Vol. 5 No. 4 April 2018 ISSN: 2321-788X UGC Approval No: 43960 Impact Factor: 3.025

DEVELOPMENT OF NOVEL TECHNIQUES TO RECOGNIZE TAMIL LANGUAGE CHARACTERS

Article Particulars: Received: 05.03.2018 Accepted: 15.04.2018 Published: 28.04.2018

Mrs. L. Suriya Kala, MCA., M.Phil.,
Research Scholar, Mother Teresa Women's University
Kodaikanal, Tamil Nadu, India

Abstract

Despite the widespread utilization of PCs in record preparing, hand-printed archives assume a significant job in our life. The machine investigation of hand-printed records has extraordinary hugeness in this period of paperless workplaces. The machine investigation of records arranged with paper and pen is alluded to as disconnected handwritten archive recognition frameworks. Character recognition is a key, however generally testing in the field of example recognition with a huge number of valuable applications. In this paper proposed to Fuzzy Neural Network utilizing Tamil script language recognition framework, a successful dynamic shape with MRF and Fuzzy Neural Network with Fuzzy C Means is used for character recognition and proficient district-based dynamic form with Improved Markov Random Field (IMRF) and support-vector-based fuzzy neural network classification (SVFNNC) system are recommended to improve the recognition exactness. The exhibition of the proposed plan is assessed utilizing different measurements, for example, Accuracy, Precision, Recall, and F-measure.

***Keywords**: Neural Network, Classification, Fuzzy Logic, Preprocessing, Segmentation, Classification, Noise Reduction.*

1. Introduction

Inscriptions are the fundamental hotspot for recreating the history and culture of ancient Tamil human advancements. Recognition of Ancient Tamil character is one of the significant undertakings to uncover the data from the ancient Tamil civic establishments [1]. The scripts of current Tamil language have developed over hundreds of years lastly changed to the present structure [2]. Present-day peruses discover trouble in translating a Tamil script of days of yore. The characters of Tamil script have changed after some time. Consequently, for perusing ancient scripts the period must be resolved, in order to know about which character set of ancient days is to be utilized for programmed perusing. The expectation of the period of a given ancient Tamil script is pursue on an individual from an ancient Tamil script recognition framework and can be utilized as a segment of the character recognition framework for ancient Tamil scripts. This knowle3dge can be utilized by archeologists and students of history for further investigations. The Tamil language is wealthy in ordinary scholarly and it was well known in different fields like prescription, stargazing and business. The information on the ancient individuals was clarified in [3]. Tamil has 12 vowels and 18 consonants. Tamil vowels are named as uyir eluttu (uyir – life, eluttu – letter). The vowels were named short (kuril) and long (five of each kind) and two diphthongs,/ai/and/auk/, and three "abbreviated" (kuril) vowels. The long (nedil) vowels are only twice the length of the short vowels. The diphthongs were articulated about 1.5 occasions as long as the short vowels, despite the fact that numerous linguistic writings place them with the long vowels. Tamil consonants are named as mey yeluttu (mey - body, eluttu - letters). The consonants are isolated into three sorts with six in each assortment: vallinam - hard, mellinam-delicate or Nasal, and itayinam - medium. These are joined together to acquire 216 composite characters and 1 uncommon character (aayutha ezhuthu) tallying to a sum of (12+18+216+1) 247 characters [4].

The content created by an individual by composing with a pen/pencil on a paper medium and which is then checked into a computerized group utilizing scanner is called Offline Handwritten Text. Online handwritten content is the one composed legitimately on a digitizing tablet utilizing stylus. The yield is a succession of x - y arranges that express pen position just as other data, for example, pressure (applied by the essayist) and speed of composing. Machine printed content can be found normally in day by day use. It is created by counterbalance forms, for example, laser, inkjet and some more. Specifically, Tamil handwritten OCR is more confounded than other related works. This is on the grounds that Tamil letters have more points and modifiers. Also, the Tamil script contains a huge number of character sets. A sum of 247 characters; comprising of 216 compound characters, 18 consonants, 12 vowels, and one extraordinary character. The principal challenges in OCR look into are because of the bends in the characters, a number of strokes and openings, sliding characters, varying composing styles so on.

