Scientific Journal Impact Factor (SJIF) : 1.711
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745 ISSN (PRINT) : 2393 - 8161

Kannada Handwritten Character Recognition Using Genetic Algorithm

Vishwaraj[1], Karthik S P[2], Sreedharamurthy S K[3]
[1]*Department of E&CE, UBDTCE, Davangere,India*
[2]*Department of E&CE, UBDTCE, Davangere, India*
[3] *Associate Professor, Department of E&CE, UBDTCE, Davangere, India*

Abstract— Now a days, Genetic algorithms (GAs) have been increasingly applied in artificial neural network (ANN) design in different ways for optimization. Few examples are control parameter optimization, topology optimization and genetic training algorithms. The method proposed in this paper basically made use of conventional feature selection techniques to select a feature vector. Subset of this feature vector is obtained by using genetic algorithm. The so selected features are given to neural network for classification. Hence the proposed method is going to exploit the advantages of both Genetic algorithm as well as neural networks. Both of which are treated as soft computing techniques in computer domain

Keywords— Kannada Characters, Genetic algorithm, neural network**.**

I. INTRODUCTION

Optical character recognition (OCR) is the most challenging area in the field of pattern recognition. This OCR system uses the scanned or captured image of a text written on a paper and process it to give recognized text in most typesetting digital document. Due to the impact of digitalization and advancement of Information Technologies, the pattern recognition has gained the wide verity of application. The work proposed here is to recognize the handwritten Kannada character. Various authors have already worked on Kannada handwritten character recognition using various methods like neural network, SVM, template matching etc. This hybrid system is to explore the advantages of the Genetic algorithm and neural network. This work is intended to make the input text image as machine editable format with software like Baraha or Nudi. Thus this not only helps in digitisation but can also make it editable. Here we are considering isolated Kannada characters only. In this work 10 samples of image are considered, thus for 49 characters, there will be 490 images.

II. METHODOLOGY.

Following block diagram defines the brief explanation about methodology

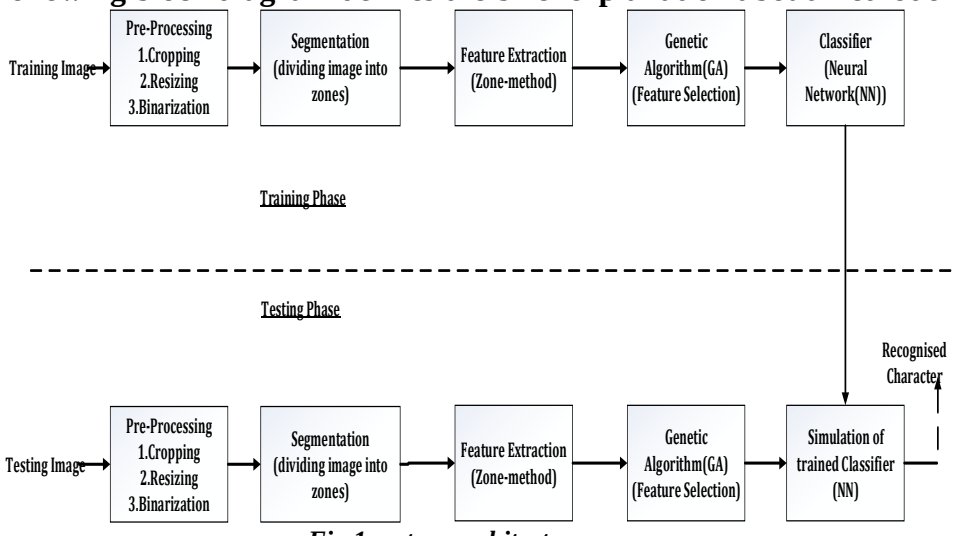

Fig.1 system architecture

The system mainly consists two phases; training phase and testing phase. Each phase has five blocks, namely, preprocessing, segmentation, feature extraction. GA, and neural network.

i. Pre-processing: Pre-processing is done remove unwanted information and retain the enhanced image which is suitable for required operation. The various pre-processing operation done here are cropping, resize, binarization and morphological thinning.

ii. Segmentation: It is portioning the image into boxes with are required for feature extraction. In this work 50x50 size image is considered, after segmentation 25 parts of image having 5x5 size are obtained.

