iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** XII **Month of publication:** December 2021

DOI: https://doi.org/10.22214/ijraset.2021.39703

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 9 Issue XII Dec 2021- Available at www.ijraset.com

Handwritten English Alphabet Recognition

Poonam Bhanudas Abhale[1], Dr. D. L. Bhuyar[2], Milind Khanderao[3]
[1]*PG. Student, Department of Electronics and Telecommunication CSMSS's CHH. Shahu College of Engineering,Aurangabad, India*
[2]*Associate Prof. Dr. D. L. Bhuyar, Department of Electronics and Telecommunication CSMSS's CHH. Shahu College of Engineering, Aurangabad, India*

Abstract: Character recognition is a process by which a computer recognizes letters, figures, or symbols and turns them into a digital form that a computer can use. In moment's terrain character recognition has gained a lot of attention in the field of pattern recognition. Handwritten character recognition is useful in cheque processing in banks, form recycling systems, and numerous further. Character recognition is one of the well- liked and grueling areas of exploration. In the unborn character recognition produce a paperless terrain. In this paper, we describe the detailed study of the being system for handwritten character recognition. We give a literature review on colorful ways used in offline English character recognition.
Keywords: Character; Character recognition; Preprocessing; Segmentation; Point birth; Bracket; neural network; Convolution neural network.

I. INTRODUCTION

Optic character recognition is the history when in 1929 Gustav Tauschek got a patent on OCR in Germany followed by Handel who attained a US Patent on OCR in USA in 1933. Since also number of character recognition systems have been developed and are in use for indeed marketable purposes also.[1] But still there's a stopgap to make some further intelligent hand written character recognition system because hand jotting differ from one person to other. [2]His jotting style, shape of rudiments and their sizes makes the difference and complexity to fete the characters. [3]Experimenters formerly paid numerous sweats in designing hand written character recognition system utmost of them cited as because of its important operation like bank checking process, reading postal canons and reading different forms [4]. Handwritten number recognition is still a problem for numerous languages like Arabic, Farsi, Chinese, English, etc [5]. A machine can perform further tasks than a mortal being in the same time;[6] this kind of operation saves time and plutocrat and eliminates the demand that a mortal perform such a repetitious task. [7]For the recognition of English handwritten characters, colorful styles have been proposed [8]. Also a many figures of studies have been reported for Farsi language [9]. In some hand- jotting, the characters are indistinguishable indeed to the mortal eye, and that [10] they can only be distinguished by environment. [11]In order to distinguish between similar analogous characters, the bitsy differences that they've must be linked. One of the major problems of doing this for hand written characters is that they don't appear at the same relative position of the letter due to the different proportions in which characters are written by different pens of the language. Indeed the same person may not always write the same letter with the same proportions. Then, the thing of a character recognition system is to transfigure a hand written textbook document on paper into a digital format that can be manipulated by word processor software. The system is needed to identify a given input character form by mapping it to a single character in a given character set. Each hand written character is resolve into a number of parts (depending on the complexity of the ABC involved) and each member is handled by a set of purpose erected neural network. The final affair is unified via a lookup table. Neural network armature is designed for different values of the network parameters like the number of layers, number of neurons in each subcaste, the original values of weights, the training measure and the forbearance of the correctness. The optimal selection of these network parameters clearly depends on the complexity of the ABC.

The first important step in any handwritten recognition system is pre-processing followed by segmentation and point birth. Pre-processing includes the way that are demanded to shape the input image into a form suitable for segmentation (8). In the segmentation, the input image . is segmented into individual characters and also, each character is resized into m x n pixels .towards the training network. The Selection of applicable point birth system is presumably the single most important factor in achieving high recognition performance. Several styles of point birth for character recognition are reported within the literature (9). The considerably used point birth styles are Template matching, Deformable templates, Unitary Image transforms, Graph . description, Projection Histograms, Contour lives, Zoning, Geometric moment invariants, . Zernike Moments, Spline wind approximation, Fourier descriptors, Grade point and Gabor Features.

