INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

HANDWRITTEN GURMUKHI CHARACTERS RECOGNITION

Er. Kulbir Kaur [1*],Er. Harpreet Kaur [2],
[1*] ECE &GCET/PTU
[2]ECE & DIET/PTU
*Correspondence Author: r_saini1888@yahoo.in

Abstract

Recently there is growing trend among worldwide researchers to recognize handwritten characters of many languages and scripts. Much of research work is done in English, Chinese and Japanese like languages. However, on Indian scripts, the research work is comparatively lagging. The work on other Indian scripts is in beginning stage. In this thesis work I have done my work on recognition of handwritten character of Gurumukhi script. Take samples on white papers written in an isolated manner. After scanning, in preprocessing stage, the samples are converted to gray scale images. Then gray scale image is converted into binary image. I also apply PSONN for increase the efficiency of character recognition.

Introduction

The objective of an OCR system is to recognize alphabetic letters, numbers, or other characters, which are in the form of digital images, without any human intervention.OCR system for handwritten character using different wavelet transform[1]. Handwritten character and numeral recognition for Roman, Arabic, Chinese and Indian scripts", had reported various works on HCR systems in four popular scripts which are Roman, Arabic, Chinese and Indian[2]. This is accomplished by searching a match between the features extracted from the given character's image and the library of image models. Ideally, we would like the features to be distinct for different character images so that the computer can extract the correct model from the library without any confusion. At the same time, we also want the features to be robust enough so that they will not be affected by viewing transformations, noises, resolution variations and other factors. The basic processes of an OCR system, as shown in fig.1.
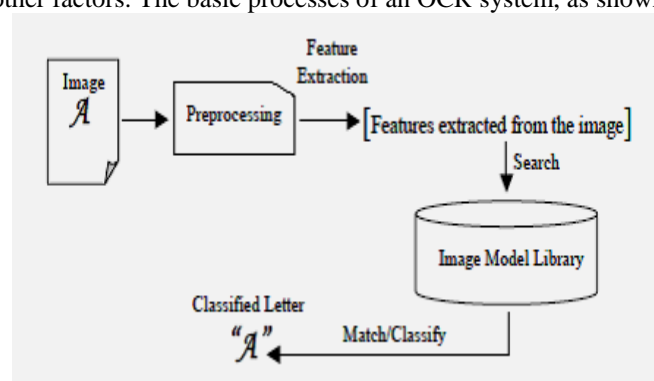

Fig.1 The basic processes of an OCR system

On-line character recognition deals with a data stream which comes from a transducer while the user is writing. The typical hardware to collect data is a digitizing tablet which is electromagnetic or pressure sensitive. When the user writes on the tablet, the successive movements of the pen are transformed to a series of electronic signal which is memorized and analysed by the computer.[3] Off-line character recognition is performed after the writing is finished. Offline handwritten Gurumukhi character recognition by using a modified division points (MDP) feature extraction technique[4].GUI fuction is used forinterfacing with matlab tool.[6,7] This technique is compared with other recently used feature extraction techniques, namely zoning features, diagonal features, directional features, intersection and open end points features, and transition features[8,9]. The major difference between on-line and off-line character recognition is that on-line character recognition has time-sequence contextual information but off-line data does not. This difference generates a significant divergence in processing architectures and methods.

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

System Overview

The idea of character recognition may be achieved by capturing the image of the handwritten character using a simple web camera and then processing it for recognition To implement my proposed work MATLAB's [6]image processing and neural network toolbox is used. Image Processing Toolbox™ provides a comprehensive set of reference-standard algorithms, functions, and apps for image processing, analysis, visualization, and algorithm development. You can perform image analysis, image segmentation, image enhancement, noise reduction, geometric transformations, and image registration. Many toolbox functions support multicore processors, GPUs, and C-code generation[10]. In computer science, local search is a metaheuristic method for solving computationally hard optimization problems. Local search can be used on problems that can be formulated as finding a solution maximizing a criterion among a number of candidate solutions .The idea may be represented in the form of chart as shown in fig2
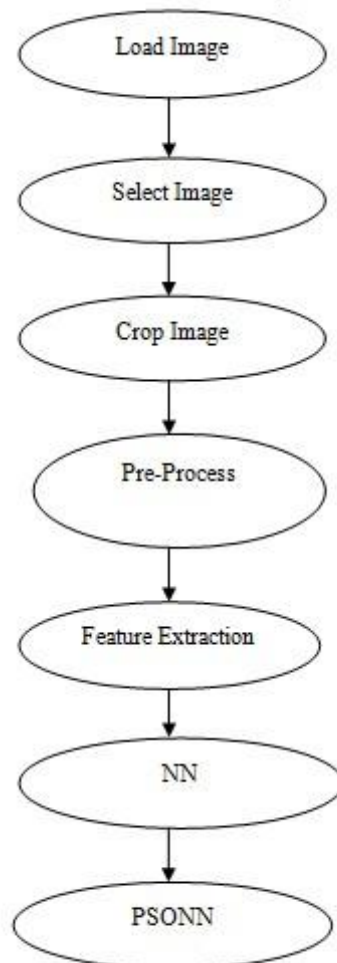

