OPTICAL CHARACTER RECOGNITION PERFORMANCE ANALYSIS OF SIF AND LDF BASED OCR

Pradeep Kumar Jena[1], Charulata Palai[2], Lopamudra Sahoo[3] and Anshupa Patel[4]

[1]Department of MCA, National Institute of Science and Technology, Berhampur, Odisha, India
pradeep1_nist@ yahoo.com
[2]Department of CSE, National Institute of Science and Technology, Berhampur, Odisha, India
charulatapalai@ gmail.com
[3,4]National Institute of Science and Technology, Berhampur, Odisha, India
06lopamudra@gmail.com
anshupa@gmail.com

ABSTRACT

The Optical Character Recognition (OCR) is becoming popular areas of research under pattern recognition and smart device applications. It requires the intelligence like human brain to recognize the various handwritten characters. Artificial Neural Network (ANN) is used to gather the information required to recognize the characters adaptively. This paper presents a performance analysis of character recognition by two different methods (1) compressed Lower Dimension Feature(LDF) matrix with a perceptron network, (2) Scale Invariant Feature (SIF) matrix with a Back Propagation Neural network (BPN). A GUI based OCR system is developed using Matlab. The results are shown for the English alphabets and numeric. This is observed that the perceptron network converges faster, where as the BPN can handle the complex script recognition when the training set is enriched.

KEYWORDS

Character recognition, perceptron network, back-propagation neural network, scale invariant feature, low dimensional feature.

1. INTRODUCTION

Automatic character recognition is a well-accepted area of research under pattern recognition. In handwritten character recognition, characters are written by different individuals that vary drastically from person to person due to variation in the writing style, its size and orientation of characters[1]. This makes the system difficult to recognize the characters. Artificial Neural Network (ANN) helps to solve the problem of identifying handwritten characters in an automated manner. ANN is an adaptive computational model which is activated by set of pixels of a specific character as features, processing of the similar and divergence information available in the features used to recognize the character.

Dhinaharan Nagamalai et al. (Eds) : ACITY, WiMoN, CSIA, AIAA, DPPR, NECO, InWeS - 2014
pp. 319–328, 2014. © CS & IT-CSCP 2014 DOI : 10.5121/csit.2014.4532

Various methods have been developed for the recognition of handwritten characters, such as the compressed Column-wise Segmentation of Image Matrix (CSIM) [2] using Neural Network. In this method, the image matrix is compressed and segmented column-wise then training and testing using a neural network for different character is performed. The Multi-scale Technique (MST)[2][3] used for high resolution character sets.

In case of feature based method, Scale Invariant features of characters such as height, width, centroid, number of bounded regions and textual features such as histogram information are used for character recognition. The features are feed to a Neural Network [4] for training and testing purpose. Hand written character recognition using Row-wise Segmentation Technique (RST) approach used to find out common features along the rows of same characters written in different hand writing styles and segmenting the matrix into separate rows and finding common rows among different hand writing styles[5]. Block-wise segmentation technique is also used for the character recognition [6] by matching the similarity among blocks of the characters.

The complex character such as Hindi, Oriya and Bangla character recognition [7][8] is more challenging at it produces very alike feature matrices for different characters due to their structural complexity. Hybrid methods are also applied to recognize the hand written characters. One such method is a prototype learning/matching method that can be combined with support vector machines (SVM) in pattern recognition [9].

It is observed that the finding the ideal feature set for particular language and normalizing the feature matrix is not easy, it requires substantial amount of processor time. In this work we have done a comparative study on character recognition using feature matrix with a Back Propagation Neural network (BPN) verses the character recognition using reduced dimensional block matrix(8x8) with a Perceptron Neural network for English hand written characters i.e. the upper case, lower case alphabet as well as digits.

2. METHODOLOGY

The sample handwritten characters are collected form ten different persons using black gel pen i.e. 10 samples of each letter each having different style. These blocks of characters were digitized using a scanner. Then each character is extracted from the scanned image automatically and saved with an appropriate name.