The significant objectives of this work are

- To structure and build up a hearty character recognition framework for Tamil script language recognition
- To accomplish a better recognition rate with high exactness
- To improve the recognition exactness of uproarious picture based character recognition framework utilizing crossbreed classification method.
- To perceive the Tamil script language with the least recognition time.
- In request to accomplish lower computational multifaceted nature
- To propose new or changed element extraction method(s) reasonable for Tamil character.
- To propose new or adjusted pre-processing algorithms appropriate for Tamil character.

2. Literature Survey

John et.al presented a 1D Wavelet change of projection profiles for disengaged handwritten character recognition [9]. The preprocessed character pictures are displayed with a projection profile. One dimensional wavelet change is applied on the projection profile. The element vector is shaped from the smooth segments of the change coefficients. A Multi-Level Perceptron network is utilized for classification. **Arora et.al** presented multiple component extraction procedures for handwritten Devanagari character recognition [10]. The structured framework utilized four-component extraction procedures, in particular, convergence; shadow includes chain code histogram and straight-line fitting highlights. Shadow highlights are registered internationally for character picture while convergence highlights, chain code histogram highlights and line fitting highlights are figured by isolating the character picture into various portions. Weighted larger part casting a ballot strategy is utilized for consolidating the classification choice got from four Multi-Layer Perceptron (MLP) based classifier. **Bhattacharya et.al** structured a two phases recognition conspire for handwritten Tamil characters [11]. A presented strategy comprises of two phases. In the primary stage, apply a solo bunching technique to make fewer gatherings of handwritten Tamil character classes. In the subsequent stage, consider a managed classification system in every one of these little gatherings for definite recognition. The highlights considered in the two phases are extraordinary. The planned two-arrange recognition plot gave satisfactory classification correctnesses on both the preparation and test sets of the present database. **Obaidullah et.al** presented a Gabor channel-based strategy for disconnected Indic script recognizable proof from handwritten archive pictures. The work is completed at report level on four well known Indic scripts, in particular, Bangla, Devanagari, Roman, and Urdu. In the structured strategy, the 8-dimensional component vector is determined based on the Gabor channel. The recurrence and direction portrayal of this channel is like those with the human visual framework. Picture morphology-based highlights are likewise utilized alongside Gabor channel based highlights for the present work [12]. This improves in the general execution of the framework contrasted with if just Gabor channel is utilized. **Pradeep et.al** corner to corner based component extraction for handwritten character recognition framework utilizing a neural network [13]. The pre-preparing is a progression of activities performed on the examined info picture. Binarization process changes over a dark scale picture into a double picture utilizing a worldwide thresholding system. Location of edges in the binarized picture utilizing the Sobel method, widening of the picture and filling the openings present in it are the activities performed in the last two phases to deliver the pre-handled picture appropriate for segmentation. In the segmentation arrange, a picture of a grouping of characters is deteriorated into sub-pictures of individual characters. Corner to corner highlight extraction conspire for perceiving disconnected handwritten characters is proposed in this work. The multilayer feed-forward neural network is utilized for classification.

3. Research Methodology

This examination work acquires various commitments a powerful method to give more character recognition precision. Phase 1 – In this phase, To presented a Tamil script language recognition framework which is based on the Fuzzy Neural Network. Phase 2 – In this phase, a powerful dynamic shape with MRF and Fuzzy Neural Network with Fuzzy C Means is used for character recognition. Phase 3 – In this phase, an effective locale-based dynamic shape with Improved Markov Random Field (IMRF) and support-vector-based fuzzy neural network classification (SVFNNC) components are proposed to improve the recognition exactness. The general procedures engaged with the proposed research work are represented in the outline in Figure 1.