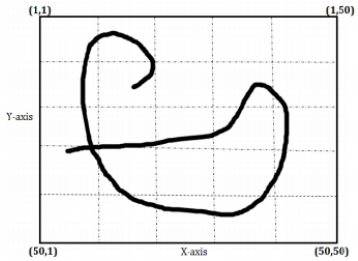

Fig 2 Segmentation

iii. Feature extraction: A pixel based method is used to extract the features. This methods considers the number of black pixels in the image which represents the character on foreground.

Feature $=\frac{number\ of\ black\ pixels}{totla\ number\ of\ pixels}$ ..... (1)

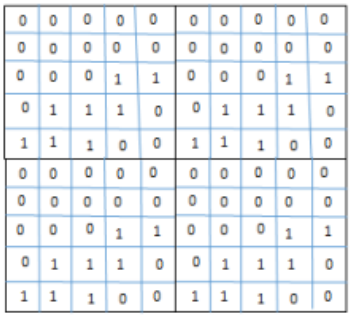

Fig 3 feature extraction

iv. Feature Selection: This is done with the help of genetic algorithm. Using above 25 features best features are selected with the application of GA.

v. Training Neural Network: The selected features from GA are trained in neural network which is the classifier.

vi. Finally, at last, a test image is considered to check the accuracy of the whole system which under goes same procedures as during taring phase.

III. GENETIC ALGORITHM

Genetic Algorithm is an evolutionary algorithm based on the law “survival of fittest”, thus it is also based on the fittest function.GA is mainly used for the optimizing the search. Inheritance, mutation, crossover, selection are the genetic operations performed in this algorithm.

Table 1 Parameters used in Genetic algorithm

Parameters	Value
Initial Population size	100
Genome length	25
Fitness Function	kNN
Generation size	300
Probability of Crossover	0.8
Probability of Mutation	0.1
Selectin type	Tournament, its size is 2
Elitism size	2

The feature selection process is shown as below.