II. SYSTEM DESIGN

The overview of the system implementation is as shown below:

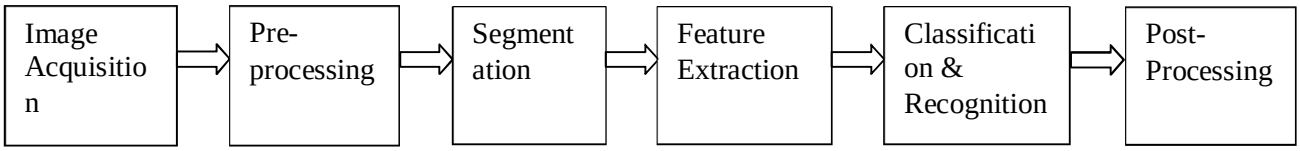


Fig.1: General Block Diagram

The steps involved in English Alphabet Recognition are:

1) *Step 1:* START
2) *Step 2:* Collect dataset i.e. English alphabets for training.
3) *Step 3:* Collect test patterns for testing.
4) *Step 4:* Apply preprocessing and extract features from collected samples.
5) *Step 5:* Define methods for training & testing using MATALB & other toolboxes.
6) *Step 6:* Training the network using defined algorithm.
7) *Step 7:* Repeat Step 5 if wants reducing the errors.
8) *Step 8:* Testing the pattern.
9) *Step 9:* Check recognized character.
10) *Step 10:* STOP

A. *Image Pre-processing*

Image pre-processing is a method where documents are processed to clean and bring them into desired format for further processing. It is series of operations. Pre-processing aims at improving data in given image by reducing or eliminating undesired distortions and enhancing important features in image.

Pre-processing is not a single step. It includes number of operations as mentioned below:

1) *Gray Scale Conversion:* The process, in which RGB image is converted to black & white, is called gray scale conversion. For binarization gray scaling is important because after doing these only gray shades remain in image and binarization of such images is efficient. Rgb2gray function is used to perform this operation. It eliminates hue and saturation and retains luminance.
2) *Binarization:* In binarization, gray image is converted into an image having pure black & pure white pixel values. In this process, pixels having intensities lower than half of full intensity are converted to black and remaining intensity values are converted to white pixels.
3) *Inversion:* In inversion, present image is just an inverse of previous one. That means, each pixel of image has colour or intensity value which is exact inverse of that of previous image. This is important because any character of image must only be extracted efficiently from its background. It is only required in case where objects of interest are of darker intensity on lighter background.
4) *Noise Removal:* Noise can occur at the time of capturing the image or image transmission. Noise is nothing but the unwanted data in input image that shows different intensities than true ones. This noise can be removed by using either of the following filtering:

a) By Mean Filtering
b) By Median Filtering
c) By Average Filtering
d) By Adaptive Filtering, etc

5) *Edge Detection:* Most of the information is enclosed in edges so edges are detected and by using filters these are enhanced. For this some masks such as prewitt, sobel, laplacain, etc are used which detect the horizontal, vertical edges, etc.

6) *Morphological Operations:* The reason behind morphological operations is filtering document by replacing convolution by logical operation. Dilation, erosion, filling, opening, closing, etc come under morphological operations. These operations smoothens contours, decay joined strokes, connect broken strokes, perform thinning, etc. hence morphological operations help in removing noise from images.These techniques use a structuring element. It is placed at all probable locations in image & it iscompared with the neighborhood of pixels. In dilation pixels are added to the object's boundary and image grows or thickens the object. Erosion removes pixels from object boundaries.Also, opening where erosion is followed by dilation & closing where dilation is followed by erosion are the typical morphological operations.

B. Feature Extraction

Feature extraction is the process where relevant information about different shapes present in the pattern is detected and put in a vector to use further for classification. Depending on the feature vectors system classifies the inputs with significant accuracy. So, the features extracted should be highly discriminative with reduced dimensions in order to reduce the computation requirements in classification. Feature extraction is done properly, and then it can also reduce the errors such as mean square error or inter-distance differences, etc. Each character has some features, which play an important role in pattern recognition. Indian Marathi language characters have many particular features. Feature extraction describes the relevant shape information contained in characters so that the task of classifying the character is made easy by a formal procedure. Feature extraction stage in Marathi language characters Character system analyses these Marathi language characters Character segment and selects a set of features that can be used to uniquely identify. Mainly, this stage is main part of system because output depends on these features.