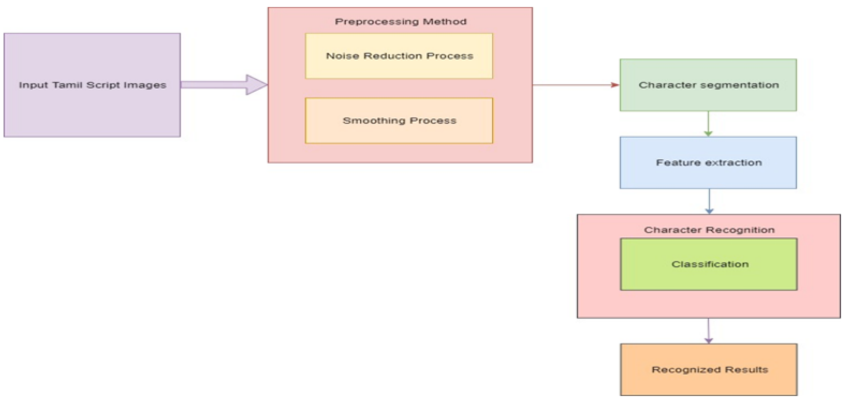

Figure 1 Flow of the proposed character recognition system

4. Proposed Work

4.1 Tamil Script Language Recognition in Degraded Stone Inscription Images

The main work presented a proficient Tamil script language recognition framework utilizing a fuzzy neural network. A contribution to the OCR framework is the examined report picture. This information picture ought to have an explicit arrangement, for example, .jpeg, .bmp and so on.

This picture is gained through a scanner, advanced camera or some other reasonable computerized input gadget.

4.1.1 Pre-processing

Preprocessing is essential for information or picture handling. The noise is presented in the picture while obtaining of picture or information, while moving the information, just as in light of changing certain parameters of securing a framework in the optical character recognition. Slant is a contortion that is frequently presented during examining or duplicating an archive and it is unavoidable. For the most part, preprocessing is utilized for smoothening and dispose of undesirable things from the info picture, specifically known as noise reduction and standardize the picture for further preparing. This proposed philosophy utilizes the Noise Reduction process as a preprocessing strategy. This noise reduction process contains two methodologies.

1. Gaussian smoothing Method
2. Bilinear filtering method

Gaussian filtering g is utilized to obscure images and expel noise and detail. Gaussian channel gives more weight to the current pixel position and afterward decreases the loads as distance increases according to the Gaussian equation. This channel can all the more likely safeguard the edges than the mean channel. This channel is applied to an image in a two phase approach. From the start, the even direction is separated by taking each pixel in the image, centering the channel on that pixel (center worth) and afterward multiplying the pixel esteems by the weight at each channel location and afterward isolates all to get the subsequent new pixel esteem. This process is then rehashed vertically on the on a level plane processed image to create the last image.

Bilinear can allude to bilinear filtering or bilinear addition. Bilinear filtering is a technique used to smooth out when they are shown bigger or littler than they actually are. Bilinear filtering utilizes points to perform bilinear introduction. This is finished by introducing between the four pixels closest to the point that best speaks to that pixel (as a rule in the center or upper left of the pixel). Bilinear introduction re-sampling takes a weighted normal of 4 pixels in the first image closest to the new pixel location. The averaging process modifies the first pixel esteems and creates entirely new advanced qualities in the yield image.

4.1.2 Segmentation using Markov Random Field (MRF)

The preprocessed image is portioned by utilizing Markov random field (MRF), which targets combining color and surface highlights. Markov Random Field (MRF) unaided segmentation approach comprises of two phases. The underlying advance is to pick a proper plan of components that can recognize the equivalent content highlights and meanwhile separate distinctive substance highlights and the second step is to apply an area developing technique over the selected highlights to accomplish a sectioned map14-15. MRF model-based segmentation approach imagines an image with the greater part of homogeneous neighboring pixels with comparable properties or highlights such as force, color, and surface. MRF model captures those comparable highlights among the pool of pixel forces to accomplish the segmentation process.

4.1.3 Chain Code Histogram based Features

The Chain Code Histogram of Character Contour is extracted to get better character recognition. Given a divided image, first, discover the contour points of the character image. We consider a 3 × 3 window encompassed by the object points of the image. In the event that any of the 4-connected neighbor points is a background point, at that point, the object point (P) is considered as contour point. The contour following procedure produces a contour portrayal called "chain coding" that is utilized for contour. Each pixel of the contour is relegated an alternate code that indicates the direction of the following pixel that has a place with the contour in some provided guidance. In this philosophy of utilizing a chain coding of connecting neighboring contour pixels, the points and the diagram coding are captured.