Fig 2 . System representation

Load image

Any scanned handwriting document can be loaded into GUI[6]. Axis to show the results of pre-processing steps is also included into GUI and push buttons are placed to give freedom to select pre-process steps. Shown in fig 3.
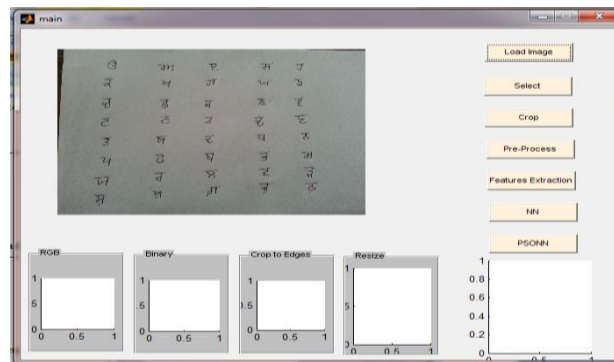
INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

Fig. 3. Load image

Select image

Selection of character into scanned document. features are extracted in features vectors selection[11] step. features are used as these exhibits two types of properties i.e. spatial locality as well as orientation property. Shown fig 4
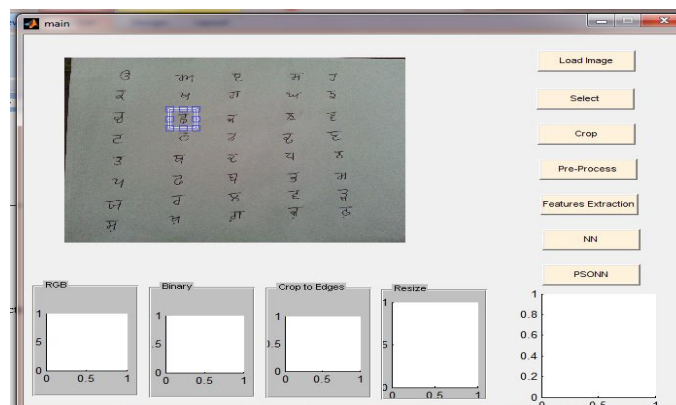

Fig 4. Select image

Crop image

For cropping Character select the block to crop image. In cropping step all edges are from particular selected character. Cropping of image is done only in the gray image[8]. So first of all RGB images are converted into gray scale image.
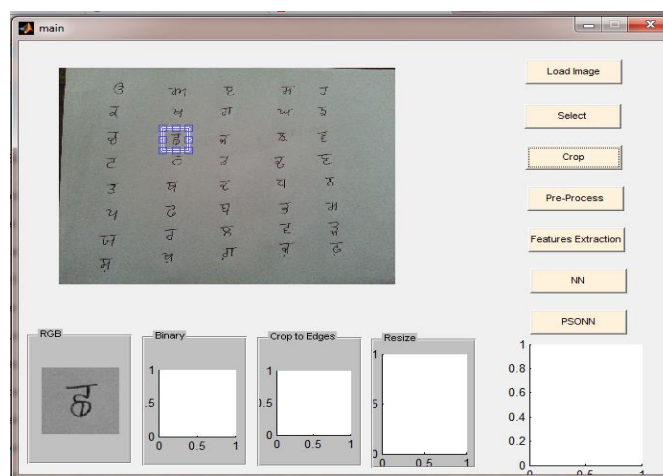

Fig 5 . crop image

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

Pre-process
In pre-processing first image is segmented for desired character for recognition. Convert the segmented character image into gray scale. Then Convert the gray scale image into binary image by thresholding[12]. Remove all white area form the binary image. Resize the image and convert it into a vector.

