A SURVEY ON DIGIT RECOGNITION USING DEEP LEARNING

[1]Aditi M Joshi, [2]kinjal Thakar,
[1]Student, [2]Assistant Professor
[1]Information Technology Department,
[1] Silver Oak College of Engineering & Technology, GTU, Ahmedabad, India

Abstract— Artificial Intelligence nowadays is so trendy, especially the neural networks, which are proving successful in a number of areas such as character recognition and speech recognition etc. This paper briefly describes how deep learning techniques used in digit recognition. It further describes the greatest advances that have occurred in the field of Medicine, forensic document analysis, number plate recognition, Banking, Healthcare, and legal industry. Digit recognition is one of the focused area of research in the field of Artificial Intelligence. Recognition of individual Digit have greater recognition accuracy than the recognition of characters. This paper presents a comprehensive review of Hindi and Sanskrit Digit Recognition using Deep learning approach.

Keywords - Deep Learning; Image Classification; Digit Recognition; Shape Based Recognition; (Half-Unit Biased) Hub; Hidden Markov Models(HMM); Pattern Recognition, Freeman Chain Code; Image processing; Convolutional Neural Networks; Power Quality.

I. INTRODUCTION

Deep learning has become the most popular approach to develop Artificial Intelligence (AI) machines that perceive and understand the world. Here a survey of deep learning methods aimed at image classification and number recognition in images is represented. Deep learning uses neural networks that pass data through many processing layers to interpret data features and relationships. Deep learning Algorithms are largely self-directed on data analysis once they are put into production. Researchers have experienced lots of ups and downs, while the first works on artificial neural networks have always been of special interest for researchers. Neural networks based methods have been successfully applied to classification, clustering, forecasting, approximation and recognition problems in medicine, biology, commerce, robotics etc. The latest advancement in the field has been caused by the invention of deep learning methods, induced by the progress of parallel computing hardware and software.

For processing numerical data, the human brain is so sophisticated that we recognize objects in a few seconds, without much difficulty. Computer vision is more ambitious. It tries to mimic the human visual system and is often associated with scene understanding. Most image processing algorithms produce results that can serve as the first input for machine vision algorithms. Image processing is a logical extension of signal processing.

When an unknown animal is encountered, we try to recognize it by comparing its features (called patterns) with known stored patterns that we already have. This process of comparing an unknown object with stored patterns to recognize the unknown object is called classification. Thus, classification is the process of applying a label or pattern class to unknown instance. In the absence of any prior knowledge of the object or stored pattern, we use a trial and error process to recognize the object. This trial and error process of grouping of objects is called clustering.

Most of the organization come across documents which are handwritten such as a forms, checks etc. Handwritten documents are further converted and stored in digital formats for easier retrieval. Without handwritten character recognition software, origination would require hiring dedicated employees for converting handwritten document in to digital format by manual entry of text. Nowadays there is a huge demand of storing the information available in paper documents in to a computer storage disk and then reusing this information by searching process. One way to store paper document information to computer system is scan the document. But the scanned image can't be edited by user. Character recognition technique is used to identify handwritten characters. For e.g. Guessing of prescribed medicine's name and does not assure the exact name because handwriting of doctors.

Handwritten character recognition can be classified as On-line handwriting recognition and Offline Handwritten character. On-line handwriting recognition involves the automatic conversion of text as it is written on a special digitizer or PDA, where a sensor picks up the pen-tip movements as well as pen-up/pen-down switching. This kind of data is known as digital ink and can be regarded as a digital representation of handwriting. Offline Handwritten character involves, from image automatic conversion of text into letter codes which are usable within computer and other text processing applications.

