SOCIETY FOR SCIENCE AND EDUCATION
UNITED KINGDOM

TMLAI TRANSACTIONS ON MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Volume 6 No. 1
ISSN 2054-7390

Comparison Performance Evaluation of Modified Genetic Algorithm and Modified Counter Propagation Network for Online Character recognition

Adigun Oyeranmi. J[1], Fenwa Olusayo D[2], Babatunde. Ronke. S[3]
[1]Department of Computer Technology, School of Technology, Yaba College of Technology, Yaba, Lagos Nigeria
[2]Department of Computer Science and Engineering, Faculty of Engineering and Technology, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.
[3]Department of Computer Science, College of Information and Communication Technology, Kwara State University, Malete, Nigeria.
odfenwa@lautech.edu.ng, ranmi.adigun@yabatech.edu.ng, ronke.babatunde@kwasu.edu.ng

ABSTRACT

This paper carries out performance evaluation of a Modified Genetic Algorithm (MGA) and Modified Counter Propagation Network (MCPN) Network for online character recognition. Two techniques were used to train the feature vectors using supervised and unsupervised methods. First, the modified genetic algorithm used feature selection to filter irrelevant features vectors and improve character recognition because of its stochastic nature. Second, MCPN has its capability to extract statistical properties of the input data. MGA and MCPN techniques were used as classifiers for online character recognition then evaluated using 6200-character images for training and testing with best selected similarity threshold value. The experimental results of evaluation showed that, at 5 x 7 pixels, MGA had 97.89% recognition accuracy with training time of 61.20ms while MCPN gave 97.44% recognition accuracy in a time of 62.46ms achieved. At 2480, MGA had 96.67% with a training time of 4.53ms, whereas MCPN had 96.33% accuracy with a time of 4.98ms achieved. Furthermore, at 1240 database sizes, MGA has 96.44 % recognition accuracy with 0.62ms training time whereas MCPN gave 96.11% accuracy with 0.75ms training time. The two techniques were evaluated using different performance metrics. The results suggested the superior nature of the MGA technique in terms of epoch, recognition accuracy, convergence time, training time and sensitivity.

Keywords: Character recognition, normalization, Modified genetic Algorithm(MGA), Modified counter propagation network (MCPN), Generation gap

1 Introduction

Characters recognition have been explored for many years and require many applications capabilities such as bank processing, person authentication, manufacturing of plastic cards, food industry etc. So far, online character recognition remains an open problem, in spite of a dramatic boost of research in this field and the latest improvement in recognition methodologies. Character recognition is an ongoing research which has motivated researchers from various aspect of human endeavors such as image processing, computer

DOI: 10.14738/tmlai.61.3976
Publication Date: 06th January, 2017
URL: http://dx.doi.org/10.14738/tmlai.61.3976

vision and machine learning [1]. Online character recognition system is the transformation process, that will allow extraction of input characters, from character image database to digitalize and translates the handwritten text into a machine editable form [2]. Due to global security threat, person authentication, and retrieving text, there is need of adopting techniques that could enhance the recognition performance of the system. Features extraction techniques and classifiers had been researched upon to have contributed to the performance of character recognition. At the same time, evaluation of the system with some selected metrics such as recognition accuracy, sensitivity, computation time etc. have been considered. Recognition of handwriting in online mode is usually accomplished using temporal spatial information obtained from the operativeness of a stylus on the surface of an electrostatic or electromagnetic tablet. Timing knowledge is accessible from Coordinate information of strokes [3]. The most popular method used in online character recognition is backpropagation network. Genetic algorithm was used to optimize backpropagation architecture and used to find an optimal solution in complex problems [4] that mimic the principle of natural genetics and natural selection that competes for survival to make up the next generation of population [5][6]. However, backpropagation neural network weakness are slow convergence and long time for recognition of characters. Character recognition consists mainly of four stages: data acquisition, pre-processing, feature extraction and classifications [7]. A modified counter-propagation (MCPN) algorithm is one the neural network classifiers with a supervised and unsupervised training which are closely related to the Nearest-Neighbor classifier. The network essentially functions as a nearest match lookup table. MCPN network has an interesting capability to extract the statistical properties of the input data, and can usually be trained very rapidly [8]. This study used statistical and structural features to extract features at MGA from the image characters and Integration of Geometrical and Statistical features had been used as feature extraction at MCPN. The classification of the developed system was carried out using both techniques. Hence, the focus of this paper is to evaluate performance of MGA and MCPN classifiers to recognize character images and establish the efficiency of the two techniques.