4.1.4 Classification using Fuzzy Neural Network

Based on the extracted highlights the classification is performed by utilizing a Fuzzy Neural Network. In this research, the fuzzy neural network model was created based on the neural networks and fuzzy frameworks. It coordinates the two speculations, by making up the defects of the neural network in fuzzy information processing and the deficiencies of fuzzy logic in learning. The fuzzy neural framework combines the learning capabilities of neural networks with the linguistic standard elucidation of fuzzy inference frameworks. The structure of FNN includes the age of IF-THEN rules. Here, the issue consists of the ideal meaning of the reason and consequent piece of fuzzy IF-THEN rules for the classification framework through the preparation capability of neural networks, assessing the mistake reaction of the framework. The SFNN is connection feed-forward architecture with five layers of neurons and four layers of connections. The primary layer of neurons receives the info data (Chain code histogram-based features). The second layer calculates the fuzzy enrollment degrees to which the information esteems have a place with predefined fuzzy participation functions, for example little, medium, or huge. The third layer of neurons speaks to associations between the information and the yield factors, fuzzy principles. The fourth layer calculates the degrees to which yield participation functions are matched by the info information, and the fifth layer does de-fuzzification and calculates values for the yield factors. Anyway it doesn't produce a high recognition rate.

4.2 An Efficient Active Contour with Markov Random Fields for Character Recognition

So as to overcome the above drawback, the proposed framework introduced an Active Contour strategy with Markov Random Fields for character recognition technique. The contribution to the OCR framework is the scanned document image.

4.2.1 Pre-processing

In the main stage, noise reduction is performed by utilizing the Laplacian smoothing and Bilateral filtering technique. Laplacian smoothing is performed on the documented images.

Laplacian Smoothing is a straightforward PDE-based smoothing approach planned as pursues:

$$v_i \leftarrow v_i + \sum_{v_j \in v_i^*} \left(\frac{v_j - v_i}{d_i}\right) \quad (1)$$

Where, $V = \{v_1 ... v_k\}$ speaks to the arrangement of vertices. This process can be done more than once to correct the location of each vertex to the geometric center of its neighboring vertices. This approach is straightforward and quick, be that as it may, it produces an over smoothing result after little cycle. The respective channel is a non-iterative, local and straightforward strategy for expelling Gaussian noise while safeguarding edges. As the name suggests Bilateral channel is a combination of range and space filtering. Conventional filtering is area filtering and enforces closeness by gauging pixel esteem with coefficients that tumble off with distance. Likewise, extend filtering can be characterized as which midpoints image esteems with loads that decay with disparity. Range channels are nonlinear because their loads rely upon image force or color. Above all, they save edges moreover.

4.2.2 Segmentation using Region based Active Contour Method with Markov Random Fields

In this work, propose a novel locale-based active contour model in a level set definition that can be utilized to section images with power non-consistency and elevated level noise. The fundamental thought of our proposed technique is to utilize Gaussian conveyances with various means and variances while incorporating a model of power non-consistency (inclination) for synchronous segmentation and force non-consistency correction. Moreover, by incorporating a Markov random field to coordinate the spatial data between neighboring pixels in the active contour model, the proposed summed up area-based model can be utilized to section images in the presence of image artifacts.

4.2.3 Feature Extraction

Chain code histogram highlights are extracted by chain coding the contour points of the scaled character bitmapped image. Minute based highlights are extracted from scaled, diminished one-pixel wide skeleton of the character image. The estimations of seven minutes speak to the basic image and are utilized to create an element vector consisting of seven qualities. Image is sectioned into nine equivalent sub-images and in each sub-image minute highlights are calculated so all out 63 highlights are framed.

4.2.4 Classification using Fuzzy Neural Network with Fuzzy C Means

Fuzzy Neural Network with Fuzzy C Means is utilized for the classification process. So as to cluster the positive and negative class the Fuzzy C Means clustering algorithm is coordinated with the yield layer. The Fuzzy C-means (FCM) clustering algorithm is the fuzzy adaptation of the K-means clustering algorithm. It targets limiting the cost function. Notwithstanding, the criterion utilized in the FCM algorithm just considers the distance from a hub to the cluster centers with their fuzzy participation. The test characters are inspected from the script automatically matched with the characters accessible for various periods utilizing machine intelligence. In this work, FNN is utilized for classification which utilizes fuzzy C-intend to gain proficiency with the parameters of the fuzzy standards, participation functions and cluster the class as positive or negative from the preparation information. Be that as it may, such sorts of learning algorithms ordinarily cannot limit the empirical hazard (preparing mistake) and expected hazard (testing blunder) at the same time, and along these lines cannot reach a decent classification performance in the testing phase.