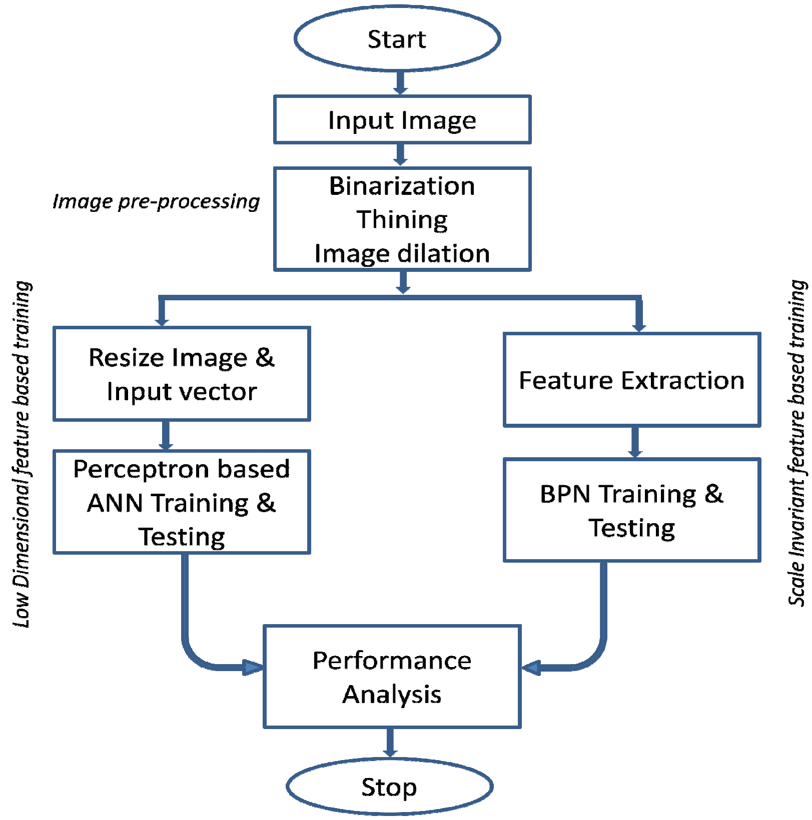


Figure 1: The Proposed Model for Analysis

2.1. Image Pre-processing

The individual character image is pre-processed to produce a skeletal template of the handwritten character. These involve various tasks such as (1) Binarization to reproduce the image with 0 (black) or 1(white), (2) Thinning to remove the thickness artefact of the pen used for writing characters, (3) image dilation to restore the continuity of the image pixels. Figure-2 show the inverted binarized data sheet of the set of handwritten characters.

2.2. Approach-1: Low dimensional feature based recognition

The images are resized into unique standard since the sample character images are different in dimension. They are converted to a reduced dimensional image matrix of size 8x8. It preserves only the highly significant features of the character that are used for the character recognition. The input vector i,e. a [64x1] matrix is prepared from the 8x8 image. It is used for the training and testing of perceptron neural network.

