MULTILINGUAL OFF-LINE HANDWRITING RECOGNITION IN REAL-WORLD IMAGES USING THE ONLINE SUPPORT VECTOR MACHINE

[1*]M.Sivasankari, [2]Dr.P.Velmani, and [3]Dr.P.Arokia Jansi Rani
[1*]Reg.No.11994, Research Scholar in Computer Science (Part-Time-Internal), Manonmaniam Sundaranar University, Abishekapatti ,Tirunelveli-627 012, Tamilnadu, INDIA. E-mail: mvsivasankari@gmail.com.
[2]Assistant professor in Computer Science, The M.D.T. Hindu College, Tirunelveli.
[3]Assistant professor in Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli.

Abstract: Manuscripts transmit messages, ideas and thoughts in writing with documentary proof which is a form of communication. This paper majorly focuses on the Feature extraction that is the process to retrieve the characters that are represented accurately from the raw data. The purpose of feature extraction is to maximize the recognition rate with the characters. It is considered to be a difficult task to obtain the features of the handwriting which has high degree of variability and imprecision. In this paper the steps namely Denoising ,binarization, noise elimination, Thresholding, and size normalization for an image are carried out then words segmentation is performed by using bat algorithm. In feature extraction stage each character's identity is represented as a feature vector. In feature extraction the set of words have been performed through structural and statistical methods. To identify the style of words and characters, fuzzy clustering algorithm has been used. The efficient recognition of word has been done through the online support vector classifier. The result shows that the proposed methods have been performed well when compared with the existing methods.

Keywords: Off-Line Handwriting Recognition, Image Denoising, Binarization, Noise Elimination, Size Normalization, Bat Algorithm, Edge Detection, Fuzzy Clustering, and Online Support Vector Machine (OSVM).

1. INTRODUCTION

Handwriting recognition is the method of identifying the words, characters, pattern and numerals of a text document, scanned images and some other records. Many diverse applications like reading medical prescription, bank cheques, and other official documents make use of this Handwriting recognition. It is used to identify the words easily. There are numerous researches and projects have been done on recognition of characters and numerals of handwriting. But still there is a problem in the field of handwritten recognition because of different styles of font which is complex to handle. As the different people use the different styles for writing, it makes recognition a difficult task. Mostly recognition of English handwritten is available in the literature [2] and some on their vernacular languages and scripts[19].This paper focuses on the recognition of multilingual handwritten mainly Tamil and English.

In order to exclude the challenges of handwriting recognition, feature extraction method has been used. The Optical Character Recognition (OCR) plays a vital role in handwriting recognition. The noise removal is the primary process involved in recognition. Feature extraction is done through normalizing the size and binarization which eliminates the noise. After that segmentation of words is performed in which the words are clustered and the particle swarm optimization algorithm is involved. The deep neural network is utilized for recognition of words. This vital process makes up the major part of recognition.

The feature extraction, edge detection and style identification are the final phase of methods involved in handwritten recognition. Extraction of knowledgeable data from raw set of data is the method involved in the feature extraction. The data extracted provides the high accuracy.

Each character in the document is denoted as feature vector in the feature extraction process which becomes its identity. The extraction is performed by means of structural and statistical features. The various structure and the styles of the words are taken into consideration in this feature. Then the result is summarized to know about the accuracy of detection in the handwriting. Style identification is done through fuzzy clustering model which defines the unique style of different users.

The final step is word identification which is performed through Online Support Vector Machine (OSVM). This OSVM model identifies the word with high accuracy by eliminating the errors. The rest of the paper is organized as follows. Section 2 discusses the literature review. Section 3 overviews the proposed technique for feature selection and classification. Experimental results of the proposed system are presented in Section 4. Section 5 concludes the paper

2. LITERATURE REVIEW

Chopra et al [1] developed a recognition system by means of Optical Character Recognition (OCR) only for the documents comprising of fixed style and font size. This model is used for offline recognition of characters and it uses database for alphabet recognition. The primary process in this study is pre-processing of document images. The important step involved is feature extraction on optical character recognition. It has the greater advantage as it provides higher scalability and it is used for predefined English documents, but it could also be configured for other documents too.

Reddy et al [2] proposed a multilayer perception model for recognition. This device is a smart technology that uses half of the human brain thinking for the recognition of both characters and digits. In order to train and test the data, MLP network uses back propagation (BP) algorithm. The ANN greatly offers an advantage of pattern recognition in an effective way.

Vithlani and Kumbharana [3] have provided a study for handwritten recognition by means of Optical Character Recognition (OCR). The pattern of the character is identified in this study. The OCR involves the various steps of process in it. They involve in Acquisition, pre-processing, Segmentation, feature extraction and classification. The various algorithm namely, statistical algorithm, template matching algorithm, neural network algorithm and structural algorithm are defined in this study for character pattern recognition.

Bobade and Sahu [4] have developed a character recognition model by means of Neural Network (NN). The feature extraction is the vital process involved in the character recognition model. In order for normalising the character image, pre-processing has been performed. For the purpose of classification back propagation has been performed. The different feature extraction based on comparative analysis in terms of accuracy has been performed. The single feature vector by combining various feature vectors is formed for hybrid method which on comparison with individual feature extraction provides high accuracy rate of character recognition.

Oujaoura et al [5] presented recognition methods for identifying the Tifinagh characters. This is made possible by means of Data mining tools which include Adaptive Neuro Fuzzy Inference System (ANFIS), Support Vector Machine (SVM), Artificial Neural Network (ANN), Classification and Regression Tree (CART), AdaBoost and K-Nearest Neighbors (K-NN). The feature extraction involves the Zernike moments and geodesic descriptors for extracting attributes. Comparison of classification data mining for recognition of Tifinagh recognition is the main objective of this study. Oracle Data mining and Oracle database are used for comparison in working environment to identify the effective outcome.

Al-Maadeed et al [6] aimed to develop a new system which selects the vital feature thereby eliminating the non-effective features. This model could be developed by means of Lukasiewicz implication on fuzzy conceptual reduction. Both English and Arabic languages are taken into consideration for handwritten recognition of characters. The database of 121 writers is considered on which the K-Nearest Neighbors (K-NN) is used to evaluate the accuracy efficiency. The left or right handedness parameter is considered for evaluation and it produces the high accuracy of 83.43%.

Liwicki and Bunke [7] used whiteboards to perform online based handwritten recognition on with the help of Hidden Markov Model (HMM). The state of the art normalization and feature extraction are used to convert text line into sequence having feature vectors. By doing added preprocessing methods, the rate of recognition has been enhanced. Classification process is made to be carried out using statistical languages and a model named Hidden Markov. In the presence of model with no language, 67.3% rate of the recognition has been achieved. On the other hand, about 70.8% of rate has been achieved in the model with language.

Dedgaonkar et al [8] have discussed about various recognition techniques used for character recognition. The recognition methods can be classified broadly under two categories. They include online and offline recognition. The various offline recognition methods namely, clustering, pattern matching; feature extraction and artificial neural network have been discussed. The Direction based algorithm and K-NN classifier are the two subdivisions used for recognising online characters. These methods are applied for the effective recognition with high accuracy.