1) *Zoning:* In the process of zoning, a binary image having Marathi language characters which are pre-processed and normalized to a size of 36 x 36 is partitioned into n number of equal-sized zones. Then features are extracted from the individual zone. The advantage of finding features from individual zone over that from the whole images is that it provides more detailed information regarding small and finer details in the skeleton of the character image. After calculating feature vectors from each zone, they are put in an array to make one feature vector representing the features for a given input image.
2) *Directional Features:* These features are extracted from the image skeleton, based on the line types forming the character skeleton. To do this, the image is zoned or partitioned into 3 x 3 sub-images and features are extracted from each of the zones.

C. Moment Invariant Features

These features are extracted from image skeleton, based on the line types forming the character skeleton. To do this, image is zoned or partitioned into 3 × 5 sub images and features are extracted from each of the zones.

Moment Invariant Features

These are the statistical features. In object recognition, moments play a significant role. This method measures the intensity function. It gives global character information.

III. CLASSIFICATIONS AND RECOGNITION

The bracket stage is the decision making part of a recognition system and it uses the features uprooted in the former stage. A feed forward back propagation neural network having two retired layers with armature of 54-100-100-38 is used to perform the bracket. The retired layers use log sigmoid activation function, and the affair subcase is a competitive subcase, as one of the characters is to be linked. The point vector is denoted as X where X = (f1, f2, …, fd) where f denotes features and d is the number of zones into which each character is divided. The number of input neurons is determined by the length of the point vectored. The total number of characters n determines the number of neurons in the affair subcase. The number of neurons in the retired layers is attained by trial and error. The most compact network is chosen and presented. The network training parameters are

1) Input bumps54/69
2) Retired bumps 100 each
3) Affair nodes38 (26 rudiments, 10 numbers, and 2 special symbols)
4) Training algorithm Grade descent with instigation training and adaptive literacy
5) Perform function Mean Square Error
6) Training Thing achieved0.000001
7) Training epochs1000000
8) Training instigation constant0.9.

IV. RESULTS

Following figure shows the home page of GUI (graphical User Interface) with various buttons on it that are used to perform various functions needed to recognize the character present in an uploaded image. Load Image will upload input image. Train button will train the neural network by using either of the two techniques. Extract Text will give characters extracted from given input image.

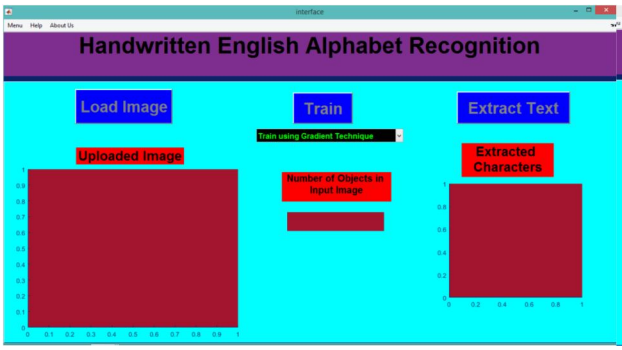


Fig.2: Home Page

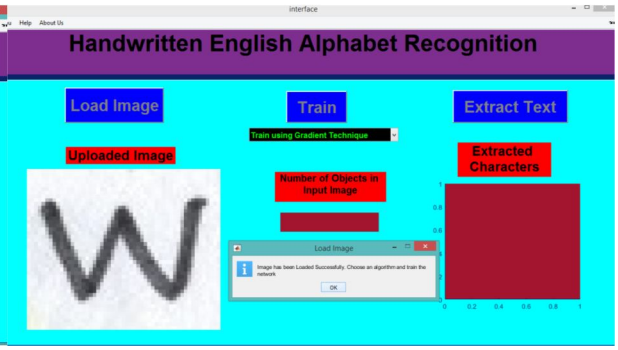


Fig.3: Uploaded Image

A. Uploaded Image

After clicking on button “Load Image”, any type of image which is of handwritten character, will be uploaded. After choosing the input image one pop-up message will be shown on screen telling that image is uploaded successfully and choose the algorithm for training the network.