Feature extraction
In feature extraction each character is assigned a set of features to identify it. This vector is used to distinguish the character[13]. For feature extraction in this work zone based features along with bounding box is used. used Filter based method for feature extraction. His database consists of 200 samples of each of basic 35 characters of Gurmukhi script collected from different writers. These samples are pre-processed and normalized to 32*32 sizes. The cross validation of whole database with SVM classifier used with RBF kernel.[14]

NN
The feature vector thus obtained is fed into neural network with maximum 100 iterations and 0.9 learning rate[15]. . The MATLAB defined user interface for neural network[6]. The mean square error between output and target vector should be minimized during neural network training, which is fulfilled in our case.

PSONN
Training of selected sample with target data in neural network gives exact recognition of Gurumukhi character in most of cases but some characters are not recognized by NN because of their shapes[17]. To overcome this drawback neural network in association with particle swarm optimization (PSO) is used. In neural network recognition 36 Gurumukhi characters are recognized correctly out of 40 characters[18,19]. The performance comparison of neural network with PSONN in terms of correct recognition of Gurumukhi characters is shown in figure 5.7.
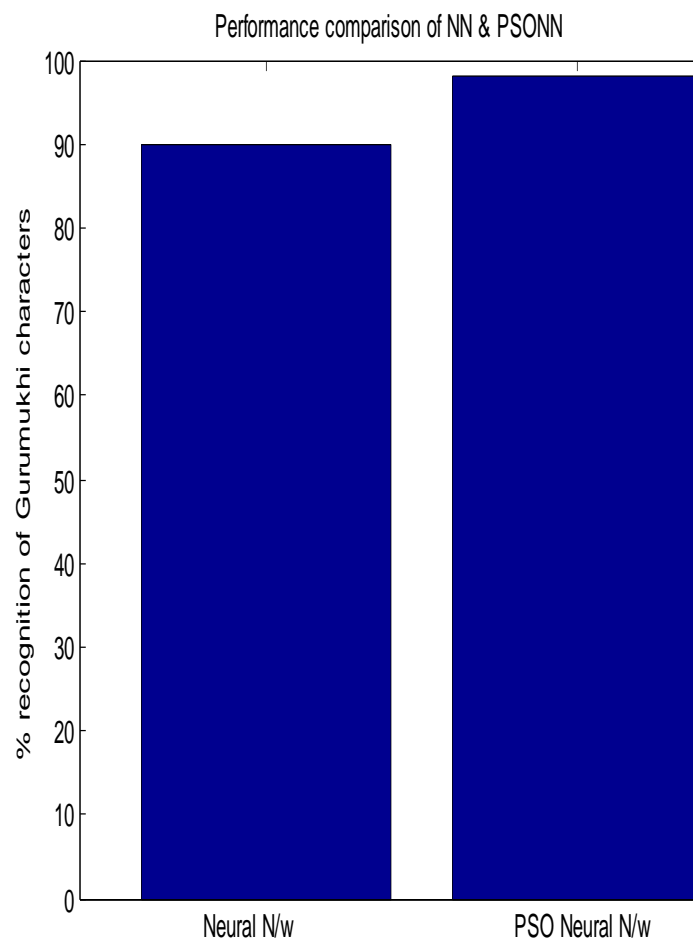