4.3 An Efficient Region based Active Contour with Improved Markov Random Field for Character Recognition

So as to overcome the above drawback, the proposed framework introduced an active contour strategy with Improved MRF for the character recognition technique. The contribution to the OCR framework is the scanned document image.

4.3.1 Preprocessing

The scanned image is preprocessed for noise expulsion utilizing low channel smoothing and advanced middle channels. Low-pass channel is a sort of channel utilized for image enhancement. It saves the smooth locale in the image and evacuates the sharp variety prompting obscuring effect. A low-pass channel is a channel that passes low-frequency flag and constricts signals with frequencies higher than the cut-off frequency. The actual measure of weakening for each frequency differs relying upon specific channel structure. Smoothing is on a very basic level a low pass activity in the frequency space. The improved middle channel algorithm is divided into three phases, which call levels A, B and C processing, respectively. The algorithm begins from a pre-processing step before the process continues to the three levels. In the preprocessing stage, the window is selected and the qualities in the window are arranged and spare in the arranged rundown. The base pixel worth, X_{min}, and the most extreme pixel esteem, X_{max}, are compared with motivation esteems K1 and K2, respectively, where K_1=0 and K_2=255. On the off chance that $X_{min} = K_1$ or in the event that $X_{max} = K_2$, at that point the window has motivation noise and processing proceeds through all levels A, B, and C.

4.3.2 Segmentation using Region-based Active Contour with Improved MRF Model

In this work, a novel district-based active contour with an improved MRF model is introduced for segmentation. The proposed segmentation technique is based on an improved MRF model, which incorporates prior and limits data of the image. Besides, by incorporating MRF to coordinate the spatial data between neighboring pixels in the active contour model, the proposed summed up locale-based model can be utilized to section images. The earlier vitality model includes more limit and power data of the image just because, along these lines, it can improve the segmentation while acquiring a decent limit especially for the noise image.

4.3.3 Feature Extraction

By and large, there are two sorts of highlights, statistical highlights, and structural highlights. Statistical highlights contain pixel thickness, minute, and mathematical change, etc. Structural highlights conclude stroke, contour, number of bifurcation points, number of circles, etc. These highlights are extracted by utilizing Independent Component (ICA) approach. The portrayal of a document image by the statistical appropriation of points deals with style varieties somewhat. In spite of the fact that this sort of portrayal doesn't permit the reconstruction of the first image, it is utilized for reducing the component of the list of capabilities giving fast and low complexity. The major statistical highlights referenced are zoning, Projections, Crossings, and Distances which are utilized for character portrayal. Structural highlights are based on the topological and geometrical properties of the character. Different worldwide and local properties of characters can be spoken to by geometrical and topological highlights with high tolerance to twists and style varieties. This kind of portrayal may likewise, encode some information about the structure of the object or may give some information regarding what kind of components make up that object. Different topological and geometrical portrayals can be assembled in four categories: they are Extracting and Counting Topological Structures, Measuring and Approximating the Geometrical Properties, Coding, and Graphs and Trees.

4.3.4 Classification using FNN with SVM

The proposed framework introduced a support-vector-based fuzzy neural network classification (SVFNNC) mechanism. The extracted highlights are given to the SVFNNC strategy as info. The SVFNNC combines the prevalent classification intensity of support vector machine (SVM) in high dimensional information spaces and the efficient human-like thinking of FNN in taking care of uncertainty data. The learning algorithm consists of two learning phases. In phase 1, the fuzzy standards and enrollment functions are automatically dictated by the clustering principle. In the second phase, the ideal parameters of SVFNNC are calculated by the SVM technique.

5. Experimental Results

This section assesses the performance of the proposed framework regarding precision, recall, f-measure and classification accuracy with respect to a number of images. The examinations are carried out in Matlab

Dataset

1. Accuracy

The weighted percentage of characters in images is correctly portioned by the estimation accuracy. It is spoken to as,