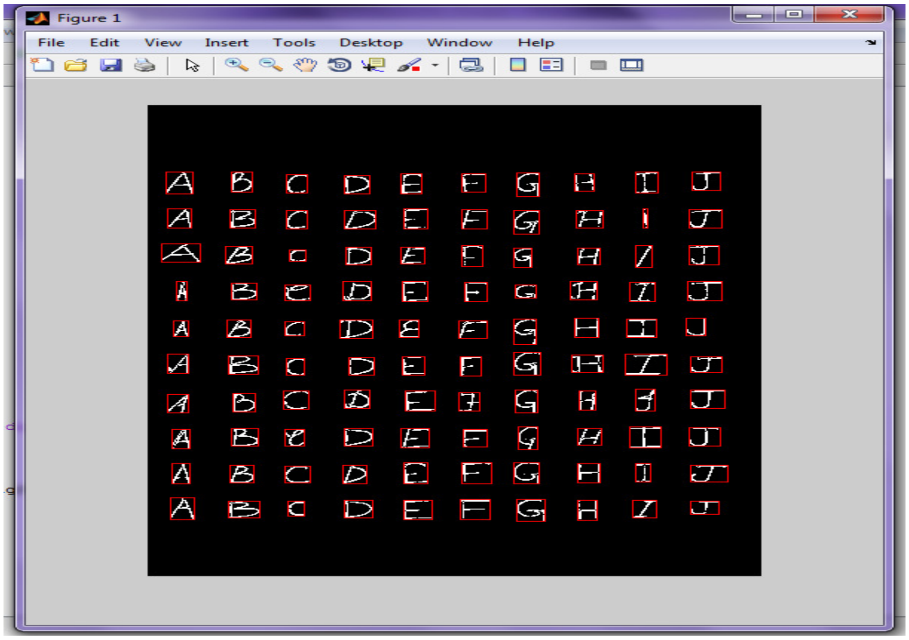


Figure 2: Processed image with bounding box

2.3. Rescaling of image matrix

Case-1: When the original image matrix is a multiple of 8.

i) Input image matrix
ii) Dimension of the original image is divided by 8 i.e. 64 x 32 will be 8 x 4
iii) Original matrix is split into ‘n’ uniform blocks of new dimension
iv) A uniform block is assigned to ‘1’ if the number of 1’s is greater than or equal to the number of 0’s. Otherwise, a uniform block is assigned to ‘0’.

Case-2: When the original image matrix is not a multiple of 8.

i) Input image matrix
ii) Dimension of the original image is first converted to the nearest multiple of 8 by appending dummy (zeros) rows and columns i.e. 60 x 50 will be 64 x 56
iii) Dimension of the revised image is divided by 8
iv) Revised matrix is split into ‘n’ uniform blocks of new dimension
v) A uniform block is assigned to ‘1’ if the number of 1’s is greater than or equal to the number of 0’s. Otherwise, a uniform block is assigned to ‘0’.

In this way images of different dimensions are resized into an 8 x 8 binary matrix. The columns of 8 x 8 matrixes are stored in a single column matrix one after other. Likewise, 10 samples of each 26 characters are considered and transformed it into column of size 64 each. As a result of a training set of 64 x 260 matrix is obtained. Accordingly the target set of 5 x 260 is generated.

2.4. Perceptron based ANN training

An Artificial Neural Network (ANN) is an adaptive computational system, it follows perceptron learning technique. The input layer consists of 64 neurons that represent one character as input, the hidden layer consists of 32 neuron, where as the output consists of 7 neurons that represents pattern of 0 and 1 which maps to an individual character. Each neuron is connected to other neurons by a link associated with weights. The weight contains information about the input, which is updated during each epoch.

2.5. Approach-2: Scale Invariant feature based recognition

Aspect Ratio: Height and width of character is obtained. The ratio of height and width remain approximately same for same person for the different characters.

$$AR = \frac{L}{W}$$

Where, AR=Aspect Ratio
L=Length of Character
W=Width of Character

Occupancy Ratio: This feature is the ratio of number of pixels which belong to the character to the total pixels in the character image. This feature provides information about character density.

Number of Horizontal Lines: It's the number of horizontal lines in a character. It's found out using a 3x3 horizontal template matrix.

Number of Vertical Lines: It's the number of vertical lines in a character. It's found out using a 3x3 vertical template matrix.

Number of Diagonal Lines: It's the number of diagonal-1 lines in a character. It's found out using a 3x3 diagonal-1 and diagonal-2 template matrix.

Number of Bounded Regions: It's the number of bounded areas found within a character image.

Number of End Points: End points are defined as those pixels, which have only one neighbor in its eight way neighborhood. Figure-3 the two end points of the image.