B. Training Neural Network

After clicking on button “Train”, to train the network using features extracted by either of the two features extraction techniques, following window is showing on screen.

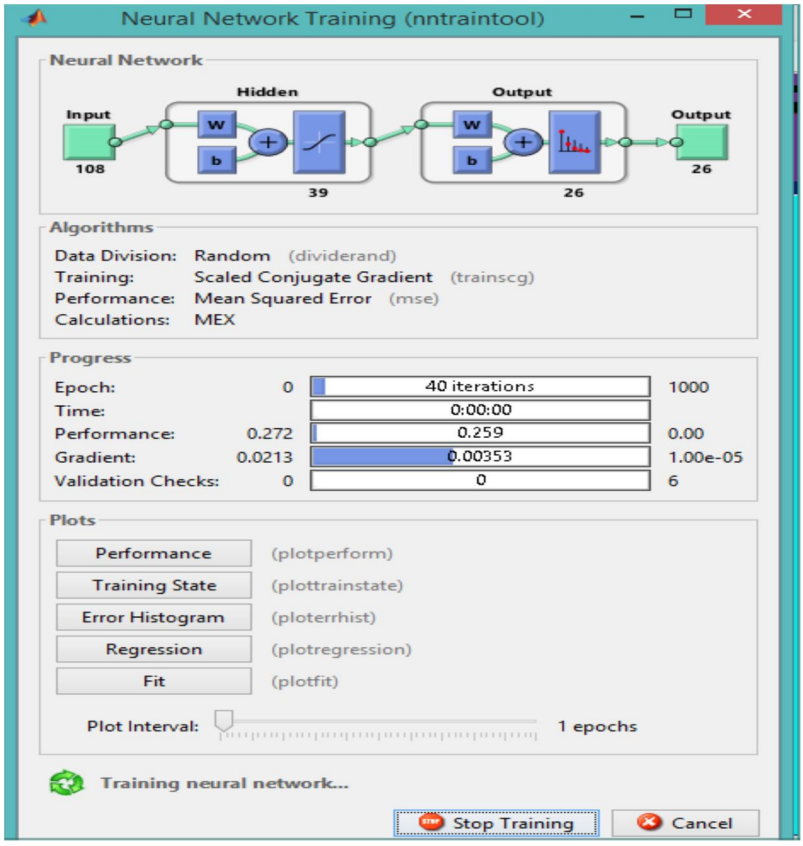


Fig.4: Training Neural Network

C. Extracted Output Character

By clicking "Extract Text" text from image will be shown in a box as shown in the figure 9.4. Also number of alphabets present in input image will be shown in box given.