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \quad (2)$$

Where,

T_P -True positive

T_N – Ture negative

F_P -False positive

F_N – False negative

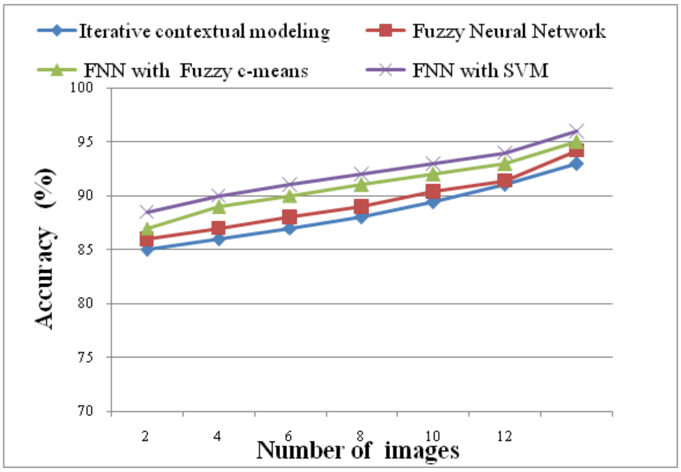

Figure 2 Accuracy Comparison

Figure 2 compares the accuracy by utilizing proposed FNN with SVM based character recognition approach and existing iterative contextual demonstrating, FNN and FNN with fuzzy c-means based character recognition strategies. The quantity of images is taken as X pivot and accuracy as y hub. In the proposed framework, Statistical and structural highlights of the characters are extracted by utilizing ICA. It improves the recognition rate. It very well may be said that the proposed FNN with SVM based character recognition approach has high accuracy esteem when compared with the other existing iterative contextual displaying, FNN and FNN with fuzzy c-means based character recognition strategies. From the chart results, it is seen that the proposed FNN with SVM based character recognition approach achieves accuracy esteem is 92% which is 1%, 0.8%, 1.2%, higher than FNN with fuzzy c-means, FNN and iterative contextual based character recognition techniques respectively.

1. Precision

Precision is characterized as a computation of correctness or quality, though recall is a computation of completeness or amount. What's more, high precision indicates that the approaches returned significantly more applicable outcomes than unessential. The precision is calculated as pursues:

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (3)$$

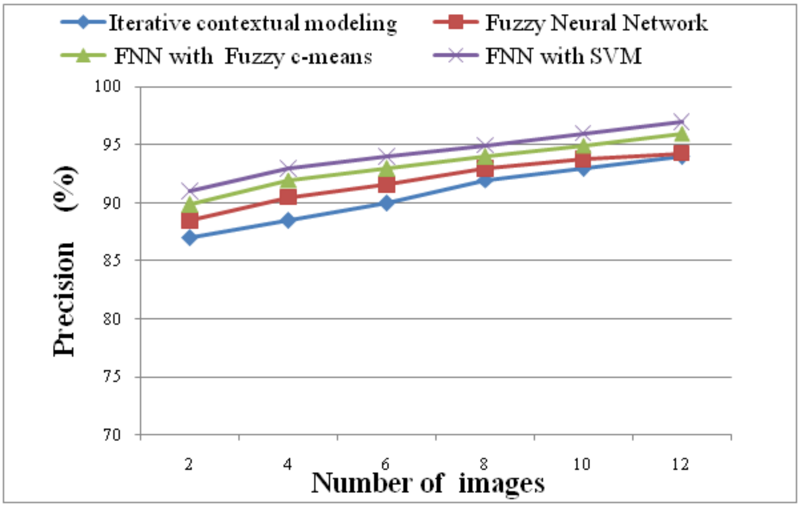

Figure 3 Precision Comparison

Figure 3 compares the precision by utilizing proposed FNN with SVM based character recognition approach and existing iterative contextual demonstrating, FNN and FNN with fuzzy c-means based character recognition strategies. The quantity of images is taken as X pivot and precision as y hub. In this proposed work, low channel smoothing and advanced middle filtering are utilized for reducing the noise and smoothen the information images. It very well may be said that the proposed FNN with SVM based character recognition approach has high precision esteem when compared with the other existing iterative contextual displaying, FNN and FNN with fuzzy c-means based character recognition strategies. From the chart results, it is seen that the proposed FNN with SVM based character recognition approach achieves precision esteem is 94.33% which is 1.01%, 1.36%, 1.2%, higher than FNN with fuzzy c-means, FNN and iterative contextual based character recognition techniques respectively.