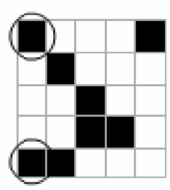


Figure 3: End point in an image

Vertical Center of Gravity: Vertical centre of gravity shows the vertical location of the character image. Vertical centre of gravity of image is calculated as follow

$$cog(v) = \frac{\sum y.Ny}{\sum Ny}$$

Where, Ny: the number of black pixels in each horizontal line Ly with vertical coordinate y,
Horizontal Center of Gravity: Horizontal centre of gravity shows the horizontal location of the character image. Horizontal centre of gravity of image calculated as follow:

$$cog(h) = \frac{\sum x.Nx}{\sum Nx}$$

Where, Nx: the number of black pixels in each vertical line Lx with horizontal coordinate x

2.6. BPN based ANN Training

Back-propagation neural network is mostly used for the handwritten character recognition since it support the real values as input to the network. It is a multi-layer feed-forward network which consists of an input layer, hidden layer and output layer. The normalized values of the scale-invariant feature matrix are given as the input for the training, the output is pattern of 0's and 1's which maps to an individual character.

2.7. Testing and Recognition

Separate test sets are prepared for testing purpose. After the training process is completed, the test pattern is fed to the neural network to check the learning ability of the trained net. The output of the simulation is compared with the specified target set. The character is recognized by selective thresholding technique.

3. RESULT ANALYSIS

The OCR system is developed using MATLAB. The experimental results are shown below which is carried out to recognize the "A" and "5" as the inputs. The Figure-4 represents the normalized values of the scale invariant features of the English Alpha-numerics. Figure-5 indicates the performance analysis of BPN network as trained with normalized feature values. The Figure-6 shows the interface to load the test image and to select the training type such as BPN based training. Figure-7 indicates the recognition of "A", The Figure-8 shows the LDF matrix, Figure-9 represents the perceptron based training and testing. The recognition of "5" is shown in Figure-10.

	A	B	C	D	E	F	G	H	I
1	HorizontalLines	VerticalLines	DiagonalLines	BoundedRegions	AspectRatio	EndPoints	OccupancyRatio	CogVertical	CogHorizontal
2	0.00349	0.00273	0.00655	0.00531	0.00005	0.85017	0.00001	0.12908	0.00261
3	0.00260	0.00383	0.00539	0.00729	0.00004	0.82175	0.00001	0.15554	0.00310
4	0.00248	0.00422	0.00593	0.00819	0.00004	0.80907	0.00002	0.16528	0.00359
5	0.00215	0.00473	0.00628	0.00905	0.00005	0.79398	0.00002	0.17804	0.00407
6	0.00218	0.00508	0.00588	0.00927	0.00004	0.78486	0.00002	0.18701	0.00411
7	0.00225	0.00413	0.00540	0.00821	0.00004	0.81779	0.00001	0.15811	0.00333
8	0.00171	0.00516	0.00570	0.00982	0.00004	0.76187	0.00003	0.20981	0.00447
9	0.00172	0.00547	0.00552	0.01017	0.00004	0.77821	0.00002	0.19325	0.00426
10	0.00140	0.00685	0.00660	0.01231	0.00006	0.74507	0.00005	0.21970	0.00559
11	0.00097	0.00887	0.00505	0.01637	0.00004	0.69268	0.00008	0.26755	0.00648
12	0.00230	0.00367	0.00625	0.00695	0.00005	0.76337	0.00002	0.21252	0.00396
13	0.00197	0.00428	0.00633	0.00788	0.00005	0.75472	0.00003	0.21917	0.00433
14	0.00209	0.00317	0.00601	0.00616	0.00005	0.76403	0.00002	0.21441	0.00367
15	0.00209	0.00302	0.00648	0.00606	0.00006	0.76450	0.00002	0.21346	0.00377
16	0.00202	0.00352	0.00609	0.00695	0.00005	0.75001	0.00002	0.22649	0.00407
17	0.00190	0.00423	0.00599	0.00766	0.00005	0.75654	0.00002	0.21854	0.00415
18	0.00180	0.00356	0.00638	0.00705	0.00006	0.74697	0.00003	0.22907	0.00423
19	0.00147	0.00459	0.00532	0.00895	0.00004	0.74794	0.00003	0.22651	0.00432
20	0.00184	0.00371	0.00615	0.00701	0.00005	0.76369	0.00002	0.21287	0.00395
21	0.00161	0.00423	0.00603	0.00803	0.00005	0.76124	0.00002	0.21367	0.00420
22	0.00162	0.00416	0.00645	0.00822	0.00005	0.82063	0.00002	0.15430	0.00360
23	0.00134	0.00492	0.00803	0.00950	0.00008	0.79639	0.00003	0.17325	0.00465
24	0.00186	0.00356	0.00700	0.00748	0.00006	0.77700	0.00002	0.19765	0.00415
25	0.00163	0.00469	0.00637	0.00967	0.00005	0.76089	0.00003	0.21033	0.00471
26	0.00166	0.00429	0.00696	0.00808	0.00006	0.81523	0.00002	0.15878	0.00377

Figure 4: Scale Invariant normalized feature matrix for 'A'-'z' & '0'-'9'

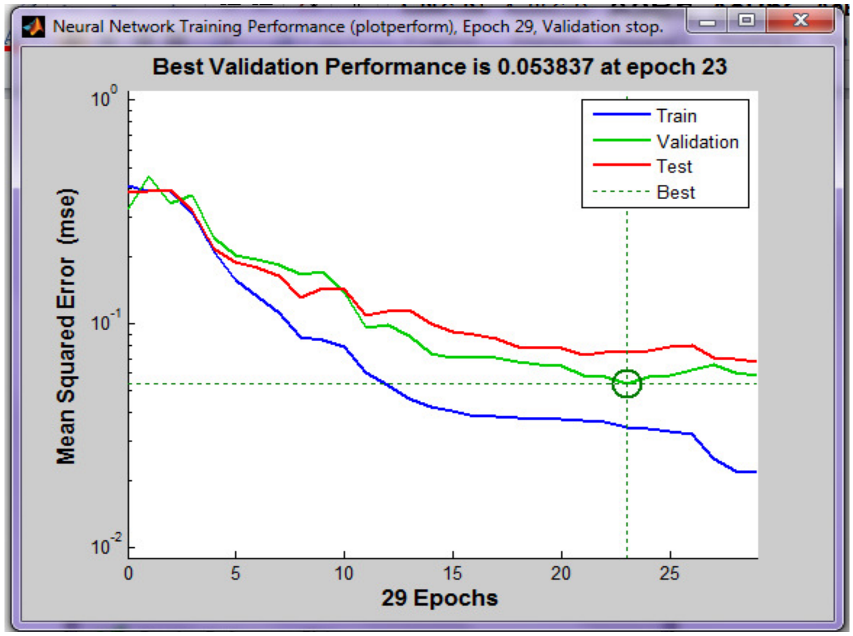


Figure 5: BPN based training scale invariant features

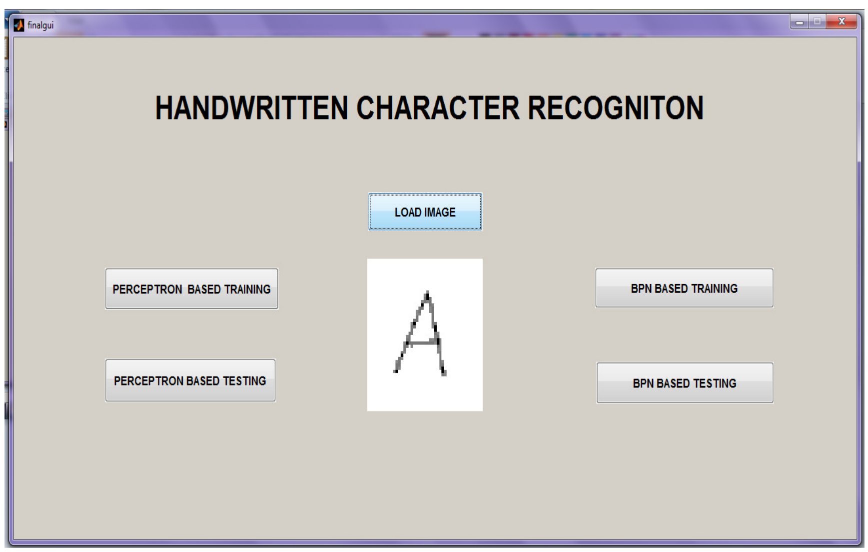


Figure 6: User Interface for Load Image

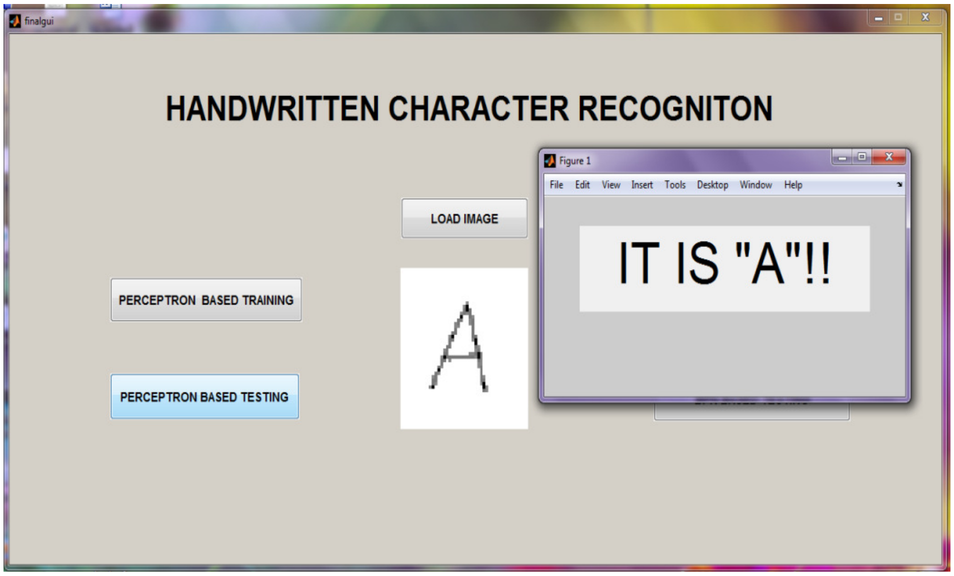


Figure 7: Recognition of character ‘A’ using BPN

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	0	0	1	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1
3	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	1	0	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
6	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
8	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0
9	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	1	0	1
10	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
11	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
13	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1	1
15	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
18	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	0	1	1	1	0	1	1	1
19	1	1	0	0	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1	0
20	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0
21	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0
22	0	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
23	0	0	0	0	1	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	0	1
24	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
26	1	1	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1
27	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0
28	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0
29	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
30	0	1	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	0	1	0	0	1	1	1	0	1	1
32	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1

Figure 8: Reduced dimensional feature matrix[64x1] for ‘A’-‘z’ & ‘0’-‘9’

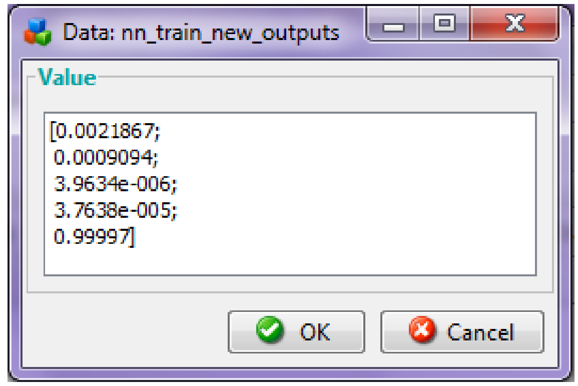
<table>
<tr>
<td>The target matrix of ‘A’ is
[0
0
0
0
1]</td>
<td>The output matrix of ‘A’ is

</td>
</tr>
</table>

Figure 9: Perception based training & testing

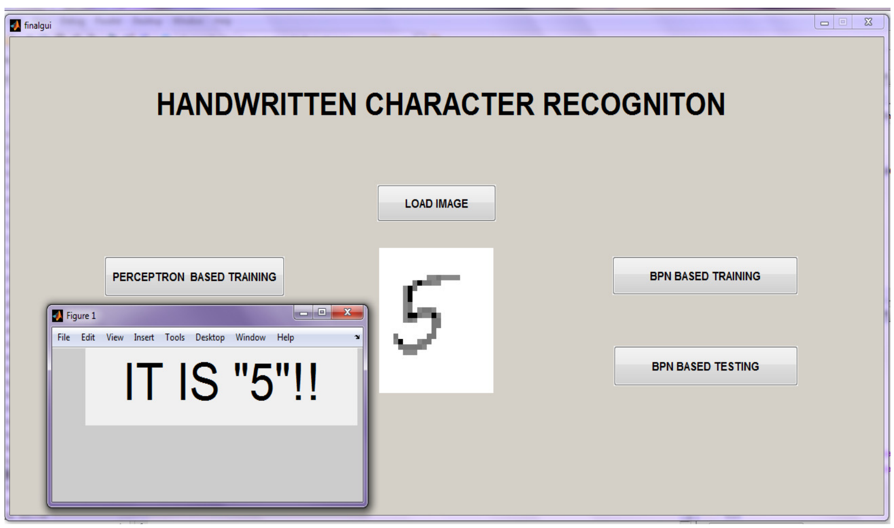


Figure 10: Recognition of character ‘5’ using perceptron network

4. CONCLUSION

In this work it has been observed that finding the reduced dimensional feature matrix of an image is easy in comparison with the scale invariant feature matrix. The training performance of the SIF matrix is much faster and reliable. The performance of the network depends upon selection of the features into the SIF matrix. The feature selection is more challenging in case of structurally complex scripts such as Bangla, Hindi and Oriya. It has been observed that using the SIF features, the English alphabets and numeric’s matches with an accuracy of 95% on average matching while the LDF match accuracy varies from 78% to 96% for different characters.

The work may be further extended to test the regional language scripts. It can be tested with different reduced dimensional matrix of different dimensions.

REFERENCES

[1] Vijay Patil and Sanjay Shimpi, “Handwritten English Character Recognition Using Neural Network”, Elixir Comp. Sci. & Engg. 41 (2011) 5587-5591, November 2011.
[2] Velappa Ganapathy, and Kok Leong Liew, “Handwritten Character Recognition Using Multi-scale Neural Network Training Technique”, World Academy of Science, Engineering and Technology 39 2008.
[3] Rakesh Kumar Mandal, N R Manna, “Hand Written English Character Recognition Using Column-Wise Segmentation of Image Matrix (CSIM)”, Issue 5, Volume 11, May 2012.
[4] Rakesh Kumar Mandal, N R Manna, “Hand Written English Character Recognition Using Row-Wise Segmentation (RST)”, International Symposium on Devices MEMS, Intelligent Systems & Communication (ISDMISC) 2011.
[5] Apash Roy, N. R. Manna, “Handwritten Character Recognition Using Block wise Segmentation Technique (BST) in Neural Network”.
[6] Soumya Mishra, Debashish Nanda, SanghamitraMohanty, “Oriya character recognition using Neural Networks”, Special Issue of IJCCT Vol. 2 Issue 2, 3, 4; 2010 for International Conference [ICCT-2010], 3rd-5th December 2010.
[7] B. Kumar, N. Kumar, C. Palai,.P. K. Jena, S. Chattopadhyay, “Optical Character Recognition using Ant Miner Algorithm: A Case Study on Oriya Character Recognition”, International Journal of Computer Applications (0975 – 8887) Volume 61– No.3, January 2013.
[8] Fu Chang, Chin-Chin Linand Chun-Jen Chen,Institute of Information Science, Academia Sinica , Taipei, Taiwan Dept. of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan”Applying A Hybrid Method To Handwritten Character Recognition”.

AUTHORS

Pradeep Kumar Jena Working as an Associate Professor in the Dept. of MCA at National Institute of Science and Technology, Berhampur. His domain of research includes Biometric AuthenticationImage Processing, Pattern Recognition and Data Mining.

Charulata Palai Working as an Assistant Professor in the Dept. of CSE at National Institute of Science and Technology, Berhampur. Her domain of research includes Soft ComputingImage Processing, Pattern Recognition.

Lopamudra Sahoo is a BTech Final Year student of NIST, Berhampur.

Anshupa Patel is a BTech Final Year student of NIST, Berhampur.