Shah et al [9] have implemented two models for handwritten recognition of diverse languages. The models include Tesseract OCR and K-Nearest Neighbor (K-NN). In this study working of both the model has been briefed and comparison is performed among both the model for character recognition to gain the accuracy of both. This model is designed in such a way that it also helps in image recognition, number plate recognition for license and extraction of text from various kinds of document.

Dash and Nayak [10] proposed a Artificial Neural Network (ANN) is purely depends on English character recognition [10]. This recognition mechanism is about offline system, if no linear relationship is found the performance of matching character is to be carrying out. In order to find out cluster presence, test has been performed over character. Mat lab platform helps to take over this process. Then the ANN model is trained and the network has been tested using English character includes fifty-two sets of the alphabets. Totally 26 small and 26 capital alphabets are comprised by sets. This Neural Network model paves a way to obtain better recognition of about 85%.

Christian Gruber, Thiemo Gruber and Sebastian Krinninger [11] presented a new technique for online signature verification or identification. It makes use of the longest common subsequences (LCSS) detection algorithm which measures the similarity of signature. It is possible to consider the local variability of signals through LCSS. It is easy to determine the similarity of two different signature in a more reliable way than with other measures.

Shanjana, C. James, Ajay [19]proposed a work of fiction method to convert handwritten Malayalam text into editable text format that can be stored by machine.

3. METHODOLOGY

The handwritten recognition is one of the most challenging tasks where many researchers have been still facing challenges for accurate prediction of words and characters. The offline recognition demands the following process for identification. The original data carries noise along with it. So in order for eliminating the noises from the source data various steps like binarization, size normalization and noise elimination are performed as are pre-processing steps. Binarization is the technique where the digitalization of the data is performed after conversion into gray scale and binary image. The de-noising is performed by eliminating loops, gaps and bumps which are general in scanned images. The variations in the images are removed by the process of size normalization without modifying the identity of words or characters.

Segmentation would be the trailing process where the clear separation of words and characters are determined by various algorithms. The BAT algorithm is one among them and provides an effective segmentation of words as it is an optimizing technique. Edge detection by means of descriptor is indulged where clear boundary separation is determined as it is the most common way of recognizing words, images and so on.

The next process is feature extraction which enhances the rate of identifying the words. Here the set of feature words are extracted by various means. The structural and statistical feature selection is involved in this domain which is of multi- resolution aids to enhance the quality of the prediction. The output from this

section is provided as input to the classifier. The classifier acts as a decision making authority in the handwritten recognition. The fuzzy logic is the process of identifying the style of words in the handwritten recognition which provides effective outcome even in the presence of noise. Online Support Vector Machine (OSVM) classifier is employed for quicker and efficient recognition of words. This classifier is also employed on Optical Character Recognition and content of text documents. The final phase of recognition is post-processing where the errors occurred during the whole process is eliminated as it contains some noises in it. This increases the performance of the recognition to be an outstanding one. The proposed system flow diagram is shown in figure 1.

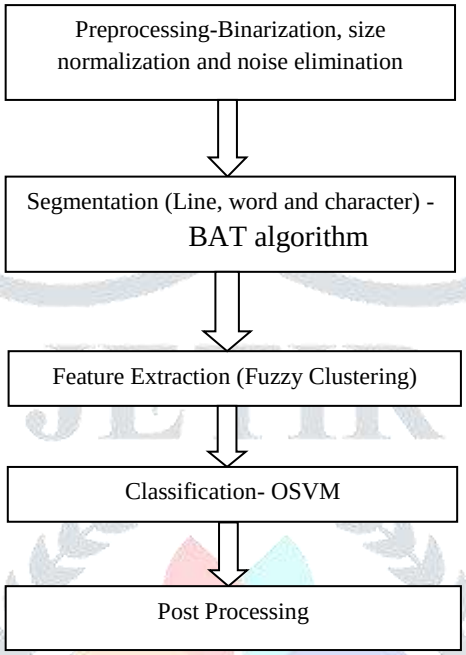

Figure1: Proposed system flow diagram

3.1. PRE PROCESSING

Pre-processing is the process where the source data is provided as input to get the data in the user understandable form on performing the following process:

3.1.1NOISE ELIMINATION

Pre-processing is the first step used for removing noise in image. Here Modified algorithm has been used and it is called as [11] K-Algorithm to perform effective reduction. In modified approach, two steps have been involved such as filtering and the binarization. Re-Sampling could be used to do filtering process. Textured images are used here for filtering process in the wish of remove noise but it may paves way for some issues. Thus binarization is used to do process and brings out its advantage in order to deal the issues that aroused in filtering. On the other hand, during this process the information considered as useful are made to retain safe. In addition to that, there are two filters involved in filtering process and they are linear and non linear type. Issues may arose while using linear filters, thus non linear filters are used to deal with those issues.

3.1.2 BINARIZATION

Binarization [12] is a methodology applicable for image denoising. This technique would remove the noise using certain advanced systems with it. That system is called to be as Modified approach named K-Algorithm. Combinations of thresholding and filtering have been bringing out to do image processing. Thus it behaves as refinement approach since it increases the quality of image. Binarization is using to avoid

some problems that relates to images. Transformation of images data in to digital data is the process taken inside binarization. Finally, the threshold value could be measured that based on color obtained.

3.1.3 SIZE NORMALIZATION

Normalization process could be carried out in image depending recognition. It mainly focuses in changing the range of intensity of images. Normalization aims to improve the sense of visibility on images in the aim of enhancing the quality in terms of images. Quality of image can be defined based on the intensity value. Normalization approach adopts the automatic processing which helps to normalize image in any format. It would obtain a result having dimension in constant state. Size normalization [13] is one type of approach in normalization mechanism that would modifies the size of character which is based on some form in means of standard. The recognition has been applied on characters depends on two direction namely horizontal and vertical.

3.1.4 SEGMENTATION- BAT ALGORITHM (BA)

The bat algorithm is the step which would carry out after noise filtering. Through preprocessing, the noise could be removed from input information which may be text or signal or image. Bat Algorithm [14] is mainly used to extract the key features with the help of various approaches. After selecting the required features, the classification has been performed. Firstly, the important features are extracted and it is gives as input to classifiers in the aim of enhancing the performance. Features extracted using this algorithm may be found as insignificant, so Improved bat algorithm [15] has been brought out. These mainly focus on measuring optimal value and reduction of feature set in terms of population. Rosen Brock is a function used for evaluating fitness of bat population and it define in 1st equation,

$$f(x) = \sum_{i=1}^{j=1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2] \quad (1)$$

Thus,

J = dimension can be represented using j

x(i) = denotes i^{th} bat.

ri = represents rate of pulse,

Ai = denotes frequency,

fmin, fmax = representation for frequency.

In Modified Bat Algorithm, initial population is made to assume and each entry involves ωi which defines frequency, v_i represents velocity.

$$\omega_{i=}\omega_{min} + (\omega_{max} - \omega_{min})\alpha \quad (2)$$

$$v_i^t = v_i^{t-1} + (G_i^t - G_{current})\omega_i$$

Thus,

$G_{current}$ = denotes global solution at present

α = function of uniform distribution ranges from 0 to1.