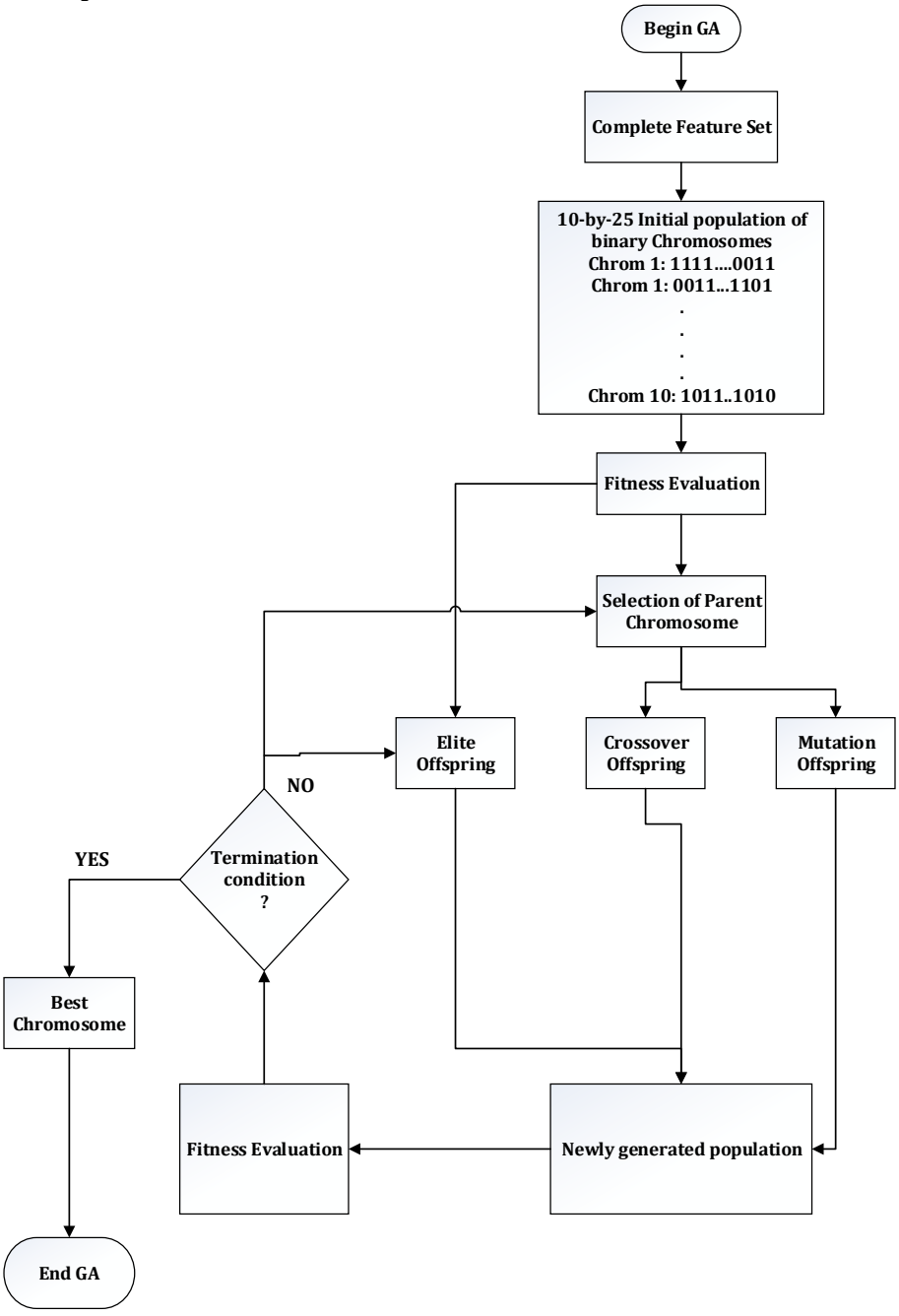

Fig 6. Feature selection method using GA.

EXPERIMENTAL RESULTS

The various statistical analysis like mean, standard deviation and variance are tabulated below. From the tables given below its evident that the close resemble in the values and very much less difference among the values

Table 2 Statistical Table for Sawaragalu

Character	Mean	Std. deviation	Variance
ಅ	0.1030	0.0813	0.0068
ಆ	0.0850	0.0780	0.0062
ಇ	0.0707	0.0640	0.0042
ಈ	0.0774	0.0810	0.0067
ಉ	0.0872	0.0706	0.0050
ಊ	0.0850	0.0709	0.0051
ಋ	0.1072	0.0735	0.0054
ಎ	0.0702	0.0735	0.0055
ಏ	0.0601	0.0689	0.0048
ಐ	0.0678	0.0672	0.0045
ಒ	0.0701	0.0718	0.0052
ಓ	0.0773	0.0814	0.0067
ಔ	0.0854	0.0799	0.0064
ಅಂ	0.1132	0.0852	0.0074
ಅಃ	0.0960	0.0830	0.0070

Table 3 Statistical Table for Vargiya Vyanjanagalu

Character	Mean	Std. deviation	Variance
ಕ	0.1001	0.0774	0.0611
ಖ	0.0954	0.0697	0.0049
ಗ	0.0440	0.0615	0.0039
ಘ	0.0838	0.0781	0.0063
ಙ	0.0864	0.0731	0.0054
ಚ	0.0588	0.0738	0.0056
ಛ	0.0879	0.0700	0.0050
ಜ	0.0778	0.0732	0.0054
ಝ	0.1027	0.0608	0.0038
ಞ	0.0808	0.0731	0.0054
ಟ	0.1007	0.0728	0.0054

ರ	0.0933	0.0758	0.0058
ಡ	0.0966	0.0809	0.0066
ಢ	0.0809	0.0737	0.0055
ಣ	0.0738	0.0715	0.0052
ತ	0.0850	0.0833	0.0071
ಥ	0.0753	0.0744	0.0056
ದ	0.0747	0.0826	0.0068
ಧ	0.0652	0.0635	0.0041
ನ	0.0903	0.0876	0.0077
ಪ	0.0839	0.0798	0.0064
ಫ	0.0787	0.0713	0.0051
ಬ	0.0818	0.0864	0.0075
ಭ	0.0825	0.0756	0.0057
ಮ	0.0914	0.0599	0.0036

Table 4 Statistical Table for Avargiya Vyanjanagalu.