2 Related Work

High recognition accuracy and less training time are the bane of online character recognition. This could be done based on the feature extraction techniques and classifiers algorithm applied. [9] compared two different algorithms with combination of genetic algorithm with backpropagation network(GABPN) and correlation method with genetic algorithm(GA) to achieve both accuracy and training swiftness for recognizing alphabets. The performance evaluation of the two algorithms showed that correlation with GA technique gave the best recognition accuracy of 95-97%. [10] implemented two different techniques (structural and statistical) and MGA was used to extract optimized feature subset of the character for classification task. Two classifiers (C1 and C2) were formulated from MGA-MOBP. The overall results indicated that C2 achieved better performance than C1. A modified CPN was proposed by Fenwa in 2012 for online character recognition, which was faster than the existing conventional CPN. In the modified CPN model, character parameters were not trained like backpropagation architecture which is an interactive method that suffer long training. The system was experimented on different handwritten characters. The performances of proposed techniques were evaluated in terms of recognition time and recognition rate. [11] implemented a research on integration of PSO with hybrid of Counterpropagation and modified Optical Backpropagation Neural Network (COMOB) model to enhancement the

performance of the classifier in terms of recognition accuracy and recognition time. Results of the proposed system show that there was a better performance of recognition accuracy. However, [12] presented a paper for recognition of English characters based on features derived from partitioning the characters into non-overlapping cells. This work applied a dynamic window sliding for feature vector generation and uses four passes of window that led to the creation of a 30-element feature vector. A neural network (multi-layered perceptron) was used for classifying the 26 alphabets of the English language. The average recognition rate achieved was about 97.33%. [13] developed a technique for recognition of an offline handwritten character using grip approach. Extracted features are trained by neural network as classifier of the character in classification stage. The overall experimental result in recognition rate was 96.9%. [14] designed handwritten digits recognition system with combination of genetic algorithm. The proposed neural network was trained with data set containing 25 sample images of each digit. The average accuracy result of the proposed system was 99% and this was then compared with currently existing techniques with various constraints.

3 Methodology

The paper focuses on implementing MGA and MCPN classifiers on image dataset, evaluates the recognition accuracy, training time performance of the algorithms and comparing their results. The images recognition process consist of four stages: image acquisition, image preprocessing, feature selection &extraction and the classification.

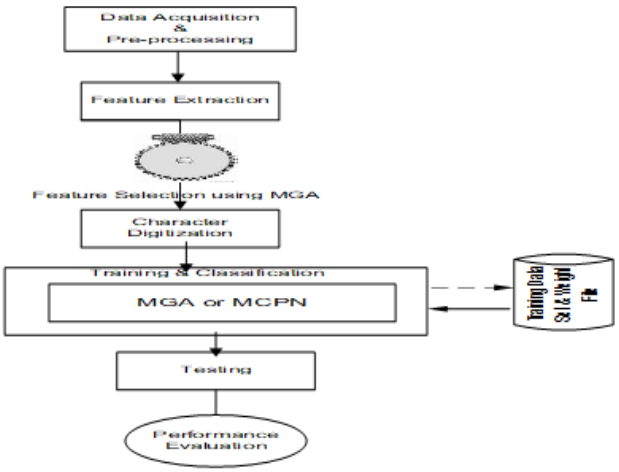


Figure 1: The Block Components of Comparison Evaluation of MGA and MCPM

3.1 Acquisition and Preprocessing

In this article, the first step is the character images were acquired using a pen digitizer from datasets of 6200 samples collected by (Adigun et al., 2016) was used for training. Three preprocessing techniques were employed: binarization, extreme coordinate measurement and grid resizing were used to convert into binary form, measure extreme coordinate of the space and matrix standard respectively. The normalization the images to remove the noise, gaps and character enhancement. Finally, classification of individual images based on input image was tested using MGA and MCPN classifiers. The extracted feature provided the characteristics of input type to classifier by considering the description of the relevant properties of image into feature space. The block diagram procedures to achieve this research work is shown in figure 1

Copyright © Society for Science and Education United Kingdom



(a) (b)