Best solution has been identified and could be updated.

$$G_{best} = G_{current} + \epsilon S_i \quad (3)$$

Where,

ϵ = denotes number randomly occurs that ranges from -1 to 1

S_i = helps to signifies the features that are similar.

Algorithm for Improved bat algorithm shown here,

Initialize the dataset: X_i where $1 \leq i \leq n$

Set Frequency ω_i velocity v_i

While $t<X_nD_0$

Update ω_i and v_i

Estimate transfer function value

Update X_i

Generate rand

Select random instance from X_i

Compute similarity

For (rand= X_i; rand<= X_n;= X_{i++})

If (rand is highly similar than the i^{th} instance)

Choose G_{best} from the available solutions

Update X_i with G_{best}

End if

Sort the features based on G_{best}

End For

End while

3.2. FUZZY CLUSTERING

Clustering mechanism having Fuzzy [16] helps to find out the function named kernel in the aim of providing hopeful performance. On Different function of kernel, experiments have been made in the sense of creating technology as identification of writer. Thus these mechanism has been using in the area of performance ? measuring. This Measuring process could purely based on identification. In Addition, Inverse and Gaussian function are mainly used for the above said process.

Gaussian Kernel Function

Assumption of the Zero mean and the Unit variance with Gaussian has been made for the purpose of distributing features. The results of exponential kernel function, is similar to the result obtained by Gaussian kernel function. This could be achieved by means of parameter which is the adjustment that made on it. This adjustment depends on two parameters namely variance and mean.

Inverse Kernel Function

This defines equation for kernel function.

$$C_{\alpha_{k_j}} = \sum_{p=1}^{M} \frac{\frac{1}{dist\ (x_{\alpha_p}, x_{\alpha_k})}}{\sum_{k=1}^{N} \frac{1}{dist\ (x_{\alpha_p}, x_{\alpha_k})}} \quad (4)$$

3.3. ONLINE SUPPORT VECTOR MACHINE

Support Vector Machine is one type of mechanism using for performing classification. Since large amount of information has been emerging in day to day life of human, everyone need a system that would do better working as then before. Thus Online Support Vector Machine [17] has been introduced in the sense of doing better and speedy classification in any field. This system also helps to deal with images in Optical Character Recognition or text content in paper. Due to the necessity in improving accuracy or performance and minimize the memory usage, this approach had brought out. Since SVM has large computational cost, even though it is good in classification it is not using that much. On the other hand, SVM is known for its strong foundation in theory and its hopeful generalized performance. Thus online support vector machine is introducing to give out trustworthy result in the field of Word Recognition. LASVM is one type of algorithm in online SVM which is an online based kernel classifier that would use less memory.

4. EXPERIMENTAL DESIGN

The data set comprised of 30 Tamil characters of 12000 characters has been grabbed from various documents. For the purpose of training this model gathered 9000 (300 samples * 30characters) samples are collected to train the model. The remaining 3000 samples (100 samples * 30 characters) are employed for the purpose of testing. 5200 samples of 26 English characters are collected from different handwritten documents. 2600 samples (100 samples * 26 characters) are gathered for training purpose, rest of the 2600 samples (100 samples *26 characters) are used for testing purpose. The Microsoft excel sheet is maintained for the feature data.

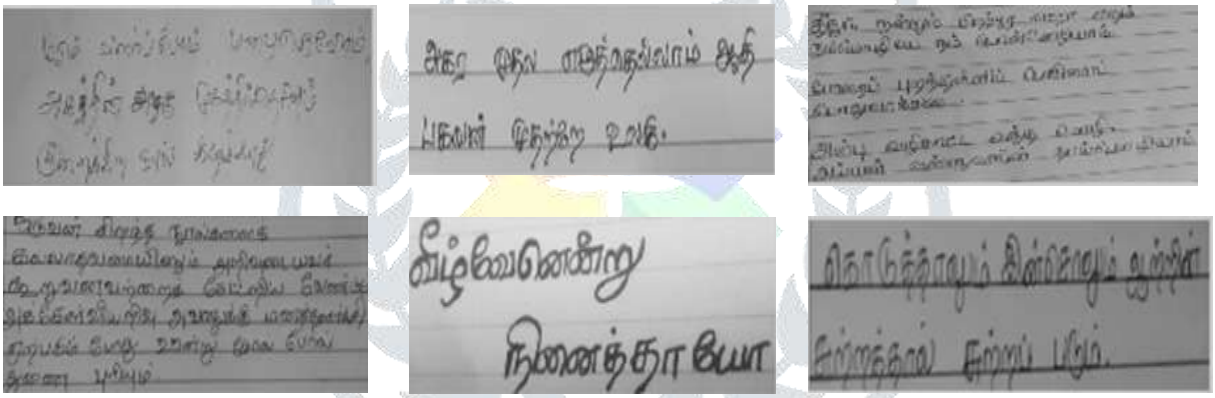

Figure 2(a) . Tamil Handwriting samples

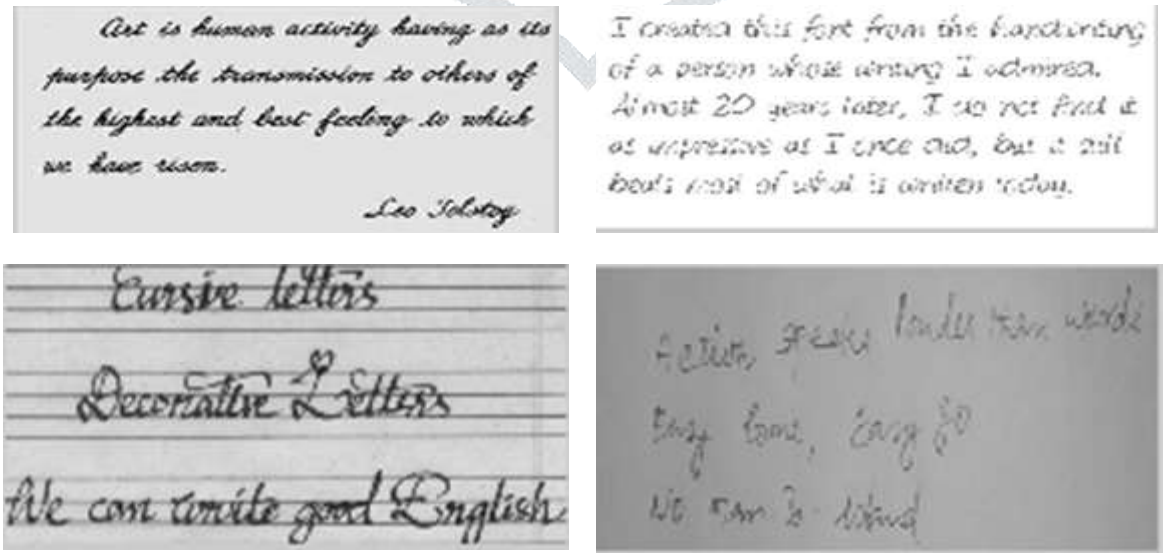

Figure 2(b). English Handwriting samples

The Figure 2(a) & 2(b) represents the images contain text data of two different languages which are provided as input to the system which improves the quality of the image by various processes and provides the better recognition of texts. The raw images by means of scanning are provided to the system as input and various steps are performed on the image to identify the words and characters in this model.