Character	**Mean**	**Std. deviation**	**Variance**
ಯ	0.0825	0.0684	0.0048
ರ	0.0794	0.0665	0.0044
ಲ	0.0760	0.0721	0.0053
ವ	0.1114	0.799	0.0065
ಶ	0.0994	0.0094	0.0099
ಷ	0.0856	0.0798	0.0051
ಸ	0.0798	0.0826	0.0069
ಹ	0.0933	0.0812	0.0067
ಳ	0.0760	0.0877	0.0964

CONCLUSION

The whole system can be ratiocinated that the proposed recognised system has about 71.63% of recognised rate. This performance is due the feature extraction method. The method, zone-method, used here is of basic one. Which explores only the black pixels and not any important or unique features like Geometric moment, etc. Some mismatch were also seen because of the close resembles in feature values. As the difference between neighbouring features values increases the accuracy also increases. This is evident from the statistical table given above. One can improve the same system with some other feature extraction techniques, which makes the feature values distinct. Due to the application of Genetic algorithm, the feature selection had been done very easily and appropriately, hence out of 25 features the only valid and best 12 features were selected for training neural network, this could also replicated in testing and training time. In total 490 images were used for training and for testing 490

images (10images/character) were considered. Among 490 images, 351 images were recognised. Thus the recognised ratio is considered as 71.63%.

REFERENCE

[1] Anshul Gupta, Manisha Srivastava and Chitralekha Mahanta, Offline Handwritten Character Recognition Using Neural network, 2011 International conference on computer Applications and Industrial Electronics (ICCAIE 2011)
[2] Pratibha A.Desai, Sumanagala N bhavikatti, and Rajashekar Patil, Neural networks based offline handwritten Character recognition system, proceeding of SARC-IRAJ International Conference, 16th June 2013, Pune,India.
[3] P.M Wagh and S.P. Ramteke, A review on optical character recognition system using Genetic Algorithm, International Journal of advanced electronics and communication system, proceedings of International Conference on modelling and simulation in Engineering and technology (ICMSET-2014) 15th-16th Feb, 2014
[4] Anjali D. Shah and A.S. Bhide, A Genetic Learning Method for OCR in English Language. PUSER 12, Vol.02, Issue:02/06 March-April; Bimonthly International Journal; Page(s) 241-247
[5] Vijaya S, Waghmare, Leena R. Ragha, Application of Genetic Algorithm for Handwritten Kannada Characters, 2013 3rd IEEE International Advance Computing Conference (IACC).
[6] Prof. M Hanmandlu, Manmeet Singh Bawa, S.Raghvan and Karthik Venkataraman H, fuxxy based approach to the recognition of handwritten numerals.
[7] Babatunde Oluleye, Armstrong Leisa, Jinsong Leng and Diepeveen Dean, A genetic algorithm based feature selection, International journal of Electronics communication and computer Engineering Vol 5.Issue 4.
[8] Vedgupta Saraf, D.S Rao, "Devanagari Script Character Recognition Using Genetic Algorithm for Get Better Efficiency", International Journal of soft Computing and Engineering, V2 issue 4, April 2013
[9] Shashank Mathur, Vaibhav Aggarwal, Himanshu Joshi, Anil Ahlawat, "Offline Handwritten recognition using Genetic Algorithm" International book series "Information Science and computing June-July 2008
[10] Rahul KALA, Harsha VAZIRANI, Anupam SHUKLA and Ritu TIWARI, "Offline Handwriting Recognition using Genetic Algorithm" IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010.
[11] E.K Vellingiriraj, P. Balasubramanie "Recognition of Ancient Tamil Handwritten Characters in Palm Manuscripts using Genetic Algorithm", International Journal os scientific Engineering and technology V2 Issue 5,pp: 342-346
[12] Huan Chen, Xiangnian Huang, Ping Cheng, Yan Hu "Offline handwritten Chinese character recognition based on GA optimization BACK PROPORTION neural network" International conference on Information, Business and education Technology, 2013