Figure 2a & b: Preprocessing System Result

3.2 Feature Selection and Extraction

This research work conducts a comparison evaluation of MGA and modified CPN techniques to recognize online characters. Summarily, MGA and MCPN were used as classifiers. The more efficient of these two techniques was checked and their recognition rate were evaluated. The character image in Fig. 2a is trained by undergoing preprocessing stage as shown above through conversion into grayscale and histogram equalization. This is to enhance the intensity of the images without losing any important information. The paper used hybrid (Struct-Statistical) Feature Extraction Algorithm developed by [4] to complement each other and reduce errors as shown figure 3. The MGA as feature selection was used to select optimized features subset to reduce redundant features which improves the recognition accuracy of the characters. The structural features used in this paper consists of stroke information, projection and invariant moments. This measures the pixel distribution around the centre of gravity of the character and allow capturing the global character shape information that improved the recognition accuracy. To calculate moment invariants as a

$$M_{pq} = \int \int \mathrm{x}^p y^q f(\mathrm{x},\mathrm{y})\mathrm{dxdy}$$

where f (x, y) is the intensity function representing the image, the integration is over the entire image and the F (x, y) is same function of x and y for example $x^p y^q$, or a sin(xp) and cos(yq). This is to determine the position, size and orientation of the character images. The statistical features adopted in this paper was hybrid Zoning algorithms of modified Image Centroid and Zone-based (ICZ) and modified Zone Centroid and Zone-based (ZCZ) distance metric feature extraction based on model developed by Fenwa *et al.,* (2012) was adopted see below:

Input: Pre-processed character image
Output: Features for Classification and Recognition
Begins
Step 1: Divide the input image into 25 equal zones
Step 2: Compute the input image centroid
Step 3: Compute the distance between the image centroid to each pixel present in the zone
Step 4: Repeat step 3 for the entire pixel present in the zone
Step 5: Compute average distance between these points
Step 6: Compute the zone centroid
Step 7: Compute the distance between the zone centroid to each pixel present in the zone.
Step 8: Repeat step 7 for the entire pixel present in the zone

Step 9: Compute average distance between these points
Step 10: Repeat the steps 3-9 sequentially for the entire zones
*Step11: Finally, 2*n (50) such features are obtained for classification and recognition. Ends*

As a result of global security threat and criminal activities, there is need of adopting techniques that could enhance the recognition performance of the system. Features extraction algorithms and classifiers had been researched upon to have contributed to the performance of the system.

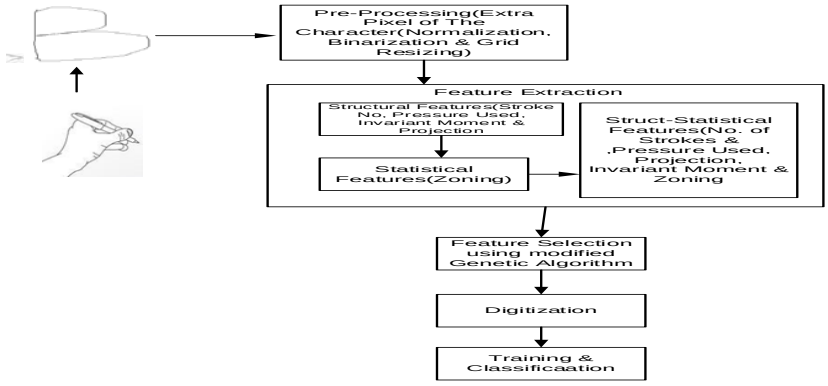


Figure 3: Developed Feature Extraction Model source: (Adigun et al., 2016)

3.3 Character Classification Method

The classification stage comprises of the training stage and the testing stage. The output of the features extracted from the training and testing images were saved and fed into the MGA and MCPN classifiers for comparison. The feature vectors generated are then compared with stored pattern and find out the best matching class for input. However, evaluation of the developed system with some selected metrics such as recognition accuracy, sensitivity, training time and computation time etc. have been considered. A threshold is determined by the continuous modification of the threshold until significant accuracy is observed. The threshold set implies correct match of a character which dependent on the minimum distance is less or equal to a threshold. The testing phase of the implementation was done in a straightforward manner. The program was coded into modular parts and the same routines of loading, analyzing, and computation of network parameters of input vectors in the training phase were reused in the testing phase as well. All the 6200 images of English characters acquired were used for training in this research work. The testing (900) images were introduced by 25 different persons one by one to see by if it will be recognized by the character recognition system.