4.1. RESULTS

The handwritten recognition takes the scanned images and text documents as input which comprises noises in it. The noise elimination is performed by means of pre-processing and trailing steps like segmentation, feature extraction; classification and word recognition are being performed in this system. The OSVM algorithm shows the optimistic increase in efficiency and accuracy of the recognition. The results of various classifiers with different metrics for Tamil and English language are shown in Table 2 and Table 3. Table 1 shows the results of two different characters and in two languages results based on the maximum and minimum number of iterations required to complete the process. Final column shows, in which iteration the overall character is recognized in the input samples for DNN, ANFIS and proposed OSVM classifier.

Table 1 .Performance of different methods with sample characters.

Methods	Sample character	No. of trained character (minimum and maximum)		No. of particular character to be recognized within the iterations
		Minimum	Maximum	
DNN		20	100	72
		17	100	80
DNN	i	15	100	85
	j	20	100	75
ANFIS		18	100	68
		15	100	75
ANFIS	i	13	100	82
	j	16	100	69
OSVM		15	100	64
		13	100	70
OSVM	i	11	100	75
	j	13	100	62

Table 2. Performance comparison results of Tamil character (OSVM)

Algorithm	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	Accuracy (%)	Time(Seconds)
OSVM	93.00	80	94.5	93.75	95.00	7.2
ANFIS	90.20	77.5	93	91.6	93.20	9
DNN	87	74.5	91	89	90.52	12
SVM	82	70	84	83	85.13	19
Naïve Bayes	85	72.3	89	87	87.00	15

Table 3. Performance comparison results of English character (OSVM)

Algorithm	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	Accuracy (%)	Time(Seconds)
OSVM	94.00	79	94.00	94.00	94.00	7
ANFIS	91.20	77	92	91.6	92.20	8.5
DNN	89	77	90	89.5	90	11
SVM	80.13	69	85	82.565	84.13	20
Naïve Bayes	84	72	88	86	86	15

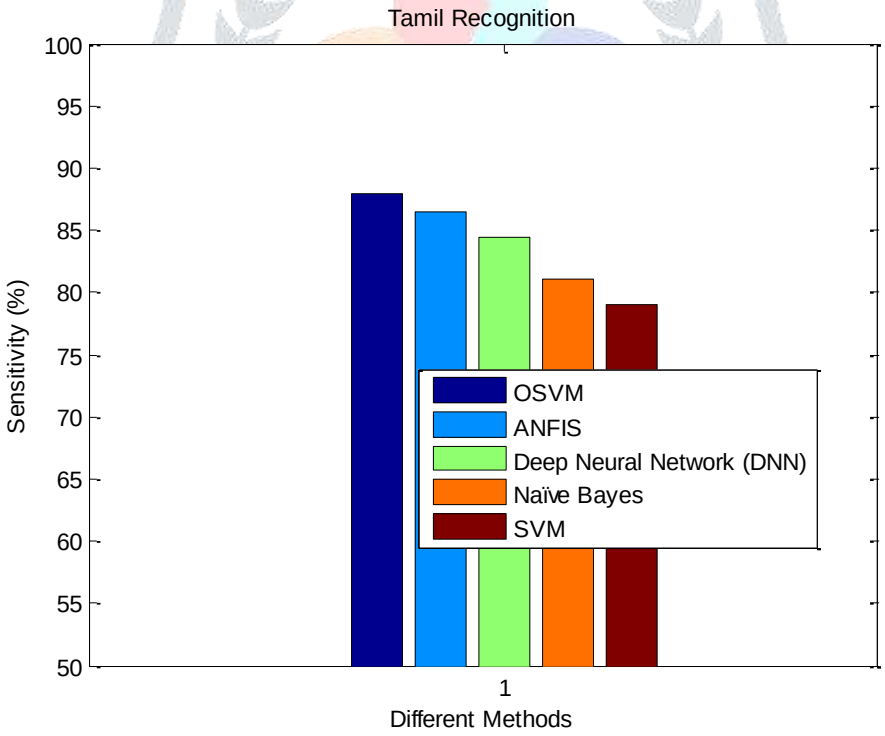

Figure 3. Sensitivity or recall comparison results of Tamil Handwritten recognition methods(Proposed OSVM)

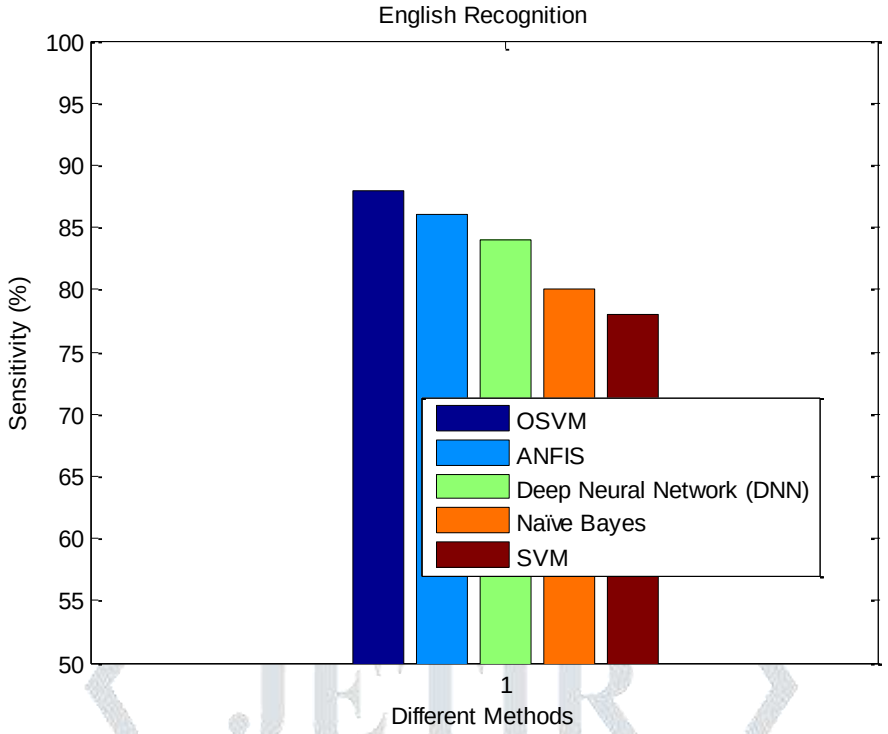

Figure 4. Sensitivity or recall comparison results of English Handwritten recognition methods (Proposed OSVM)