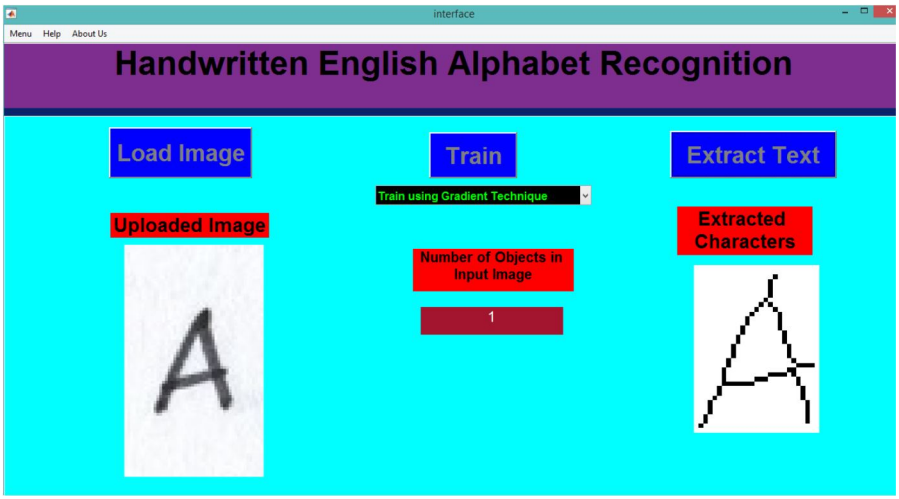


Fig.5: Character Extraction

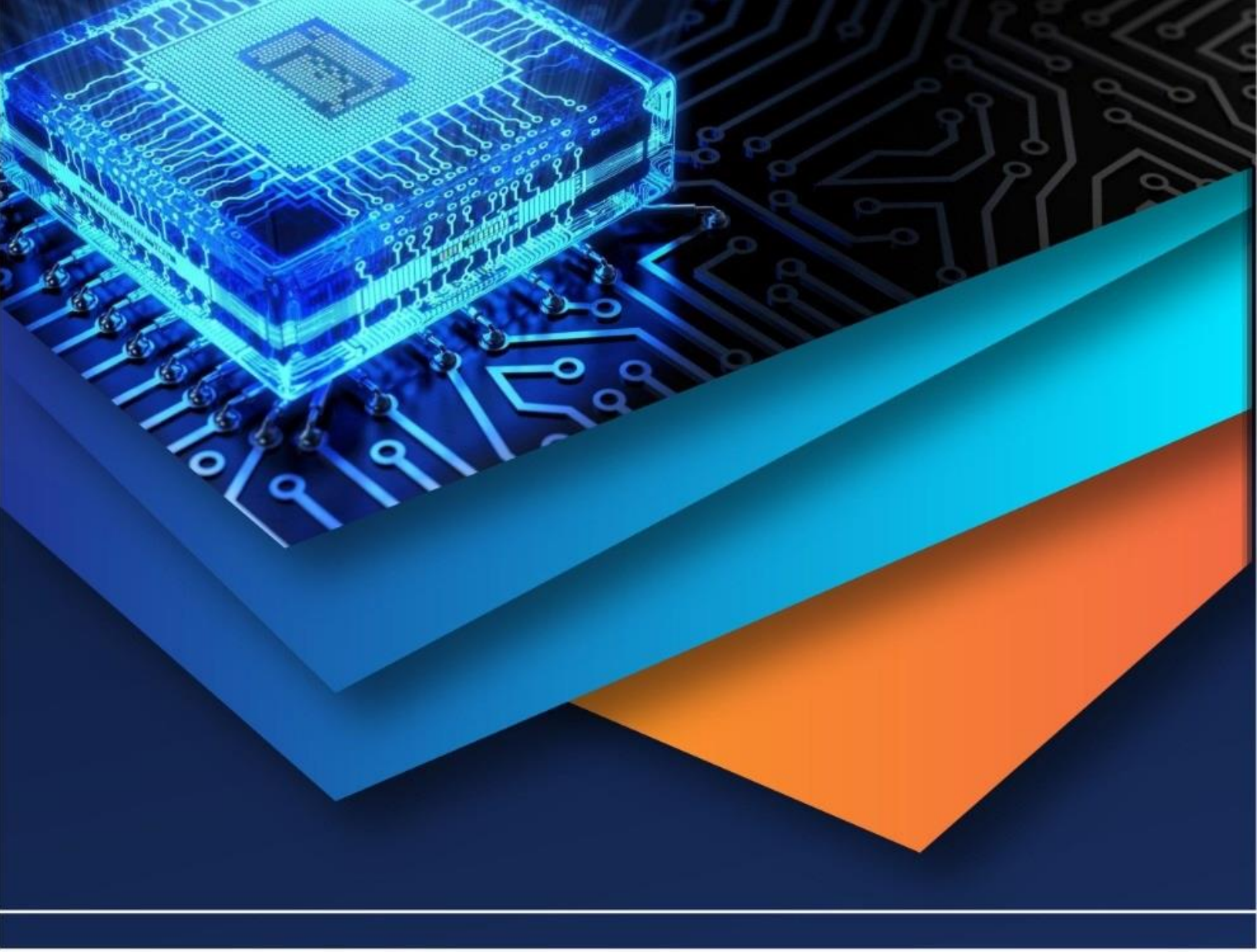
V. CONCLUSION

This project is intended to build a system that recognizes the Handwritten English alphabet. This method produces good results for handwritten text images which are written by different persons in different handwriting styles. Neural networks are commonly used for character recognition because of their property to tolerate high noise. Extracting features is one of the most essential steps in the character recognition domain. A strongly chosen feature set provides a good recognition rate. Without effective pre-processing, the extracted features of images are of low quality. The gradient feature extraction technique is highly accurate than the geometric feature extraction technique. Gradient feature extraction technique is faster to process high-end data.