Handwritten character recognition system consists of pre-processing, segmentation, feature extraction, training and recognition and post processing. Handwritten character recognition system consists of pre-processing, segmentation, feature extraction, training and recognition and post processing.
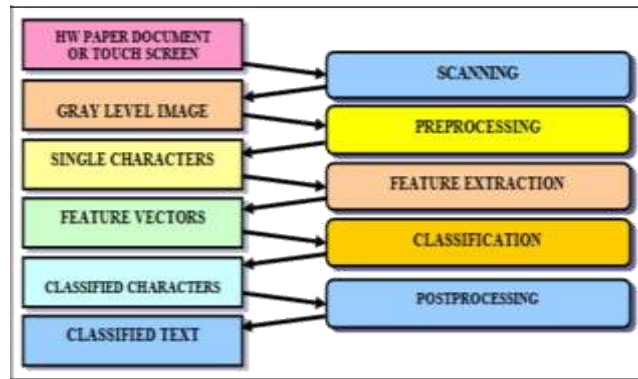
Figure. 2 Steps for Recognition

Freeman Chain Code (FCC) introduced by Freeman, 1961, the evolution and improvement of chain code representation scheme has been widely used as a topic of research [9]. Freeman codes or chain code are used to represent the boundary of the image. There are two types of chain codes corresponding to 4-neighbourhoods and 8-neighbourhoods. The boundary can be traced and allotted numbers based on the directions. Freeman chain code is the best proven method for number recognition.
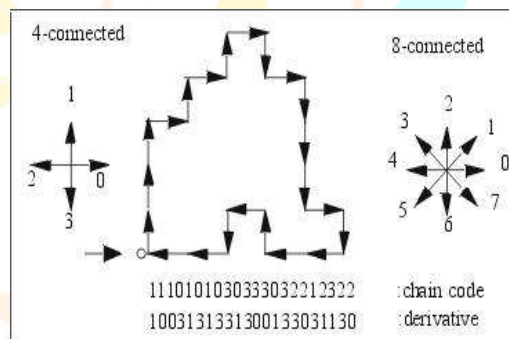

Figure. 3 Image chain code

a) 4- Directional code b) 8 – Directional code
The basic principle of chain codes is to separately encode each connected component in the image where chains represent the border of the objects.
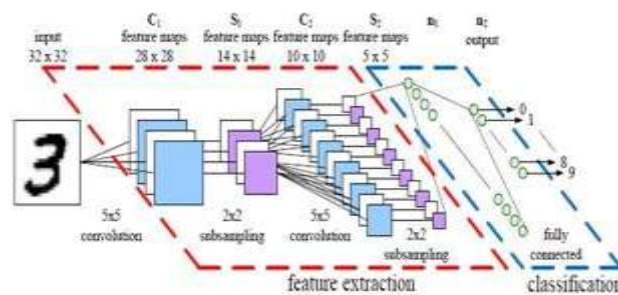

Figure. 4 Architecture of Convolutional Neural Networks

Neural networks also recognize patterns with extreme variability (such as handwritten characters), and with robustness to distortions and simple geometric transformations. Generating rules in neural networks is not a straightforward process. The inputs to neural networks should be numeric. Other data types should be converted to a number first. Convolutional Neural Networks (CNN) are a special kind of multi-layer neural networks that are designed to recognize visual patterns directly from pixel images with minimal preprocessing. Convolutional Neural Networks (CNN) also used for speech recognition. Deep Convolutional Neural Network (CNN) have shown superior results to traditional shallow networks in many recognition tasks [4]. Deep Neural Networks do not require any feature to be explicitly defined, they work on the raw pixel data generating the best features and using them to classify the inputs into different classes. Neural networks are robust in handling noisy data and improve their performance gradually. They have a low error-rate and high accuracy.LeNet-5 was designed by Lecun et al and is the latest convolutional network designed for handwritten and machine-printed character recognition. Lenet – 5 has 7 layers (2 Convolutional layers, 2 Subsampling layers, 2 hidden layers and 1 output layer) and not work on large images. It is one of the first shallow Convolutional neural network specifically designed to classify handwritten digit. It is trained and tested on the MNIST data set to classify the input into one of the ten classes representing 0-9 digits.