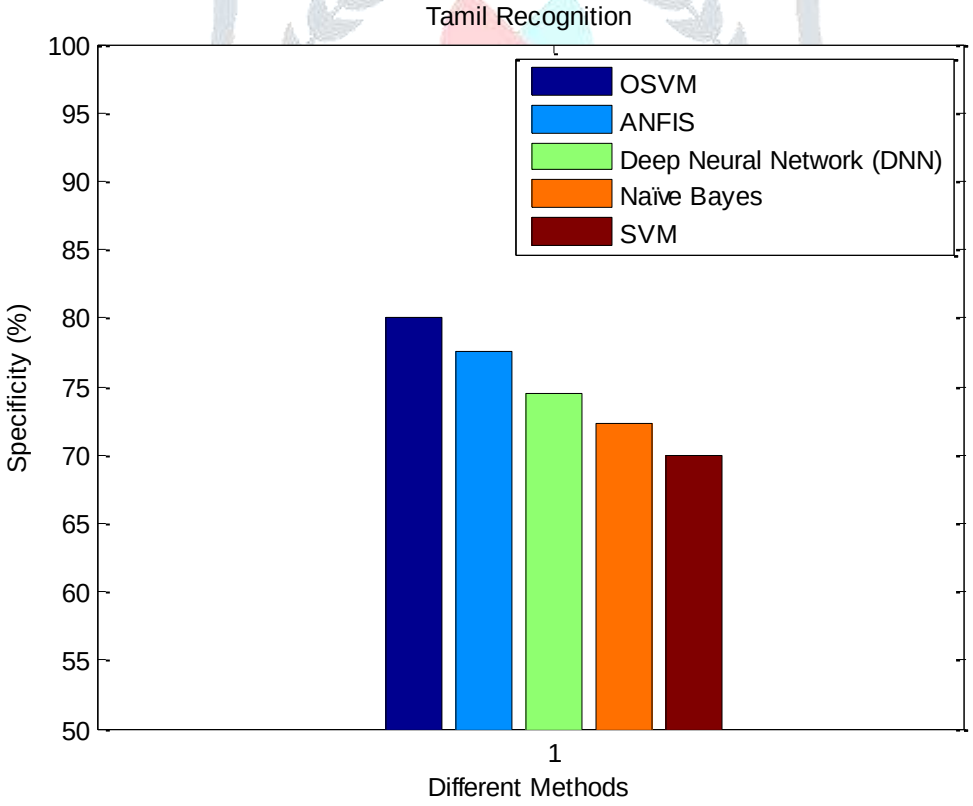

Figure 5. Specificity comparison results of Tamil Handwritten recognition methods(Proposed OSVM)

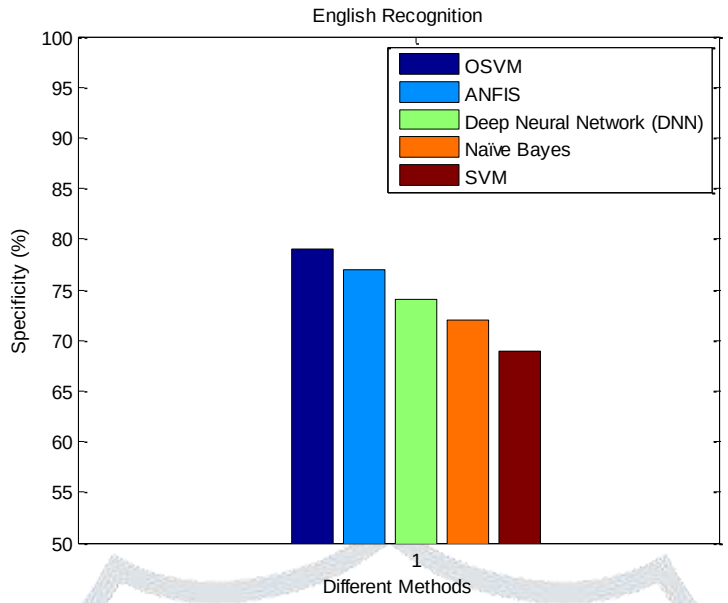

Figure 6. Specificity comparison results of English Handwritten recognition methods(Proposed OSVM)

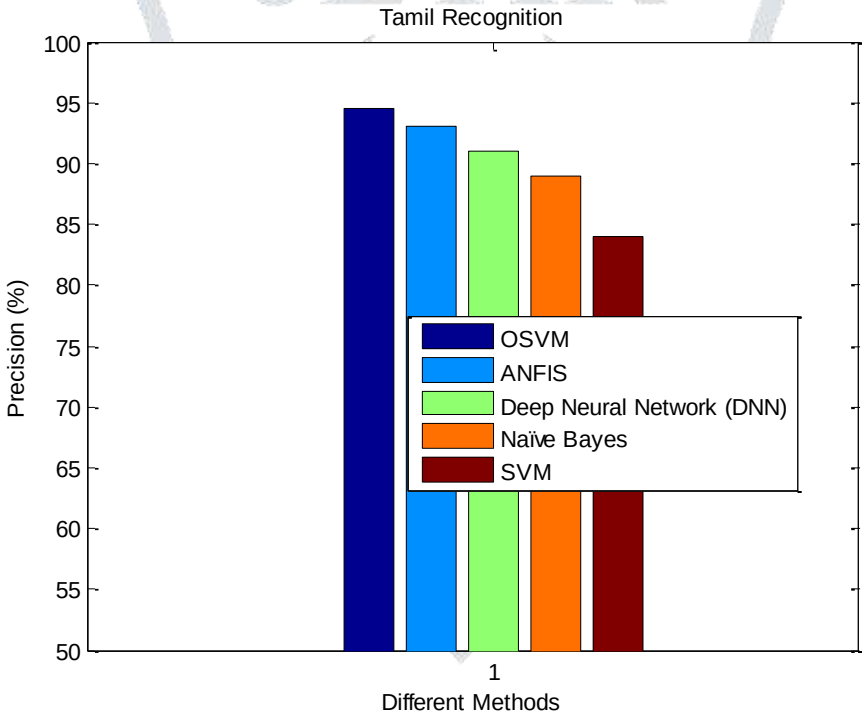

Figure 7. Precision comparison results of Tamil Handwritten recognition methods (Proposed OSVM)

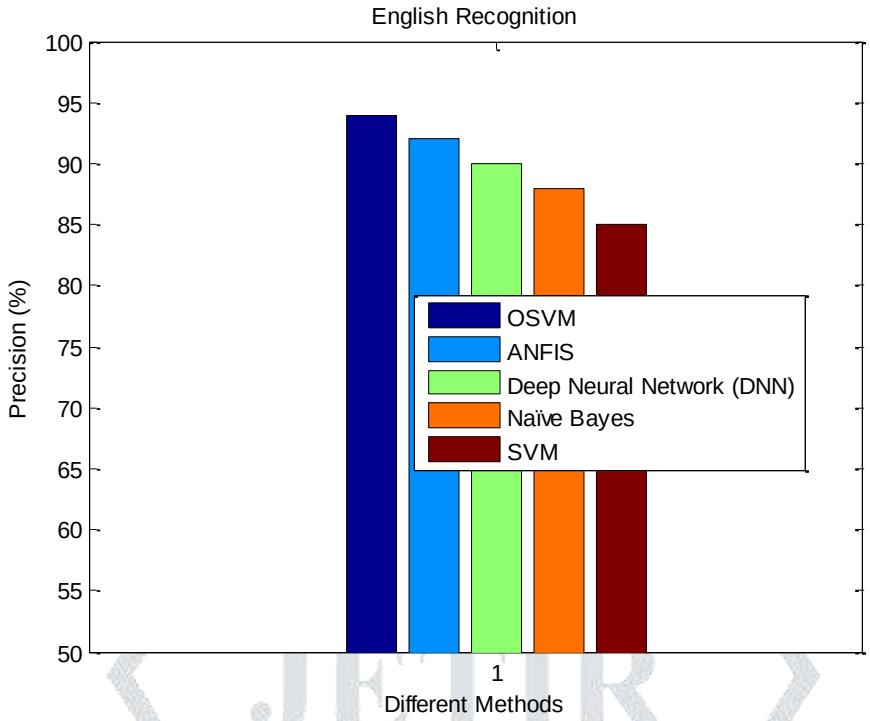

Figure 8. Precision comparison results of English Handwritten recognition methods(Proposed OSVM)

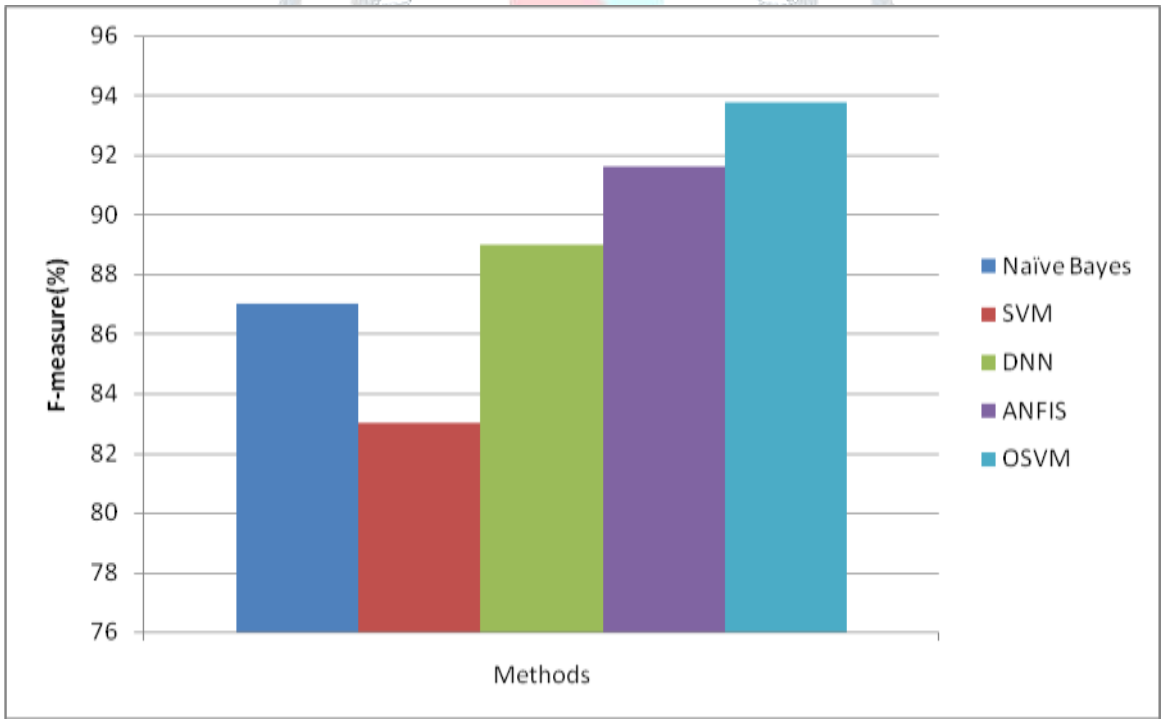

Figure 9. F-measure comparison results of Tamil Handwritten recognition methods(Proposed OSVM)

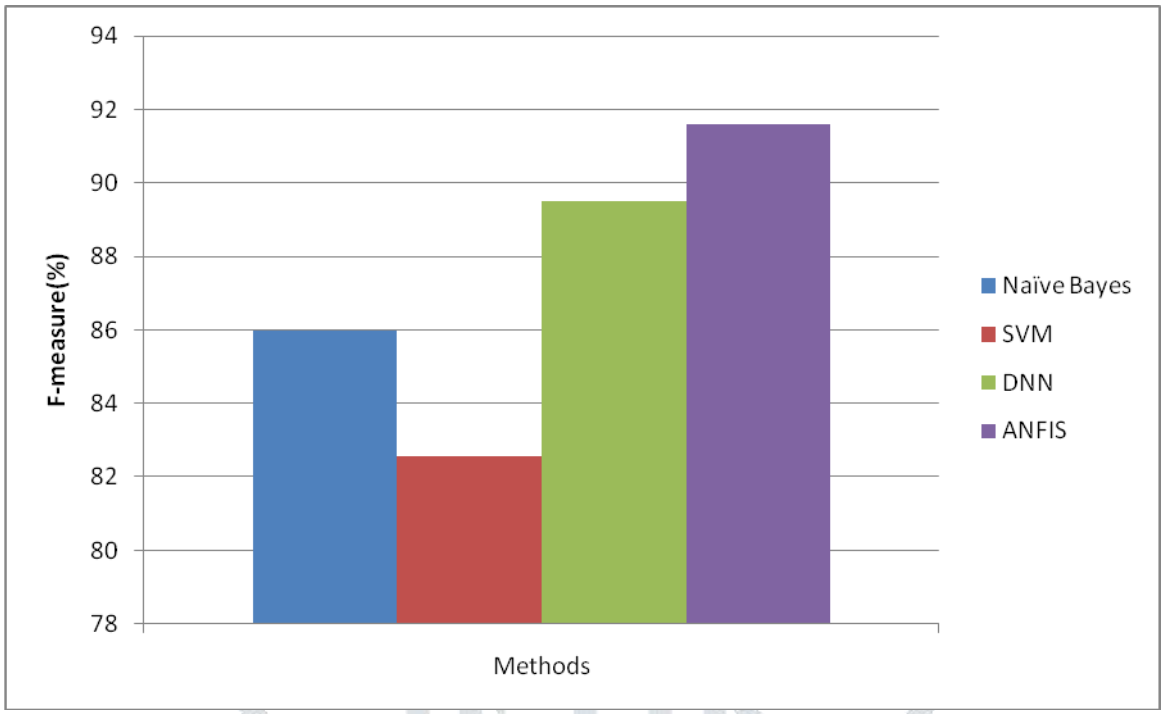

Figure 10. F-measure comparison results of English Handwritten recognition methods(Proposed OSVM)

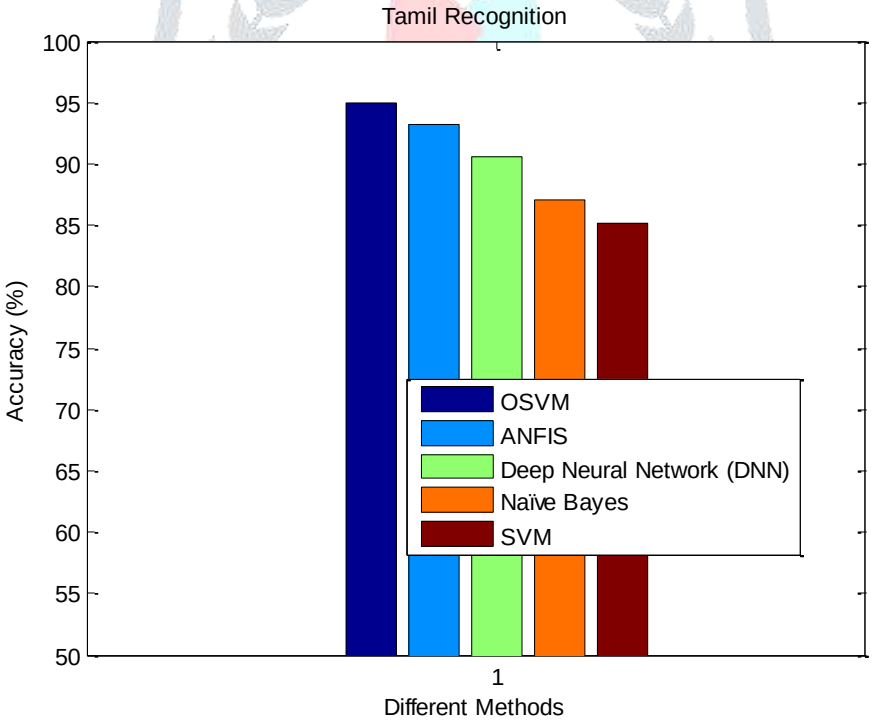

Figure 11. Accuracy comparison results of Tamil Handwritten recognition methods(Proposed OSVM)

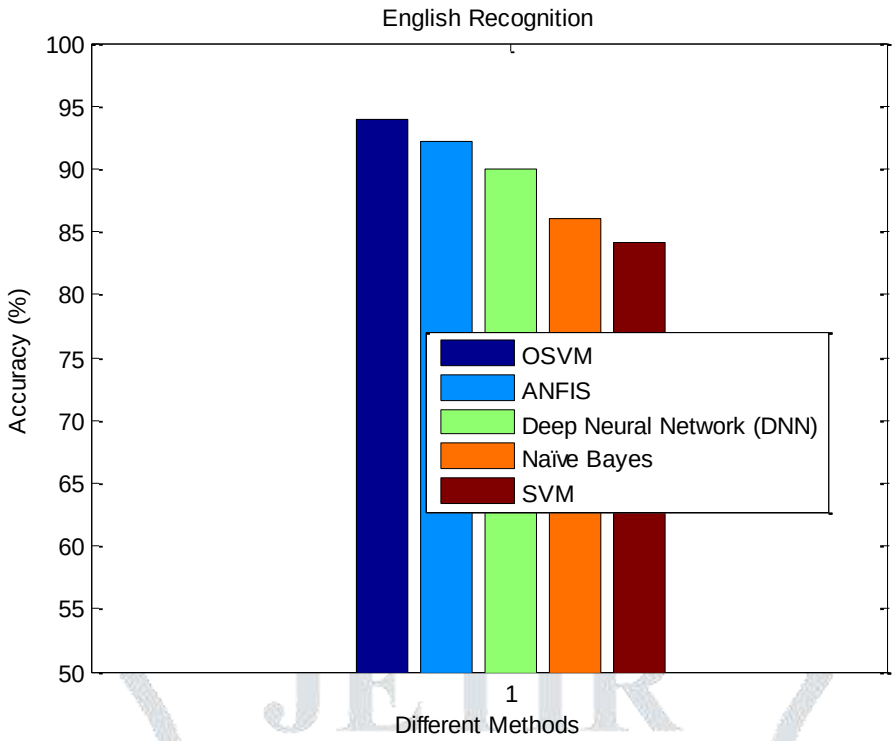

Figure 12. Accuracy comparison results of English Handwritten recognition methods(Proposed OSVM)

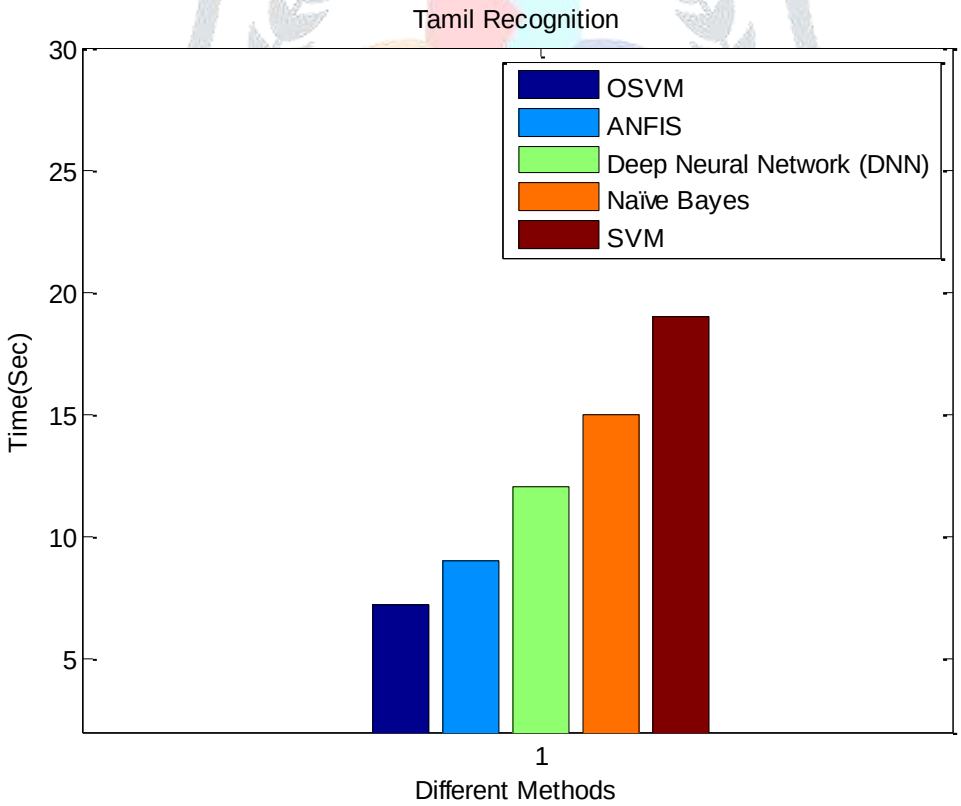

Figure 13. Classification time results of Tamil Handwritten recognition methods(Proposed OSVM)

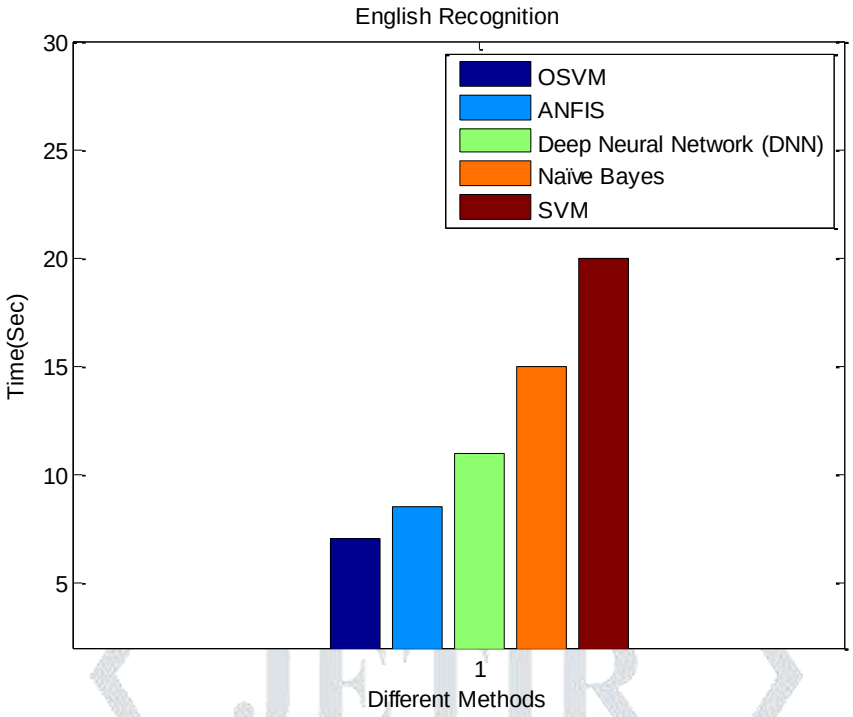

Figure 14. Classification time results of English Handwritten recognition methods (Proposed OSVM)

Figure 3, Figure 5, Figure 7, Figure 9, Figure 11 and Figure 13 shows the results of sensitivity, specificity, precision, F-measure, accuracy and classification time of Tamil handwritten documents with recognition methods. From the results it concludes that the proposed **OSVM** classifier performs better for all metrics and takes lesser computation time when compared to other three classifiers. Figure 4, Figure 6, Figure 8, Figure 10, Figure 12 and Figure 14 show the results of sensitivity, specificity, precision, F-measure, accuracy and classification time of English handwritten documents with recognition methods. From the results it concludes that the proposed **OSVM** classifier performs better for all metrics and lesser computation time when compared to other three classifiers. The values of these metrics are discussed in Table 2 and Table 3.

5. CONCLUSION

This paper elucidates the various phases of recognition of off-line recognition system. By the means of noise elimination, binarization and size normalization, the noise is eliminated from the raw data in the pre-processing phase. The Segmentation of characters and words has been done through the BAT algorithm. To determine the boundaries for recognition, the edge detection is done through the descriptor. The feature extraction has been performed through the structural and statistical methods where the feature set of words has been extracted. Fuzzy clustering is the algorithm used for identification of style of words and characters. The Online Support Vector Machine algorithm (OSVM) is the classifier installed in this work where efficient recognition of words is done. Then the final phase of post-processing eliminates the errors and increases its accuracy rate.

REFERENCES

1. Shalin A. Chopra, Amit A. Ghadge, Onkar A. Padwal, Karan S. Punjabi, Prof. Gandhali S. Gurjar, Optical Character Recognition, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 1, pp.4956-4958, 2014.
2. Reddy, K.V., Rao, D.R., Ankaiah, U. and Rajesh, K., 2013. Handwritten Character and Digit Recognition Using Artificial Neural Networks. International Journal of Advanced Research in Computer Science and Software Engineering, 3(4), pp.140-143.

3. Vithlani, P. and Kumbharana, D.C., 2015. A Study of Optical Character Patterns identified by the different OCR Algorithms. International Journal of Scientific and Research Publications, 5(3), pp.2250-3153.
4. Bobade, H. and Sahu, A., 2013. Character Recognition Technique using Neural Network. International Journal of Engineering Research and Applications, 3(2), pp.1778-1783.
5. M. Oujaoura, R. El Ayachi, O. Bencharef, Y. Chihab, B. Jarmouni, Application of Data Mining Tools for Recognition of Tifinagh Characters, IJACSA Special Issue on Selected Papers from Third international symposium on Automatic Amazigh processing (SITACAM), pp.1-4,2013.
6. Al-Maadeed, S., Ferjani, F., Elloumi, S. and Jaoua, A., 2016. A novel approach for handedness detection from off-line handwriting using fuzzy conceptual reduction. EURASIP Journal on Image and Video Processing, 2016(1), pp.1-14.
7. Liwicki, M. and Bunke, H., 2006, HMM-based on-line recognition of handwritten whiteboard notes. In Tenth international workshop on frontiers in handwriting recognition. Suvisoft.
8. Dedgaonkar, S.G., Chandavale, A.A. and Sapkal, A.M., 2012. Survey of methods for character recognition. International Journal of Engineering and Innovative Technology (IJEIT), 1(5), pp.180-189.
9. Shah M, Sahil Mehta, Pranav Mody, Akhil Sen Roy, Sunil P. Khachane, Handwriting Recognition of Diverse Languages, International Journal of Computer Science and Mobile Computing(IJCSMC), vol.7, no.4, pp.109-114, 2018.
10. Dash, T. and Nayak, T., 2013. Non-Correlated Character Recognition using Artificial Neural Network. arXiv preprint arXiv:1306.4629.
11. Bansal, K. and Kumar, R., 2013. K-Algorithm A Modified Technique for Noise Removal in Handwritten Documents. arXiv preprint arXiv:1306.1462.
12. Ntogas, N. and Veintzas, D., 2008, A binarization algorithm for historical manuscripts. In WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering (Vol. 12, pp. 41-51). World Scientific and Engineering Academy and Society.
13. Kumar, G., Bhatia, P.K. and Banger, I., 2013. Analytical review of preprocessing techniques for offline handwritten character recognition. International Journal of Advances in Engineering Sciences, 3(3), pp.14-22.
14. Kora, P. and Kalva, S.R., 2015. Improved Bat algorithm for the detection of myocardial infarction. SpringerPlus, 4(1), pp.1-18.
15. Jeyasingh, S. and Veluchamy, M., 2017. Modified bat algorithm for feature selection with the wisconsin diagnosis breast cancer (WDBC) dataset. Asian Pacific journal of cancer prevention: APJCP, 18(5),pp.1257–1264.
16. Tan, G., Viard-Gaudin, C. and Kot, A., 2008, Online writer identification using fuzzy c-means clustering of character prototypes. In ICFHR (pp. 475-480).
17. Davy, M., Desobry, F., Gretton, A. and Doncarli, C., 2006. An online support vector machine for abnormal events detection. Signal processing, 86(8), pp.2009-2025.
18. C. Gruber, T. Gruber, S. Krinniger and B. Sick, “Online Signature Verification With Support Vector Machines Based on LCSS Kernel Functions,” in IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), vol.40, no.4, pp.1088-1100, Aug. 2010.
19. Shanjana C, James, Ajay,” Offline Recognition of Malayalam Handwritten Text” Procedia Technology 19 (2015) 772 – 779.