REFERENCES

[1] S. A. Yadav, S. Sharma and S. R. Kumar, "A robust approach for offline English character recognition," 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), Noida, 2015, pp. 121 126.
[2] G. Katiyar and S. Mehfuz, "MLPNN based handwritten character recognition using combined feature extraction," International Conference on Computing, Communication & Automation, Noida, 2015, pp. 1155-1159.
[3] M. S. Ali and M. N. I. Mondal,"Character Recognition System: Performance Comparison of Neural Networks & Genetic Algorithm," 2015 International Conference on Computer and Information Engineering (ICCIE), Rajshahi, 2015, pp. 91-94.
[4] R. S. Hussien, A. A. Elkhidir and M. G. Elnourani, "Optical Character Recognition of Arabic handwritten characters using Neural Network," 2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering (ICCNEEE), Khartoum, 2015, pp. 456-461.
[5] V. J. Dongre and V. H. Mankar, "Devanagari offline handwritten numeral and character recognition using multiple features and neural network classifier," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2015, pp. 425-431.
[6] Verma B.K, "Handwritten Hindi Character Recognition Using Multilayer Perceptron and Radial Basis Function Neural Network", IEEE International Conference on Neural Network, 4, pp. 2111-2115, 1995.
[7] Sutha.J, Ramraj.N, "Neural Network Based Offline Tamil Handwritten Character Recognition System", IEEE International Conference on Computational Intelligence and Multimedia Application, 2007, 2, 13-15, Dec.2007, Page(s): 446-450, 2007.
[8] Yuelong Li Jinping Li Li Meng, "Character Recognition Based on Hierarchical RBF Neural Networks" Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference, 1, On Page(s): 127-132, 2006.
[9] Dayashankar Singh, Maitreyee Dutta and Sarvpal Harpal Singh, "Comparative Analysis of Handwritten Hindi Character
[10] Recognition Technique", IEEE International Advanced Computing Conference (IACC'09), March 6-7, 2009, Thaper University, Patiala.
[11] Dayashankar Singh, Maitreyee Dutta and Sarvpal H. Singh, "Neural Network Based Handwritten Hindi Character Recognition", ACM International Conference (Compute 09), Jan. 9-10, 2009, Bangalore.
[12] F.Kimura, T.Wakabayashi, S.Tsuruoka, and Y.Miyake,"Improvement of Handwritten Japanese Character Recognition using Weighted Direction Code Histogram," Pattern Recognition, 30, No.8, pp. 1329-1337, 1997.
[13] N.Kato, M.Suzuki, and S.Omachi, "A Handwritten Character Recognition System Using Directional Element Feature and Asymmetric Mahalanobis Distance", IEEE Trans. on PatternAnalysis and Machine Intelligence, 21, No.3, pp. 258-262, 1999.
[14] C.C. Tappert, C.J. Suen and T. Wakahara, "The State of the Art in Outline Handwriting Recognition," IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-12, No.8, pp.707-808, 1990.
[15] D.S. Yeung, "A Neural Network Recognition System for Handwritten Chinese Character Using Structure Approach," Proceeding of the World