Algorithm:

The basic steps in testing input images for characters can be summarized as follows:

i. Start
ii. Load character from the database
iii. Convert images into grayscale, then map into an input vector form and normalize image vector
iv. Apply Stat-struct technique for feature extraction and MGA for feature selection and store features
v. Apply MGA or MCPN classifier and compute output
vi. Match the introduced image with the ones in the database template

Copyright © Society for Science and Education United Kingdom

vii. Classify the image into (CR, FR and RF)

viii. Convert the binary output to the corresponding character and display to a message box

ix. Test the next character image and repeat until all characters were visited

The simulation tool used for this research is C# programming language, an objected oriented programming language, derived from C++. It supports window based application, 64bits operating system, 8.00GB RAM and run under Windows 7 Professional Operating system on Intel® Core(TM) i5-42000M CPU @2.50GHz processor and built for .Net platform. The performance of the results was evaluated using generation gap, database sizes and overall accuracy (correct recognition(CR), false recognition(FR) and recognition failure(RF) to determine the performance and accuracy of the system. The sensitivity of the character recognition is determined by

$$\text{Sensitivity} = \frac{CR}{CR + FR}$$

4 Results and Discussion

The performance of MGA and MCPN on trained and recognized character were measured to determine its efficiency in terms of training time, epoch, convergence time, sensitivity and recognition accuracy,540 images were used for testing. The results obtained by using MGA and MCPN classifiers with respect to metrics mentioned above were evaluated as follows: Training time was estimated for different database sizes of the same 5*7-pixel resolution. The result obtained was as shown in Table 1, in which the training time of MGA was smaller than MCPN due to its stochastic capability and ability to achieve feature selection reduction. The average training time of the developed system is very much less when compared with MCPN classifier. The MGA recorded 0.62 ms averagely with 1240 database size, 4.53 ms with 2480 database size, 61.20 ms with 6200 database sizes whereas MCPN at 1240 database size had 0.75ms averagely, at 2480 the training time of 4.98ms and 62.46 with 6200 database sizes.

Table 1 Evaluation variation of database size on training time (milliseconds) accuracies of MGA and MCPN

Character Samples		MCPN				MGA		
	Training time	CR	FR	RF	Training time	CR	FR	RF
1240	0.75	865	30	5	0.62	868	27	5
2480	4.98	867	27	4	4.53	870	26	4
6200	61.20	877	21	1	61.20	881	18	1

From Table 2, the recognition accuracy obtained using MGA with different generation gap threshold values of 0.1, 0.3, 0.5 and 0.7and the study revealed that MGA has better performance in convergence time and accuracy than MCPN as computed in section 3.1. The recognition accuracy at 6200 database sizes with MGA recognition accuracy of 97.89% at 0.1, 97.34% at 0.3, 96.81% at 0.5 and 96.18 % at 0.7 whereas, MCPN obtained 97.44 % at 0.1, 97.1% at 0.3, 96.37% at 0.5 and 95.93% at 0.7 generation gap. Table 2 deduced the performance of MGA against MCPN different database sizes of 1240, 2480, and 6200.The average convergence time reported at 6200 database sizes with MGA are 193.41ms, while MCPN produced 199.01 ms. However, MGA was noticed to classified faster because of its feature selection

reduction capability than MCPN. Table 2 shows that values generated in terms of convergence time by MGA took less time than that of MCPN at different database sizes level employed. Furthermore, Table 2 shows that MGA in terms of sensitivity at different sizes had an increase value compared with MCPN. The MGA generated high recognition accuracy at a less time with MCPN. The MCPN also has its capability to extract statistical properties of the input data. Finally, the results of evaluation showed that MGA distinctively outperformed MCPN in terms of recognition accuracy, faster convergence time and less training time.

Table 2 showing combined results with MGA and MCPN at best selected threshold value

.

Database Size	Algorithm	Epoch	Convergence Time (milliseconds)	Sensitivity (%)	Accuracy (%)
1240	MGA	403	49.79	96.98	96.44
	MCPN	413	49.98	96.65	96.11
2480	MGA	771	201.19	97.1	96.74
	MCPN	780	200.46	96.9	96.33
6200	MGA	1838	193.41	98.0	97.89
	MCPN	1898	199.01	97.66	97.44

It was shown in Table 2 that the higher the database sizes, the better the recognition accuracy due to fact the network training was able to attribute the test character to larger character sample in the vector space. Usually, the complex and large sized input sets require a large topology network with more number of iterations (Epochs). The epochs are directly proportional to the training time, this implies that the larger the image size, the more the training time. However, this would also imply more number of iterations were required to reach its optimal state.