2. Recall

The recall is described as the number of pertinent documents recovered through a search separated by the complete number of accessible applicable documents. The recall is additionally the number of genuine positives isolated through the complete number of components that effectively have a place with the positive class. The calculation of the recall esteem is done as pursues:

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (4)$$

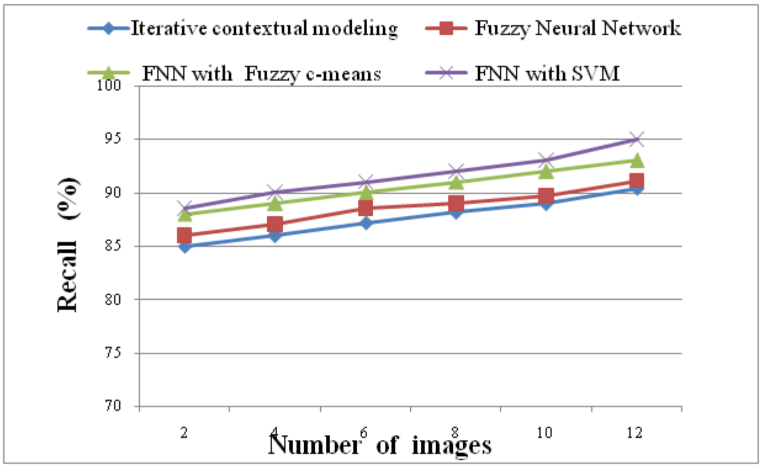

Figure 4 Recall Comparison

Figure 4 compares the recall by utilizing proposed FNN with SVM based character recognition approach and existing iterative contextual demonstrating, FNN and FNN with fuzzy c-means based character recognition techniques. The quantity of images is taken as X hub and recall as y hub. In this proposed work, the segmentation technique is based on an improved MRF model, which coordinates priori and limit data of the image to improve the segmentation accuracy. It very well may be said that the proposed FNN with SVM based character recognition approach has high recall esteem when compared with the other existing iterative contextual demonstrating, FNN and FNN with fuzzy c-means based character recognition techniques. From the diagram results, it is seen that the proposed FNN with SVM based character recognition approach achieves recall esteem is 91.58% which is 1.08%, 1.94%, 0.91%, higher than FNN with fuzzy c-means, FNN and iterative contextual based character recognition strategies respectively.

3. F-measure

It computes the combined estimation of precision and recall as the harmonic mean of precision and recall. The f-measure esteem is gotten as pursues:

$$F = 2.\frac{Precision\ .Recall}{Precision\ +Recall} \qquad (5)$$

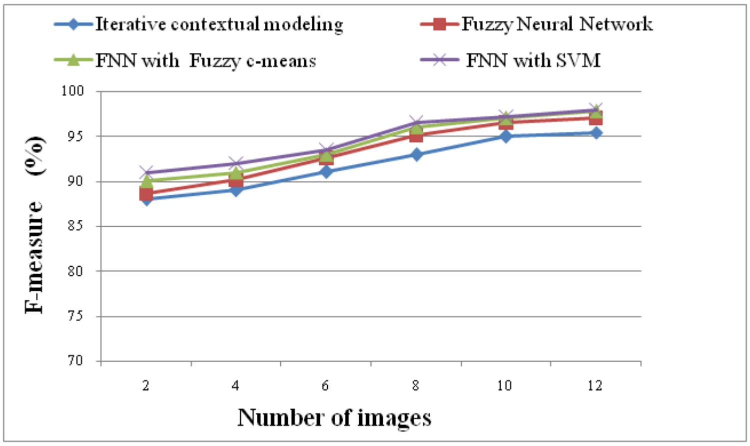

Figure 5 F-Measure Comparison

Figure 5 compares the F-measure by utilizing proposed FNN with SVM based character recognition approach and existing iterative contextual demonstrating, FNN and FNN with fuzzy c-means based character recognition techniques. The quantity of images is taken as X pivot and f-measure as y hub. So as to smoothen and evacuate the noise the proposed framework utilized low channel smoothing and advanced middle filtering strategies. The preprocessed image is divided by utilizing a district-based active contour strategy with IMRF. After the element extraction, characters are recognized by utilizing FNN with SVM. It tends to be said that the proposed FNN with SVM based character recognition approach has high f-measure esteem when compared with the other existing iterative contextual demonstrating, FNN and FNN with fuzzy c-means based character recognition techniques. From the chart results, it is seen that the proposed FNN with SVM based character recognition approach achieves f-measure esteem is 94.7% which is 0.56%, 0.8%, 1.41%, higher than FNN with fuzzy c-means, FNN and iterative contextual based character recognition strategies respectively.

