very few Samples	Not Specified	Very few Samples	TIF, JPG or GIF input format
12 [8]	Applied for 250 "Thirukural"	92% to 94 % (50 kurals - 94.1%, 100 Kurals - 94.3, 150 Kurals - 92.5%, 200 Kurals - 94.2%, 250 Kurals - 92.5%)	Accuracy rate, Handwritten character not specified	-
13 [9]	Tested Many Hand written of different individuals, Olaichuvadi, Scanned and machine printed documents, more compatible for other Indian scripts	93% to 98% (25 words - 94.41%, 50 words - 96.52% , 100 words - 98.48% ,150 words - 95.07%, 200 words - 93%)	.But Given example is few hand written scanned documents only, Not considering the Continuous writing (If there are no space between the characters) and sliding characters	-

14 [11]	collected 1500 Tamil numeral samples from 150 people for recognize, 10 sample characters are considered	96%	10 characters only considered for recognition	Input format is BMP
15 [23]	Tested and recognize 2500 sample of 28 characters. Recognized 30 degree sliding characters also	76-94%	Tested 28 characters only. problems to more curves and joined letters.	Input size 15 pixels height and width 30 pixels

The following challenges are identified from various papers which may provide more lively interest to the researchers to carry out the research work in this area. They are

- Recognized limited number of Tamil characters (40 Characters has been identified)
- Minimum number of testing samples from various handwritten documents has been considered
- Difficult to identify the abnormal writing and similar shaped characters
- Font variation and sliding letters – not dealt
- Low accuracy rate in recognition
- Old handwritten character set which are available in palm leaf, historical documents, mother documents and stone carvings are not dealt until now.
- Distorted characters which are on the damaged palm leaf and damaged documents are not considered till now.

## 7. CONCLUSION

A lot of research work exists in the survey for Tamil Handwritten recognition. However, there is no standard solution to identify all Tamil characters with reasonable accuracy. Various methods have been used in each phase of the recognition process, whereas each approach provides solution only for few character sets. Challenges still prevails in the recognition of normal as well as abnormal writing, slanting characters, similar shaped characters, joined characters, curves and so on during recognition process.

In this paper, we have projected various aspects of each phase of the offline Tamil character recognition process. Researchers have used minimal character set. Coverage is not given for different writing styles and font size issues. The following key challenges can be further explored in the future by researchers

- Curves in the Tamil characters
- Very large character set
- Complex letter structure
- Significant variation in writing styles due to complex letter structures
- Increased number of stroke and holes
- Mixed words (English and Tamil)
- Extreme font variability
- Difficulties faced in viewing angles, shadows and unique fonts