Fig 6 . Select image

Result & Discussion

First of all I show the result of recognition of Gurumukhi character with only NN algo. After then I show the result of PSONN and also comparison of NN and PSONN. Now i show my first result of NN , as shown in fig 7.
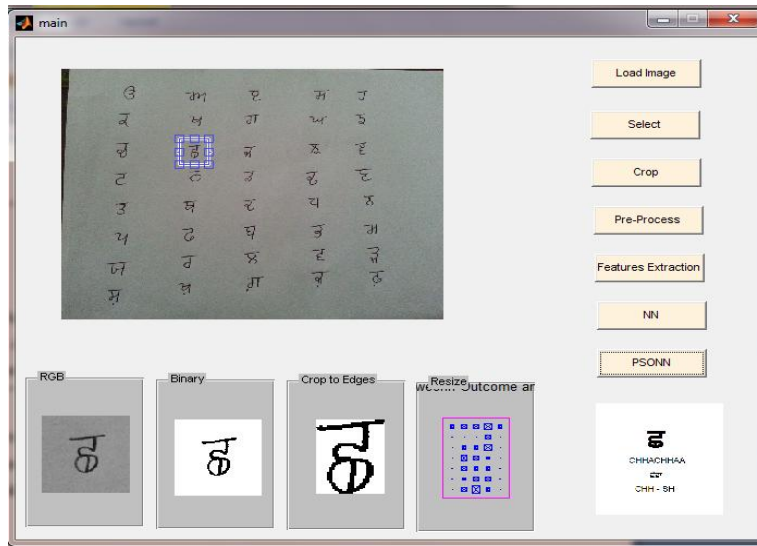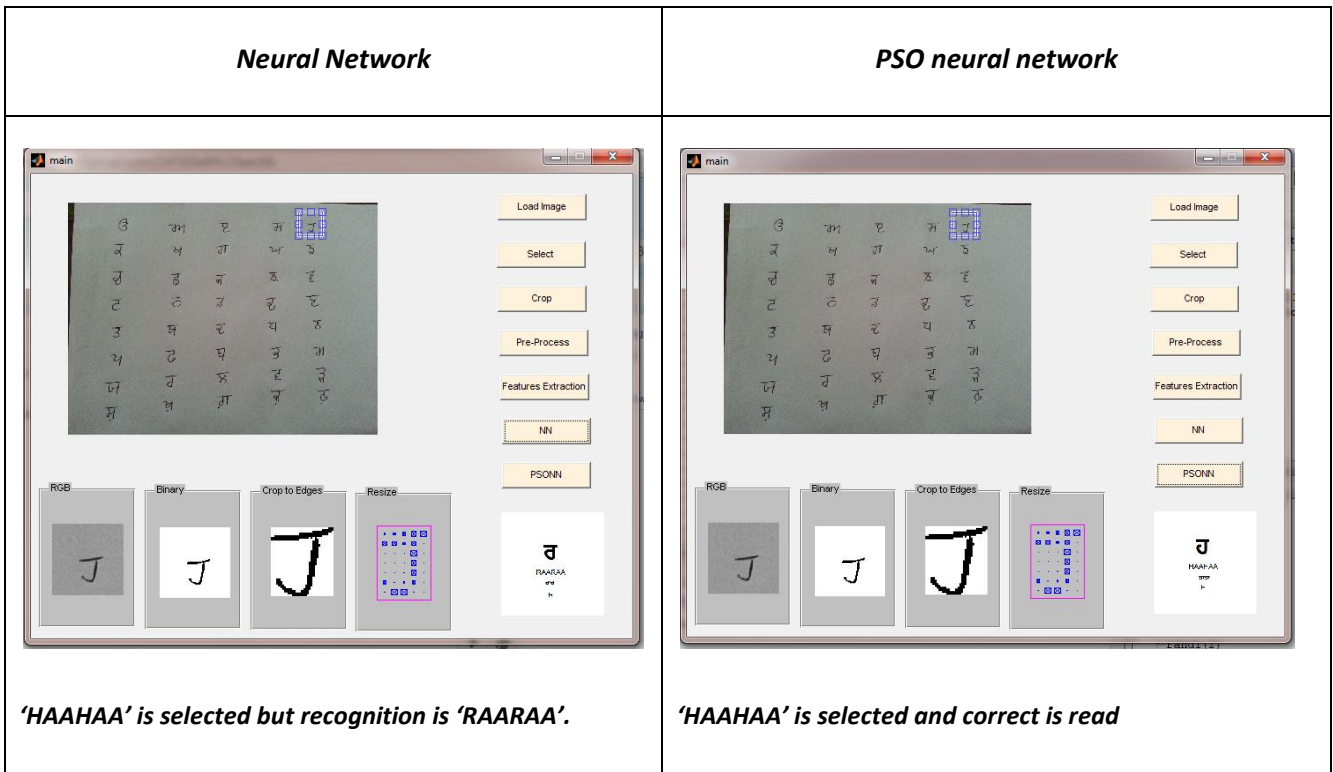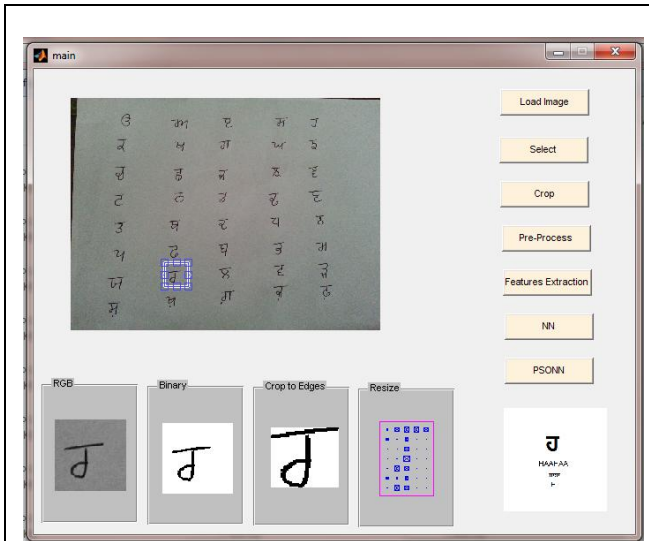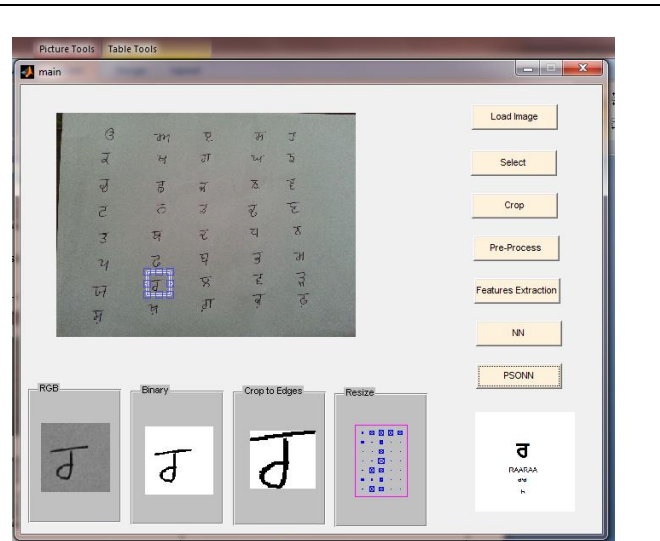