Conclusion

In the main work, an efficient Tamil script recognition framework is introduced utilizing a Fuzzy Neural Network with K-means. After the completion of preprocessing the characters are sectioned utilizing MRF. At that point, the chain code histogram highlights are extracted from sectioned images. At long last, the Classification is performed by utilizing FNN. Anyway, it doesn't produce a satisfactory outcome. To overcome the previously mentioned issues, the second work introduced an active contour with Markov random fields for character recognition. So as to expel the undesirable noise from input image the Laplacian smoothing and respective filtering strategy is performed. By utilizing Active Contour strategy with MRF technique the preprocessed image is portioned. The Chain Code Histogram of Character Contour and minute highlights are extracted for classification. At long last, the classification is performed by utilizing a Fuzzy Neural Network with fuzzy c-means. Be that as it may, it cannot reach a decent classification performance in the testing phase. To tackle this issue, the third work introduced an efficient district-based active contour with IMRF for character recognition. In the main stage, and info image is preprocessed by utilizing low channel smoothing and advanced middle filtering techniques. The locale based active contour strategy with the IMRF technique is used for an image that is portioned. At that point there are two sorts of highlights such as statistical and structural highlights are extracted. At last, the classification is performed by utilizing the SVFNNC mechanism. The trial results show that the proposed framework achieves better performance compared with the current framework as far as precision, recall, f-measure, and classification accuracy. Our future work expects to improve recognition rate and furthermore to grow new zone-based component extractions algorithms, which gives efficient outcomes.

References

1. Kanimozhi.V.M, Muthumani.I 2014, 'Optical Character Recognition for English and Tamil Script', International Journal of Computer Science and Information Technologies, vol.5, no. 2, pp 1008 – 1010.
2. Dan Claudiu Cires¸an and Ueli Meier and Luca Maria Gambardella and Jurgen Schmidhuber 2011, 'Convolutional Neural Network Committees for Handwritten Character Classification', International Conference on Document Analysis and Recognition, IEEE, 2011.
3. Chinnuswamy P and Krishnamoorthy S G 1980, 'Recognition of Hand printed Tamil Characters', Pattern Recognition, pp.141-152.

4. Aparna K H, Jaganathan S, Krishnan P & Chakravarthy VS 2002,' An Optical character recognition system for Tamil newsprint', In International Conference on Universal Knowledge and Language-2002, pp. 881-886.
5. Sutha J and Rama Raj N 2007, 'Neural network based offline Tamil handwritten character recognition System' , International Conference on Conference on Computational Intelligence and Multimedia, vol: 2, pp. 446 – 450, 2007.
6. StutiAsthana, FarhaHaneef and Rakesh K Bhujade 2011,' Handwritten Multiscript Numeral Recognition using Artificial Neural Networks', Int. J. of Soft Computing and Engineering ISSN: 2231-2307, vol.1, no.1.
7. Tay, Yong Haur, Pierre-Michel Lallican, Marzuki Khalid, Christian Viard-Gaudin, and S. Kneer 2001, 'An offline cursive handwritten word recognition system', Proceedings of IEEE Region 10 International Conference on Electrical and Electronic Technology, vol. 2, pp. 519-524.
8. LiuH & Ding, X 2005, 'Handwritten character recognition using gradient feature and quadratic classifier with multiple discrimination schemes', In Document Analysis and Recognition, pp. 19-23.
9. John R, Raju G and Guru DS 2007, '1D Wavelet transform of projection profiles for isolated handwritten character recognition', pp.481-485.
10. AroraS, Bhattacharjee D, Nasipuri M., Basu DK & Kundu M2008, 'Combining multiple feature extraction techniques for handwritten devnagari character recognition', In Industrial and Information Systems, pp. 1-6.
11. Bhattacharya U, Ghosh S K, & ParuiS 2007, 'A two stage recognition scheme for handwritten Tamil characters', In Document Analysis and Recognition, vol. 1, pp. 511-515.
12. Obaidullah SM, Nibaran D & Kaushik R 2014,' Gabor filter based technique for offline indic script identification from handwritten document images', pp. 1-5.
13. Pradeep J, Srinivasan E & Himavathi S 2011,' Diagonal based feature extraction for handwritten character recognition system using neural network', In Electronics Computer Technology (ICECT), vol. 4, pp. 364-368.