II. STRUCTURE OF CHARACTER RECOGNITION SYSTEM

Character Recognition system is the process of recognizing character from the image of printed document or the image of handwritten document. The major problem in handwritten character recognition is variety of handwriting styles, which is completely different from writer to writer. The block diagram of the character Recognition system is shown in Figure 5. Classification is the supervised learning method. In the training phase, the classifier algorithm is fed with a large set of known data. This dataset is called training data or labeled data.
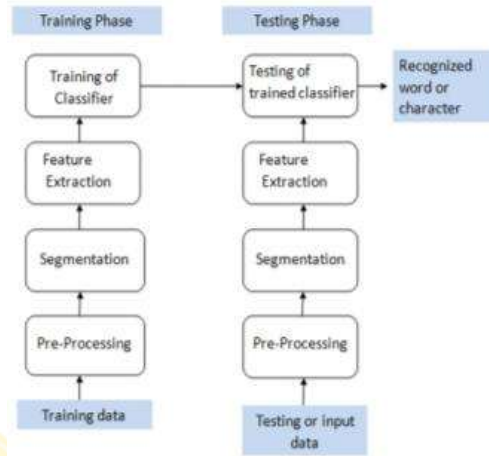

Figure.5 Block diagram of Character Recognition System [7].

The training data should generally large and representative in nature. Collected datasets are divided in to two parts Training data and testing data. For train the system training data are used. And this trained system further used to recognize test data. Generally, the performance of the classifier depends on factors such as the nature of data and the nature of learning.

A. Preprocessing

This step involves a basic processing of the image before it is used for recognition by the system. It has to be processed in such a way that it is suitable for the system to understand. The steps involved are:

Noise removal

There are several reasons due to which noise gets added in the images. Noise removal is the process of removing noise from scanned image by using appropriate filter. Smoothing is used for reduce noise and removal of small details from the image extracting large object. It could be from the mechanism that is used to acquire the image, the film grain or the electronic transmission of the image. The noise can be removed by linear filtering, median filtering or adaptive filtering.

Binarization

Binarization converts a gray scale image into a binary image using global thresholding technique like Otsu's method of thresholding. Otsu's method provides optimum value of threshold. Binarization is a process where each pixel in the image is converted into one bit and assigned a value of 0 or 1 depending upon the mean value of all the pixels. If the value is greater than mean values, then value assigned is 1 else 0.
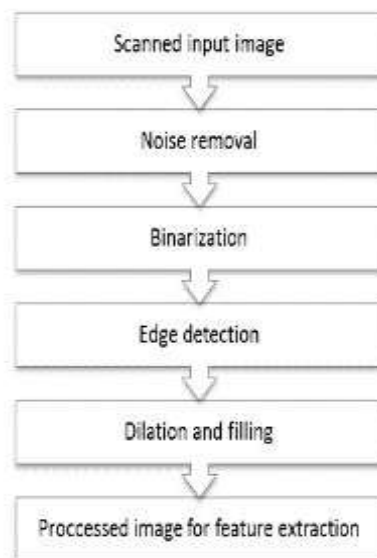

Figure.6 Steps for preprocessing [8].

Edge Detection

Edges play a very important role in many image processing applications. They provide an outline of the object. In the physical plane, edges correspond to the discontinuities in depth, surface orientation, change in material properties, and light variations. An edge is a set of collected pixels that lies on the boundary between two regions that differ in grey value. The pixels on an edge are called edge points. When edge is detected, the unnecessary details are removed, while only important structural information is retained. An edge is typically extracted by computing the derivative of the image function.

Dilation and filling

The basic effect of using dilation operator on binary image is to gradually enlarge the boundary of regions of the foreground pixels. Thus the areas of foreground pixels grow in size while holes within those regions become smaller.

Processed image for feature extraction

In this stage, the features of the characters that are crucial for classifying them at recognition stage are extracted. This stage is an important as its effective functioning improves the recognition rate and reduces the misclassification.

B. Segmentation

Image segmentation has emerged as an important phase in image based applications. Segmentation breaks the words or sentences in to pieces such that a clear boundary is set between the characters. Segmentation is the process of partitioning an image into multiple regions and extracting a meaningful region known as the region of interest. Region of interest (ROI) may vary with applications. For example, if the goal of doctor is to analyze the tumor in a computer tomography (CT) image, then the tumor in the image is the ROI. Similarity, if the image application aims to recognize the iris in eye image, then the iris in the eye image is the required ROI.

C. Feature Extraction

Image pre-processing and feature extraction techniques are mandatory for any applications that is based on image. Feature Extraction is a process of extraction and generation of features to assist the task of object classification. This phase is critical because the quality of the features influences the classification task. The extended set of features is stored as a vector called as the feature vector. The classifier takes the feature vector as input and performs the classification.