**REFERENCES**

- [1] Akshay Apte and Harshad Gado, "Tamil character recognition using structural features" ,2010
- [2] Banumathi P and Nasira G.M, "Handwritten Tamil Character Recognition using Artificial neural networks", International Conference on Process Automation, Control and Computing (PACC), page(s): 1 – 5, 2011
- [3] Bhattacharya U, Ghosh S.K and Parui S.K, "A Two Stage Recognition Scheme for Handwritten Tamil Characters", Ninth International Conference on Document Analysis and Recognition, Vol: 1 page(s): 511 – 515, 2007
- [4] Bremananth R and Prakash A, "Tamil Numerals Identification", International Conference on Advances in Recent Technologies in Communication and Computing, page(s): 620 – 622, 2009
- [5] Hewavitharana S and Fernando H.C, "A Two Stage Classification Approach to Tamil Handwritten Recognition", Tamil Internet, California, USA, 2002
- [6] Indra Gandhi R and Iyakutti K, "An attempt to Recognize Handwritten Tamil Character using Kohonen SOM", Int. J. of Advance d Networking and Applications, Volume: 01 Issue: 03 Pages: 188-192, 2009
- [7] Jagadeesh Kannan R and Prabhakar R, "An improved Handwritten Tamil Character Recognition System using Octal Graph", Int. J. of Computer Science, ISSN 1549-3636, Vol 4 (7): 509-516, 2008
- [8] Jagadeesh Kumar R and Prabhakar R, "Accuracy Augmentation of Tamil OCR Using Algorithm Fusion", Int. J. of Computer Science and Network Security, VOL.8 No.5, May 2008
- [9] Jagadeesh Kumar R, Prabhakar R and Suresh R.M, "Off-line Cursive Handwritten Tamil Characters Recognition", International Conference on Security Technology, page(s): 159 – 164, 2008
- [10] Paulpandian T and Ganapathy V, "Translation and scale Invariant Recognition of Handwritten Tamil characters using Hierarchical Neural Networks", Circuits and Systems, IEEE Int. Sym. , vol.4, 2439 – 2441, 1993
- [11] Rajashekararadhya S.V and Vanaja Ranjan P, "Efficient Zone based Feature Extraction Algorithm for Handwritten Numeral Recognition of Four Popular south Indian Scripts". Int. J. of Theoretical and Applied Information Technology, pages: 1171 – 1181, 2008
- [12] Rajashekararadhya S.V and Vanaja Ranjan P, "Zone-Based Hybrid Feature Extraction Algorithm for Handwritten Numeral Recognition of two popular Indian Script", World Congress on Nature & Biologically Inspired Computing, page(s): 526 – 530, 2009.
- [13] Rajashekararadhya S.V, Vanaja Ranjan P and Manjunath Aradhya V. N, "Isolated Handwritten Kannada and Tamil Numeral Recognition a Novel Approach", First International Conference on Emerging Trends in Engineering and Technology, page(s): 1192 – 1195, 2008
- [14] Ramanathan R, Ponmathavan S, Thaneshwaran L, Arun.S.Nair, and Valliappan N, "Tamil font Recognition Using Gabor and Support vector machines", International Conference on Advances in Computing, Control, & Telecommunication Technologies, page(s): 613 – 615, 2009
- [15] Sarveswaran K and Ratnaweera , "An Adaptive Technique for Handwritten Tamil Character Recognition", International Conference on Intelligent and Advanced Systems, page(s): 151 – 156, 2007
- [16] Shanthi N and Duraiswami K, "A Novel SVM -based Handwritten Tamil character recognition system", springer, Pattern Analysis & Applications, Vol-13, No. 2, 173-180,2010

- [17] Shanthi N and Duraiswami K, "Performance Comparison of Different Image size for Recognizing unconstrained Handwritten Tamil character using SVM", Journal of Computer Science Vol-3 (9): Page(3) 760-764, 2007
- [18] Sigappi A.N, Palanivel S and Ramalingam V, "Handwritten Document Retrieval System for Tamil Language", Int. J of Computer Application, Vol-31, 2011
- [19] Stuti Asthana, Farha Haneef and Rakesh K Bhujade, "Handwritten Multiscript Numeral Recognition using Artificial Neural Networks", Int. J. of Soft Computing and Engineering ISSN: 2231-2307, Volume-1, Issue-1, March 2011
- [20] Subashini A and Kodikara N.D , " A Novel SIFE-based Codebook Generation for Handwritten Tamil character Recognition" , 6th IEEE Int. Conf. on Industrial and Information Systems (ICIIS), Page(s): 261 – 264, 2011
- [21] Suresh Kumar C and Ravichandran T, "Handwritten Tamil Character Recognition using RCS algorithms", Int. J. of Computer Applications, (0975 – 8887) Volume 8– No.8, October 2010
- [22] Suresh R.M, Arumugam S and Ganesan L, "Fuzzy Approach to Recognize Handwritten Tamil Characters", Third International Conference, Proc. on Computational Intelligence and Multimedia Applications, page(s): 459 – 463, 1999
- [23] Suresh R.M, "Printed and Handwritten Tamil Characters Recognition using Fuzzy Technique", Pro. of the Int. Multi Conference of Engineers and Computer Scientists, Vol I, 19-2, March, 2008
- [24] Sutha J and RamaRaj N, "Neural network based offline Tamil handwritten character recognition System" , International Conference on Conference on Computational Intelligence and Multimedia Vol : 2, page(s): 446 – 450, 2007
- [25] Venkatesh J and Suresh Kumar C, "Tamil Handwritten Character Recognition using Kohonon's Self Organizing Map", Int. J. of Computer Science and Network Security, VOL.9 No.12, December 2009