Fig 6. Select image

Training of selected sample with target data in neural network gives exact recognition of Gurumukhi character in most of cases but some characters are not recognized by NN because of their shapes. To overcome this drawback neural network in association with particle swarm optimization (PSO) is used. PSO, as discussed, is a bio optimized technique which falls under the category of global optimization. In this the weights and biases of neural network are trained with PSO rather than back propagation algorithms, usually used in NN training.

Neural Network	PSO neural network
'HAAHAA' is selected but recognition is 'RAARAA'.	*'HAAHAA' is selected and correct is read*

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT
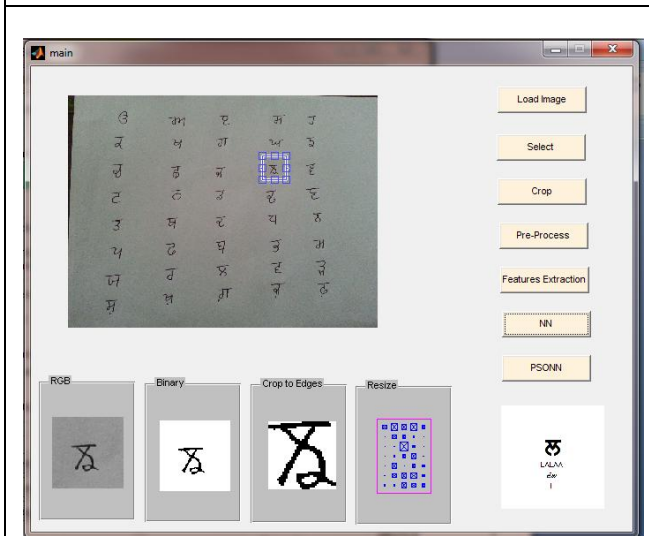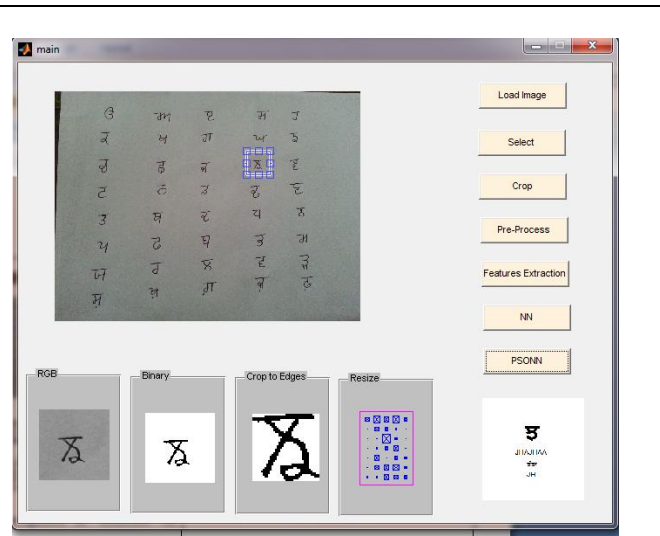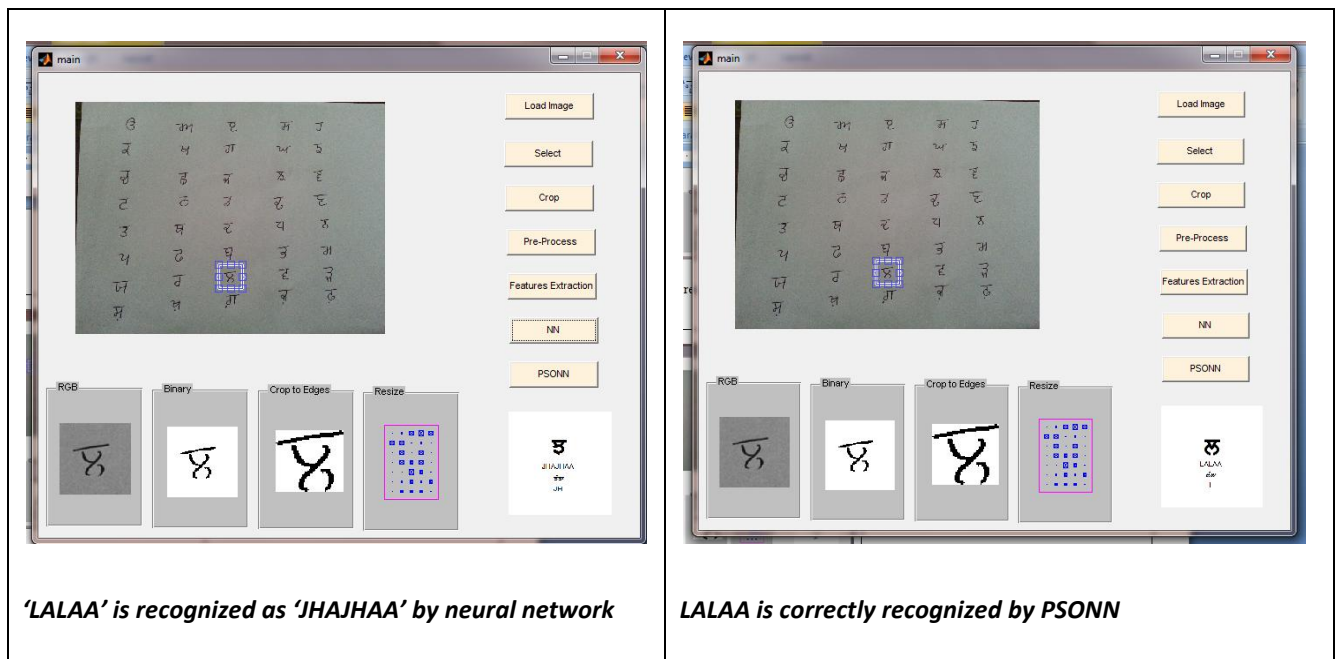

Wrong recognition of RAARAA by neural network

RAARAA is correctly recognized by PSONN

'JHAJHAA' is recognized as 'LALAA' by neural network

JHAJHAA is correctly recognized by PSONN

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

'LALAA' is recognized as 'JHAJHAA' by neural network   LALAA is correctly recognized by PSONN

Conclusion

There are four cases where neural network is failed to recognize the character but PSONN recognized those correctly. Two pairs of cases generated in NN which are not recognized because of similarity in shape of characters as shown in above table but this is not the problem with PSO optimized NN. The only limitation in case of PSONN is that the selection rectangle has to be kept very small in size otherwise it also give erroneous result as image cropped to edges will not be perfect. The selection of white space in Gurumukhi character selection is covering another character part also, so it is giving a wrong interpretation. That's why white area is selected as small as possible so that no other character part is covered in main selection