Congress on Computational Intelligence, 7, pp. 4353-4358, Orlando, USA, June 1994.
[16] D.Y. Lee, "Handwritten Digit Recognition Using K Nearest- neighbor, Radial Basis Function and Backpropagation Neural Networks, "IEEE Neural Computation, 3, Page(s) 440- 449.
[17] Aiquan Yuan, Gang Bai, Po Young, Yanni Guo, Xintig Zhao,"Handwritten english word recognition based on convolution neural networks", 2012 International conference on trontiers in handwriting recognition.
[18] Pranab k Charles,V. Harsh, M. Swath, CH. Deepth,"A review on the various techniques used for opticle character recognition", International journal of engineering research and application, vol.2, pp.659-662, jan-Feb 2012.
[19] R. Plamondon and S.N. Srihari,"Online and offline handwritten character recognition: A comprehensive survey", IEEE transaction on pattern analysis and machine intelligence, vol22, no1, pp-63-84,2000.
[20] Nafiz Arica and Fatos T. Yarman," An overview character recognition focused on offline handwriting", IEEE transaction on system. Man and Cybeznetic, vol.31.no.2. May 2011.
[21] J.Pradeep, E.Srinivasan, S.Himawathi, "Perfomance analysis of hybrid feature extraction technique for recognizing english handwritten character",978-1-4673-4805-8/12/2012IEEE.
[22] Anshul Gupta, Manish Srivastva, Chitraleka Mahanta,"Offline handwritten character recognition using neural network", 2011 International conference on computer application and industrial electronics.
[23] Aiquan yuan, Gang Bai, Lijing Jiao, Yajie Liu,"Offline handwritten english character recognition based on convolution neural network, 2012, 10th IAPR International workshop on document analysis system."
[24] I. Guyon, L. Schomaker, R. Plamondon, M. Liberman, and S. Janet,"Unipen project of on line data exchange and recognizer benchmarks". In Pattern Recognition 1994, vol.2-conference B:
[25] Velappa Ganapathy, Kok Leong Liew, "Handwritten Character Recognition Using Multiscale Neural Network Training Technique", World Academy of Science, Engineering and Technology, pp. 32-37, 2008.
[26] J Pradeep, E Shrinivasan and S.Himavathi, "Diagonal Based Feature Extraction for Handwritten Alphabets Recognition System Using Neural Network", International Journal of Computer Science & Information Technology (IJCSIT), vol. 3, No 1, Feb 2011.
[27] Om Prakash Sharma, M. K. Ghose, Krishna Bikram Shah, "An Improved Zone Based Hybrid Feature Extraction Model for Handwritten Alphabets Recognition Using Euler Number", International Journal of Soft Computing and Engineering (ISSN: 2231 - 2307), Vol. 2, Issue 2, pp. 504- 508, May 2012.
[28] Anshul Mehta, Manisha Srivastava, Chitralekha Mahanta "Offline handwritten character recognition using neural network" IEEE 2011 International conference on computer applications and Industrial Electronics.
[29] Serrano, Nicolás, Adrià Giménez, Jorge Civera, Alberto Sanchis, and Alfons Juan. "Interactive handwriting recognition with limited user effort."International Journal on Document Analysis and Recognition (IJDAR) 17, no. 1 (2014): 47-59.
[30] Amma, Christoph, Marcus Georgi, and Tanja Schultz. "Airwriting: a wearable handwriting recognition system." Personal and ubiquitous computing 18, no. 1 (2014): 191-203.
[31] Minhua Li; Chunheng Wang; Ruwei Dai, "Unconstrained handwritten character recognition based on WEDF and Multilayer Neural Network", 7th World Congress on Intelligent Control and Automation, 2008.
[32] Yusuf Perwej and Ashish Chaturvedi, "Neural Networks for Handwritten English Alphabet Recognition", International Journal of Computer Applications (0975 – 8887) Volume 20 – No.7, 2011, pp.1-5.
[33] Salma Shofia Rosyda1 and Tito Waluyo Purboyo, "A Review of Various Handwriting Recognition Methods", International Journal of Applied Engineering Research ISSN 0973-4562 Volume13-No.2, 2018 pp. 1162-1163.
[34] Anita Pal and Dayashankar Singh, "Handwritten English Character Recognition Using Neural Network", International Journal of Computer Science & Communication Vol. 1, No. 2, July December 2010,pp.141-144.
[35] Amandeep Kaur, Seema Baghla, Sunil Kumar, "Study Of Various Character Segmentation Techniques For Handwritten Off-Line Cursive Words: A Review", International Journal of Advances in Science Engineering and Technology, Vol. 3, Issue-3, July-2015.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY