Classification and Recognition Phase

Step 11-Feed Forward Back-propagation Neural Network is trained.
Step 12- The trained neural network is used for classification and recognition.

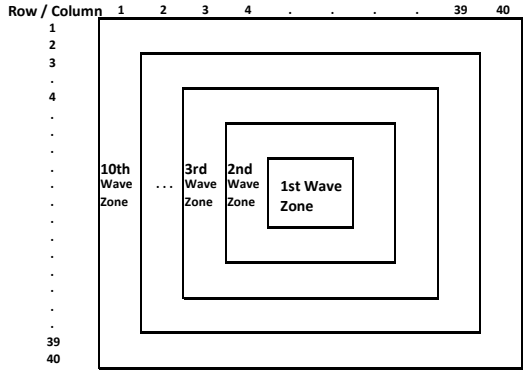

Figure 3: Division of an Image into 10 Wave Zones

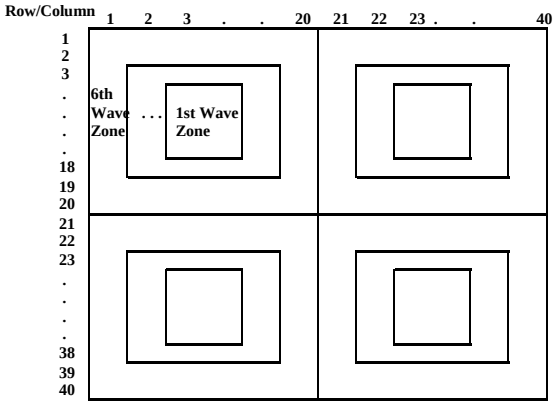

Figure 4: Division of Image into 4 Quadrants and Division of Each Quadrant into 6 Wave Zones

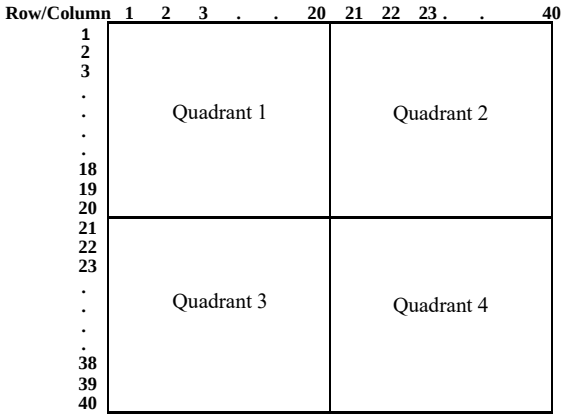

Figure 5: Division of Image into 4 Quadrants without wave zones

5. Implementation

Samples of all (13) Hindi vowels (SWARs) is collected on the white paper sheet of A4 size from different persons (150) of varying age group to create a database of 1950 samples. All the samples taken on the paper are converted to digital form by scanning the sheets through a flatbed scanner with an optical resolution of 2400 x 4800 dpi. Samples of handwritten Hindi SWARs collected from three different persons are shown in Table 1.

Table 1. Samples of Handwritten Vowels in Hindi Script Collected from Three Persons

S. No.	Sample 1	Sample 2	Sample 3
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			

To study the efficacy of the created feature vectors in offline handwritten character recognition following four strategies are considered:

Strategy I- Consisted of global wave feature vector A of length 10 for training the network.

Strategy II- Consisted of local wave feature vector B of length 24 for training the network.

Strategy III- Consisted of combined wave feature vector C = {A} {B} of length 34 for training the network.

Strategy IV- Consisted of combined feature vector D = {A} U{B}U{C} of length 38 for training the network.

MATLAB 2013a is used to carry out the experiments using the sample set of vowels. Keeping in view the complexity of character recognition problem, neural network architecture with two hidden layers is adopted. For training phase, Back-propagation Neural Network (BPNN) with two hidden layers (H1 and H2) is used. The output layer (O) of neural network consisted of thirteen neurons for representing the thirteen different classes of Hindi vowels whereas the number of neurons in the input layer (I) depended upon the lengths of feature vectors. Table 2 depicts the features, feature length and the architecture of neural network for different strategies.

Table 2. Features, Feature length and Architecture of Neural Network for Different Strategies

Strategy	Features	Length of Feature Vector	Neural Network Architecture (I-H1-H2-O)
I	Global wave feature	10	10-5-5-13
II	Local wave feature	24	24-12-12-13
III	Combined wave feature	34	34-24-13-13
IV	Combined feature	38	38-24-13-13

Once the neural network is trained to classify and recognize the new handwritten vowels (test samples) of Hindi script. Neural networks is trained for all the strategies I, II, III and IV. During the testing phase, 650 samples are selected randomly out of total 1950 samples.

5. Experimental Results

The performance of various methods in terms of training time (in second) and recognition rate (in %) is presented in Table 3. It is obvious from Table 3 that global wave features (strategy I) and local wave features (strategy II) individually achieved lower recognition rate 73.2% and 84.5%, respectively. But, when global wave features and local wave features are combined (strategy III) a higher recognition rate of 95.7% is achieved. Slightly better recognition rate of 96.2% is achieved in strategy IV when average intensity feature vector of four quadrants without wave zone is combined with feature vector of strategy III.

Table 3. Performance of Different Strategies

Strategy	Training Time (in seconds)	Recognition Rate (in %)
I	81	73.2
II	107	84.5
III	149	95.7
IV	161	96.2

Recognition rates (accuracies) of individual characters using Back-propagation Neural Network with four strategies are shown in Table 4. For strategy IV, more confusion was found in recognizing characters अ, ए, ओ and ऋ. Comparatively less confusion was recorded in recognizing characters ई, उ, ऐ, ओ, अं, and अः. The characters आ, ई and ऊ are recognised with 100% accuracy.

Table 4. Recognition Rate (in %) of Individual Characters with Different Strategies

S. No.	Hindi SWARs (Vowels)	Strategy I	Strategy II	Strategy III	Strategy IV
1	अ	70	90	98	90
2	आ	82	84	96	100
3	इ	60	76	98	96
4	ई	98	98	98	100
5	उ	86	90	96	96
6	ऊ	84	96	98	100
7	ए	48	98	100	92
8	ऐ	66	86	98	96
9	ओ	88	84	94	98
10	औ	82	88	94	94
11	अं	82	84	100	98
12	अः	56	62	98	98
13	ऋ	50	62	76	92

The performance comparison of the proposed approach with some existing approaches is presented in Table 5.

Table 5: Summary of Offline Handwritten Hindi/ Devanagri Script Character Recognition Systems

S. No.	Classifier Used	Character Set	Feature Extraction	Database	Author	Recognition Rate (%)
1	Fuzzy Sets	Handwritten Hindi Characters	Modified exponential functions.	Own	Hanmandlu et al. [8]	90.65
2	Fuzzy Sets	Hindi Numerals	Normalize distance features, Modified exponential functions	Own	Hanmandlu et al. [9]	96.0
3	Fuzzy Sets	Isolated, Devanagari Characters	Structural Features	Own	Mukherji and Rege [11]	86.4
4	Multi-layer Perceptron	Devanagari Characters	Three different feature sets	Own	Sandhya Arora et al. [12]	92.16
5	Backpropagation Neural Network	Handwritten Hindi Characters, only Five Starting Consonants	Binary vector of image	Own	Gunjan Singh et al. [14]	93.0
6	KNN	Handwritten Devanagari Vowels	Feature extraction using recursive sub-division techniques	ISI Kolkatta	Rakesh Rathi [15]	96.14
7	SVM	Hindi Characters	Diagonal feature extraction	Own	Sonika Dogra et al. [16]	93.06
8	Backpropagation Neural Network	Isolated Offline Handwritten Devanagari Numerals	Wavelets features	ISI Kolkatta	Gaurav Y. Tawde [17]	60 -70
9	Backpropagation Neural Network	Isolated Offline Handwritten Hindi Script Vowels	Wave based feature extraction	Own	Proposed Method	96.2

5. Conclusion

The work aims at proposing an efficient approach for feature extraction to enhance the accuracy offline handwritten character recognition of Hindi characters. A novel technique 'TARANG', inspired by wave motion in a medium, for feature extraction is proposed to recognize offline handwritten Hindi 'SWARs' (vowels).

It is shown that when a feature vector obtained by using proposed wave based technique is used for training of Backpropagation Neural Network, a recognition rate as high as 95.7% is achieved. Also, when a feature set obtained by combining features with and without wave based approach is used for training neural network, a recognition rate as high as 96.2% is achieved, which is comparatively very high and also comparable with the accuracies some existing approaches. Further, better recognition rate is expected by dividing the image into smaller sized portions and then using the local wave features and global wave features to form the combined wave feature vector. A better accuracy is also expected by increasing the data set for training and testing the neural network. It is, therefore, concluded that the combined wave based feature vector is capable of producing even better results. The proposed approach is only implemented and tested for the samples of Hindi SWARs. In future, the proposed approach may be used with the entire character set of Hindi script as well as for other scripts using neural network or other classification scheme.

References

[1] Laurene Fausett, Fundamentals of Neural Networks, Architectures, Algorithms, and Applications, Addision Weslley Publishing Company, December, 1993.

[2] Birijesh K. Verma, "Handwritten Hindi Character Recognition Using Multilayer Perceptron and Radial Basis Function Neural Networks", IEEE, 0-7803-2768-3,1995, pp. 2111-2115.

[3] J.A. Anderson, Introduction to Neural Networks, MIT Press, 1995.

[4] Robert J. Schalkoff, Artificial Neural Networks, McGraw-Hill International Editions, Computer Science Series, 1997.

[5] L. S. Oliveira, N. Benahmed, R. Sabourn, F. Bortolozzi, and C. Y. Suen, "Feature Subset Selection Using Genetic Algorithms for Handwritten Digit Recognition", IEEE, 1530-1834, Vol. 37, 2001, pp. 362-369.

[6] Utpal Garain, Mangal P. Chakraborty, Dipankar Dasgupta, "Recognition of Handwritten Indic Script Using Clonal Selection Algorithm", Springer Berlin/Heidelberg, 2006, pp. 256-266.

[7] Ranadhir Ghosh, Moumita Ghosh, "An Intelligent Offline Handwriting Recognition System Using Evolutionary Neural Learning Algorithm and Rule Based Over Segmented Data Points", Journal of Research and Practice in Information Technology, Vol. 37, No. 1, February 2005, pp. 73-87.

[8] M. Hanmandlu, O.V. Ramana Murthy, Vamsi Krishna Madasu "Fuzzy Model Based Recognition of Handwritten Hindi Characters", Digital Image Computing Techniques and Applications, IEEE, 0-7695-3067-2, 2007, pp. 454-461.

[9] M. Hanmandlu, A.V. Nath, A.C. Mishra, V.K. Madasu, "Fuzzy Model Based Recognition of Handwritten Hindi Numerals Using Bacterial Foraging", 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS), IEEE, 0-7695-2841-4, 2007.

[10] U. Pal, T. Wakabayashi, F. Kimura, "Comparative Study of Devnagari Handwritten Character Recognition using Different Feature and Classifiers", IEEE, 10th International Conference on Document Analysis and Recognition, 2009, pp. 1111-1115.

[11] Prachi Mukherji, Priti P. Rege, "Shape Feature and Fuzzy Logic Based Offline Devnagari Handwritten Optical Character Recognition", Journal of Pattern Recognition Research, 2009, pp. 52-68.

[12] Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri, Dipak Kumar Basu, Mahantapas Kundu, "Multiple Classifier Combination for Offline Handwritten Devanagari Character Recognition", arxiv.org/pdf/1006.5913, June, 2010, pp. 01-06.

[13] Sheetal Dabra, Sunil Agrawal, Rama Krishna Challa, "A Novel Feature Set for Recognition of Similar Shaped Handwritten Hindi Characters Using Machine Learning", CCSEA, CS & IT 02, 2011, pp. 25–35.

[14] Gunjan Singh, Sushma Lehri, "Recognition of Handwritten Hindi Characters using Backpropagation Neural Network", International Journal of Computer Science and Information Technologies, Vol. 3 (4) , 2012, pp. 4892-4895.

[15] Rakesh Rathi, Ravi Krishan Pandey, Vikas Chaturvedi, Mahesh Jangid, "Offline Handwritten Devanagari Vowels Recognition using KNN Classifier", International Journal of Computer Applications, 0975–8887, Volume 49, No. 23, July 2012, pp. 11-16.

[16] Sonika Dogra, Chandra Prakash, "Pahchaan: Hindi Handwritten Character Recognition System Based on SVM", International Journal on Computer Science and Engineering (IJCSE), 0975-3397 Vol. 4, No. 05, May 2012, pp. 718-722.

[17] Gaurav Y. Tawde, "Optical Character Recognition for Isolated Offline Handwritten Devanagari Numerals Using Wavelets", International Journal of Engineering Research and Applications, 2248-9622, Vol. 4, Issue 2 (Version 1), February 2014, pp.605-611.

[18] Kulkarni Sadanand A., Borde Prashant L., Manza Ramesh R., Yannawar Pravin L., "Offline MODI Character Recognition Using Complex Moments", Second International Symposium on Computer Vision and the Internet (VisionNet'15), Procedia Computer Science, 58516–523, 2001, pp-516-523.

Ajay Indian received the M.C.A. degree in 2001 from Dr. B.R. Ambedkar University, Agra, India and M.Tech. (Computer Science & Engineering) in 2010 from Gautam Buddha Technical University, Lucknow, India. Currently he is an Associate Professor at Department of Computer Science & Applications, Invertis University, Bareilly, India and a research scholar at Department of Computer Science, Gurukula Kangri Vishwavidyalaya, Haridwar, India. His research interests include Image Processing, Artificial Neural Networks and Data Mining.

Karamjit Bhatia received M.Phil. degree in Computer Applications in 1988 from University of Roorkee (now IIT Roorkee), Roorkee, India and received Ph.D. (Computer Science) in 2001 from Gurukula Kangri Vishwavidyalaya, Haridwar, India. Currently he is a Professor at Department of Computer Science, Gurukula Kangri Vishwavidyalaya, Haridwar, India. His research interests include Distributed Systems, Image Processing and Mobile Ad-hoc Networks. He has published around 40 research papers in International Journals of repute and National/ International conference proceedings. Five Ph.D. degrees have been awarded under his supervision.