References

1. Singh, P., Budhiraja, S. (2012), "Offline Handwritten Gurmukhi Numeral Recognition using Wavelet Transforms" I.J.Modern Education and Computer Science, 2012
2. Azmi, A.N., Nasien, D., Shamsuddin, S.M (2013)," A review on handwritten character and numeral recognition for Roman, Arabic, Chinese and Indian scripts" International Journal of advanced studies in Computers, Science & Engineering (IJASCSE), Volume 2, Issue 4, 2013
3. B. Saranya Bargavi, C. Santhi (2014),"Global and Local Facial Feature Extraction using Gabor Filters" International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 4, April 2014
4. Bayat, A.B., (2013)" Recognition of Handwritten Digits Using Optimized Adaptive Neuro-Fuzzy Inference Systems and Effective Features" Journal of Pattern Recognition and Intelligent Systems, Vol. 1 Issue. 2, Aug. 2013, PP. 25-37
5. Diego J. Romero, Leticia M. Seijas, Ana M. Ruedin (2007)," Directional Continuous Wavelet Transform Applied to Handwritten Numerals Recognition Using Neural Networks" JCS&T Vol. 7 No. 1, April 2007
6. http://www.mathworks.in/help/imaq/imaqhwinfo.html
7. Haseen, H., Khemakhem,M. (2011)," Comparative study of Arabic handwritten characters invariant feature" International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 2, No. 12, 2011
8. Kumar,M., Jindal, M.K. and Sharma, R. K. (2013)," MDP Feature Extraction Technique for Offline Handwritten Gurmukhi Character Recognition" Smart Computing Review, vol. 3, no. 6, December 2013
9. Kumar,M., Jindal, M.K. and Sharma, R. K. (2013)," PCA-based Offline Handwritten Character Recognition System" Smart Computing Review, vol. 3, no. 5, October 2013
10. Muhammad, N.A., Javed, I. (2012), "Handwritten Character Recognition Using Multiclass SVM Classification with Hybrid Feature Extraction" Pak. J. Engg. & Appl. Sci. Vol. 10, Jan., 2012.
11. Patel, D.K., Som, T., Yadav, S.K, Singh, M.K.," Handwritten Character Recognition Using Multiresolution Technique and Euclidean Distance Metric" Journal of Signal and Information Processing, 2012, 3, 208-214

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

12. Roy, A., Dutta, D., Choudhury, K. (2013), "Training Artificial Neural Network using Particle Swarm Optimization Algorithm" International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Volume 3, Issue 3, March 2013

13. Sahu, V.L., Kubde, B. (2013)," Offline Handwritten Character Recognition Techniques using Neural Network: A Review" International Journal of Science and Research (IJSR), Volume 2 Issue 1, January 2013

14. Singh, S., Aggarwal, A., Dhir, R. (2012),"Use of Gabor Filters for Recognition of Handwritten Gurmukhi Character" International Journal of Advanced Research in Computer Science and Software Engineering(IJARCSSE), Volume 2, Issue 5,

15. Singh, G., Kumar, C.J., Rani, R. (2013)," Feature Extraction of Gurmukhi Script and Numerals: A Review of Offline Techniques" International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Volume 3, Issue 1, January 2013

16. Singh, P., Budhiraja,S. (2012)," Handwritten Gurmukhi Character Recognition Using Wavelet Transforms" International Journal of Electronics, Communication & Instrumentation Engineering Research and Development (IJECIERD), Vol.2, Issue 3 Sep 2012

17. Abed, Ali, M., Hamid Ali Abed Alasadi (2013)," Simplifying Handwritten Characters Recognition Using a Particle Swarm Optimization Approach" European Academic Research, VOL. I, Issue 5, August 2013

18. Siddharth, K.S., Jangid, M., Dhir, R., Rani, R. (2011)," Handwritten Gurmukhi Character Recognition Using Statistical and Background Directional Distribution Features" International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 6 June 2011

19. Dr. Hanan A. R. Akkar, Hussain, S.S. (2013)," Integration of Swarm Intelligence and Artificial Neural Network for Medical Image Recognition" Eng. &Tech. Journal, Vol.31, Part (A), No.13, 2013