D. Classification

The classification stage is the decision making part of a recognition system and it uses the features extracted in the previous stage. The feature vector is denoted as X where $X = (f1, f2,....., fd)$ where f denotes features and d is the no. features extracted from character. Based on the comparison of feature vector characters are efficiently classified into appropriate class and recognized

E. Post – processing

In this stage accuracy of character recognition is further increased by connecting dictionary to the system in order to perform Syntax analysis, semantic analysis kind of higher level concepts, which is applied to check the recognized character. This stage is not mandatory in character recognition system.

III. RELATED WORK

Recognition of characters from image is depends on the sensitivity of selected features and which kind of classifier they used for image recognition. Following are the papers which perform character recognition with different algorithms and they also used different deep learning approaches for better accuracy and recognition. Freeman chain coding is used to create contour of the object in the image, which is than proceeds further for finding modified set of Fragments. For the purpose of scoring functio ns Modified Needleman – Wunsch algorithm is used [1].

Rounding number calculations with Hub formats. This is well-suited for Digital circuit equations. This paper has suggested the use of solving complex numerical issues for binary and numeric values processing [2].

Power quality waveforms are used while the image characters are wavy in format. In this papers possible PQ waveform deviation are found from internet. For this, previously recorded data and correct identification of current data is performed using signals. Google search engine tool perform its process with this format [3].

They introduce a new public image dataset for Devanagari script: Devanagari Handwritten Character Dataset (DHCD). In Devanagari script, the base form of consonant characters can be combined with vowels to form additional characters which is not explored in this research. Deep neural network was trained on the DHCD as a multi-class classification problem. Devanagari characters are least explored for image processing. Handwritten and printed both have different way to recognize them. Deep learning with CNN and multilayer perception network is used in this paper for accurate recognition of character [4].

Hidden markov model in machine learning is used to recognize the series of states from a series of observations. For example: We don't get to observe the actual sequence of states. Rather, we can only observe some outcome generated by each state. There can be Viterbi or Baysen model to find the maximum likelihood state assignment. Pattern of Devanagari is explored with hmm and ANN in this case to find out results of character recognition from an online isolated character set of Devanagari [5].

To process the images and characters from cheques, this paper has used k-space auto encoders. And the hidden layers of auto encoders are found with CNN for single digit analyzer. They have done modification to perform soft-tuning of the weights of

each stacked layer separately from other hidden layers. Further it can be improved to find modified process to improve the feature extraction process and representation efficiency [6].

They are going to take the MNIST dataset for training and recognition. The primary aim of this dataset is to classify the handwritten digits 0-9.They have a total of 70,000 images for training and testing. Each digit is represented as a 28 by 28 grey scale pixel intensities for better results. The digits are passed into input layers of LeNet and then into the hidden layers which contain two sets of convolutional, activation and pooling layers. Then finally it is mapped onto the fully connected layer and given a softmax classifier to classify the digits. They are going to implement this network using keras deep learning inbuilt python library [12].

The chain code descriptor of various hand shapes was calculated and tested with different classifier such as k-nearest-neighbour (KNN), Support vector machine (SVM) and Naive Bayes. Principal component analysis (PCA) was applied after the chain code description [14].

Table 1 Literature Survey

Author & Year	Approaches	Open issues	Accuracy
Ema Rachmawati, Masayu Leylia khodra, Iping Supriana (2016 | IEEE)[1]	1.Freeman chain code 2.Modified Needleman - Wunsch 3.Scoring Function 4.Shape Based Recognition	Predefined shape. Only characters are used. Length of chain code is ranging from 100 to 300 characters. Generating modified set of fragments is a lengthy process. Using polygonal shape, vectors are created which does not a give assurance of generation of exact image.	They are able to recognize the correct fruit object from the same variety, with Average recall ranging from 0.733 to 1,with detection rates at the standard 0.3/0.4 FPPI that range 0.3 to 1.
Javier Hormigo, Julio Villalba (2015 | IEEE)[2]	1.(Half Unit Biased) Hub Formats 2.Binary Canonical Round to nearest Representation	Implement Systems for computation with real number under round to nearest.	It is clearly observed that the three cases of addition using HUB numbers produce less error thanusingRN-representation. The best result is produced when adding HUB numbers and the output is provided in a conventional format.
Leandro Rodrigues Manso Silva, Paulo Fernando Ribeiro (2016 | IEEE)[3]		Using powerful tool, here named iPQ Google, and which has the possibility to help utilities, customers and researchers to investigate and easily find, compare and diagnose possible PQ waveform deviations saved on the internet.	The preliminary results show the possibility of using the iPQ-Google to identify and classify electrical disturbance.

Shailesh Acharya, Ashok Kumar Pant, Prashnna Kumar Gyawali (2015 | IEEE)[4]	1.CNN or ConvNet (convolutional Neural Network) 2.Deep CNN	Explore the challenges in recognition of Devanagari characters.	Using Deep Cnn Increases test accuracy by nearly 1 percent. Architecture scored highest test accuracy of 98.47% on our dataset.

Author & Year	Approaches	Open issues	Accuracy
Ashish Tripathi, Sankha Subhra Paul, Vinay Kumar Pandey[5]	Hidden Markov Models (HMM) Artificial Neural Network (ANN)	Iphone does not support devanagri script.Training and recognition of isolated Devanagari character on the basis of the HMM Classes.	A few character among 20 most accruing character are taken has samples and their manual study generates certain stroke classes.
Raid Saabni (2015 | IEEE)[6]	1. Neural Network 2.Back propagation 3.Sparse auto encoders	Algorithm that trains k-sparse auto encoders and used their hidden layers to be stacked as pre trained hidden layers into a deep neural network. Trained the system on an extended version of the MNIST and CVL benchmarks and evaluated it in recognizing single handwritten digits and digit strings from CVL and ORANDCAR data sets.	Results in terms of error rate, show that extending the size of the training set improves recognition rates witnessed by the better results rather than the previous one.

IV. CONCLUSION

Deep learning have shown superior results to traditional approaches in many recognition tasks. Neural networks are robust in handling noisy data, they have low error rate and higher accuracy. Layers in deep learning creats an artificial neural network that can learn and make intelligent decision on its own. Deep learning has applied to so many applications like Handwritten character recognition, Machine printed character recognition, signature verification etc. deep learning's primary motive is to develop a network which is better to every performance measure and provide results for all kinds of datasets which can be trained and recognized. Word based recognition methods is somewhat similar to character recognition, Sentences have higher variability than single character or digit. In image processing, Recognition of characters or sentences is tougher task than recognition of single character or digit. Template matching, statistical methods, structural methods, neural networks etc. are the various recognition methods that is used for character recognition in image processing.

V. ACKNOWLEDGMENT

REFERENCES

[1] Ema Rachmawati, Masayu Leylia khodra, Iping Supriana,School of Electrical Engineering and Informatics Institut Teknologi Bandung,"*Shape based recognition using Freeman chain code and Modified Needleman - Wunsch* " 978-1-5090-4139-8/16/$31.00 ©2016 IEEE.

[2] Hormigo, J. and Villalba, J. (2016). "*New Formats for Computing with Real-Numbers under Round-to-Nearest.*" IEEE Transactions on Computers, 65(7), pp.2158-2168.

[3] Leandro Rodrigues Manso Silva, Paulo Fernando Ribeiro, Federal University of Juiz de Fora Juiz de Fora – MG, Brazil ,"*Power Quality Waveform Recognition Using Google Image Search Engine (iPQ-Google)*" 978-1-5090-3792-6/16/$31.00 ©2016 IEEE.

[4] Shailesh Acharya , Ashok Kumar Pant , Prashnna Kumar Gyawali, Institute Of Engineering Tribhuvan University Kathmandu, Nepal,"*Deep Learning Based Large Scale Handwritten Devanagari Character Recognition*" 978-1-4673-6744-8/15/$31.00 ©2015 IEEE

[5] Ashish Tripathi, Sankha Subhra Paul, Vinay Kumar Pandey, CSED, MNNIT, Allahabad, India "*Standardization of stroke order for online Isolated Devanagari Character Recognition for iPhone*"

[6] Raid Saabni School of Computer Science Tel-Aviv Yaffo Academic College, Tel-Aviv, Israel Triangle Research & Development Center.*"Recognizing handwritten Single Digits and Digit Strings Using Deep Architecture of Neural Networks*" 978-1-4673-9187-0 ©2016 IEEE

[7] Monica Patel, Shital P. Thakkar,*"Handwritten Character Recognition in English: A Survey"* International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015

[8] Hirali S. Amrutiya, Payal P. Moghariya and Vatsal H. Shah *"handwritten character recognition"* International Journal of Advance Engineering and Research Development (IJAERD) Special Issue, Volume 1,Issue 4, April 2014, e-ISSN: 2348 – 4470.

[9] Ibrahim, Ratnawati & Abu Hassan, Mohd Fadzil & Zainudin, Nurnashriq & Syahidi Aisha, Mohd. (2011). "*Number Recognition System Using Chain Code Technique*" December 2011.

[10] Dhannoon, Ban N. & H. Al, Huda. (2013). *"Handwritten* Innovation and Applied Studies. *Hindi Numerals Recognition*." International Journal of

[11] Pulipati Annapurna, Sriraman Kothuri, and Srikanth Lukka, "*Digit Recognition Using Freeman Chain Code*" International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 8, August 2013.

[12] T Siva Ajay,*"Handwritten Digit Recognition Using Convolutional Neural Networks"* International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 07, July -2017.

[13] Apoorva Chaudhary and Mr.Roshan lal chhoker ,"*Handwritten Hindi Numeric Character Recognition and comparison of algorithms"* 978-1-5090-3519-9/17/$31.00©2017 IEEE

[14] Fating, Kshama & Ghotkar, Archana. (2014).*"Performance Analysis of Chain Code Descriptor for Hand Shape Classification."* International Journal of Computer Graphics & Animation. 4. 9-19. 10.5121/ijcga.2014.4202.

[15] T.N. Sainath, A.–R. Mohamed, B. Kingsbury and B. Ramabhadran, "Deep convolutional Neural Networks for LVCSR," IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP), pp. 8614-8618, 2013

[16] H. Larochelle, Y. Bengio, J. Louradour and P. Lamblin "Exploring Strategies for Training Deep Neural Networks," The Journal of Machine Learning Research Vol. 10, pp. 1-40, 2009

[17] D. Erhan, Y. Bengio, A. Courville, P. Manzagol, P. Vincent and S. Bengio, "Why Does Unsupervised Pre-training Help Deep Learning?" The Journal of Machine Learning Research, Vol. 11, pp. 625-660, Feb 2010

[18] S. R. Fischer, "A History of Writing," Reaktion Books, 2004

[19] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," In Proc. 27th International Conference on Machine Learning, 2010

[20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov "Dropout: A simple way to Prevent Neural Networks from Overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929-1958, 2014

[21] R. Jayadevan, Satish R. Kolhe, Pradeep M. Patil and Umapada Pal, "Offline Recognition of Devanagari Script: A Survey," IEEE transactions on systems, man, and cybernetics-part c: applications and reviews,vol. 41, no. 6, November 2011, pp. 1094-6977.

[22] Rachid Hedjam, Hossein Ziaei Nafchi, Margaret Kalacska, and Mohamed Cheriet, "Influence of Color-to-Gray Conversion on the Performance of Document Image Binarization: Toward a Novel Optimization Problem," IEEE transactions on image processing, vol. 24, no. 11, November 2015, pp. 1057-7149.

[23] Inderpreet Kaur, Saurabh Mahajan, "Bilingual Script Identification of Printed Text Image," Volume: 02 Issue: 03- June-2015.

[24] Umapada Pal, Ramit Kumar Roy, Fumitaka Kimura, "Bangla and English City Name Recognition for Indian Postal Automation," IEEE DOI 10.1109, pp. 1051-4651/10.

[25] Ying-Ho Liu, Fu Chang and Chin-Chin Lin, "Language Identification of Character Images Using Machine Learning Techniques," Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition IEEE.

[26] Pankaj Kumar, Sheetal Srivastava, Monica Joshi, "Syntax Directed Translator for English to Hindi Language,"2015 IEEE International Conference on Research in Computational in Intelligent and Communication Networks.

[27] Shalini Puri, S.P.Singh, "A Technical Study and Analysis of Text Classification Techniques in N - Lingual Documents," 2016 International Conference on Computer Communication and Informatics, Jan. 07-09-2016, Coimbatore.