5 Conclusion

The study has presented a comparative performance evaluation of modified genetic algorithm and modified counter propagation Neural Network as well as their application with online handwritten character recognition. MGA, an optimization technique was used to extract salient features from online handwritten character images at the initial stage before the application of MGA and MCPN classifiers. This was done to reduce the insignificant features for enhancing and efficient character recognition. This research work was implemented and evaluated in order to determine their effectiveness. The results suggest that MGA recorded better convergence time and recognition accuracy than MCPN. In view of the MGA would find the near global optimal solution in a large solution space quickly. It can also be used extensively in many application areas, such as image processing, pattern recognition, feature selection, criminal activities and machine learning.

REFERENCES

[1] Abed Majida Ali, Ismail Ahmad Nasser and Hazi Zubadi Matiz., *Pattern recognition Using Genetic Algorithm,* International Journal of Computer and Electrical Engineering, 2013. 2(3): p. 583-588.

[2] Adigun J.O., Fenwa O.D., Omidiora E.O., Olabiyisi S.O. Optimized features for genetic based neural network model for online character recognition. British Journal of Mathematics & Computer Science, 2016, 14(6):1-13.

Copyright © Society for Science and Education United Kingdom

[3] Herekar Rachana R. , S. R. Dhotre, *Handwritten Character Recognition Based on Zoning Using Euler Number for English Alphabets and Numerals*. In IOSR e-ISSN: 2278-0661, 2014, 16(4), p.75-88.

[4] Adigun J.O., Omidiora E.O., Olabiyisi S.O., Fenwa O.D., Oladipo O, Rufai M.M. *Development of a Genetic based neural network system for online character recognition*. International Journal of Applied Information Systems (IJAIS), 2015, 9(3), p.1-8.

[5] Noaman, Khaled M.G, Saif, Jamil Abdulhameed M., Alqubati, Ibrahim A.A, *Optical Character Recognition Based on Genetic Algorithms*, Journal of Emerging Trends in Computing and Information Sciences, 2015, 6(4), p.204-208.

[6] Yeremia Hendy, Niko Adrianus Yuwono, Pius Raymond and Widodo Budiharto, *Genetic Algorithm and Neural Network for Optical character recognition*, Journal of computer science, 2013, 9 (11),1435-1442.

[7] Fenwa, O.D., Omidiora, E.O. and Fakolujo, O.A. *Development of a Feature Extraction Technique for Online Character Recognition System*, Journal of Innovative System Design and Engineering, International Institute of Science, Technology and Education, New York, USA, 2012, 3(3), p.10-23.

[8] Kumar D, Rai CS, Kumar S. *An experimental comparison of unsupervised learning techniques for face recognition*. International Journal of Computer and Information Science and Engineering. 2007,1(3):158-166.

[9] Inamdar Farhan, and Bagal, S. B. (2016): *Comparative Study of Optical Character Recognition Techniques,* International Journal of Innovative Research in Computer and Communication Engineering, 2016,4(11), p.19831-19837

[10] Oyeranmi Adigun, Elijah Omidiora and Mohammed Rufai, *Modified Genetic Algorithm Parameters to Improve Online Character Recognition*, British Journal of Applied Science & Technology, 2016, 18(5): 1-8.

[11] Olusayo D. Fenwa, Funmilola A. Ajala and Alice O. Oke, *A PSO-Based Modified Counter Propagation Neural Network Model for Online Handwritten Character Recognition System*, International Journal of Emerging Technology and Advanced Engineering, 2014, 4(6), p.768 -776

[12] Biswas, Mithun and Parekh, Ranjan, *Character Recognition using Dynamic Windows,* International Journal of Computer Applications, 2012, 41(15):47-52.

[13] Chaudhari Prasad P, Sarode KR. *Offline handwritten character recognition by using grid approach.* International Journal of Application or Innovation in Engineering & Management, 2014, 3(4), p.71-73.

[14] Kaur Tarandeep and Chabbra Amit, Genetic Algorithm Optimized Neural Network for Handwritten Character Recognition, International Journal of Computer Applications, 2015, 119(24), p.22-26