A GA-FUZZY Approach to Evolve Hopfield Type Optimum Networks for Object Extraction

Susmita Ghosh (nee De)[1] and Ashish Ghosh[2]

[1] Department of Computer Science & Engineering,
Jadavpur University, Kolkata 700 032, India
susmita_de@hotmail.com
[2] Machine Intelligence Unit,
Indian Statistical Institute, Kolkata 700 035, India
ash@isical.ac.in
http://www.isical.ac.in/~ash

Abstract. Earlier attempts are made to design Hopfield type neural network architecture for object extraction using Genetic Algorithms (GAs). Energy value of the neural network was taken as the index of fitness of the GA. In the present article fuzzy logic reasoning is incorporated into this Neuro-GA hybrid framework to remove some of the drawbacks of earlier attempts. Here, GAs have been used to evolve Hopfield type optimum neural network architecture for object background classification. Each chromosome of the GA represents an architecture. The output status of the neurons at the converged state of the network is viewed as a *fuzzy set* and measure of fuzziness of this set is taken as a measure of fitness of the chromosome. The best chromosome of the final generation represents the optimum network configuration. When the input images are less noisy, the evolved networks are found to have less (compared to the corresponding energy based objective evaluation) connectivity for providing comparable outputs.

1 Introduction

Neural networks (NNs) are designated by the network topology, connection strengths between pairs of neurons/nodes, node characteristics and rules for updating status. In spite of the wide range of applicability of NNs, there is no formal procedure to design an optimum neural network for a given problem. Recently, some attempts are made in this regard using Genetic Algorithms (GAs) [1]- [4]. Each chromosome of a GA represents a network architecture. In one of our earlier attempts, optimum Hopfield type networks [5] were evolved using GAs for extracting object regions from gray images [6, 7]. In that work, the energy value (which involves the connectivity also) of the converged network was taken as a measure of fitness of the corresponding chromosome. Thus, two different output images may have the same energy value and also the same set of output images may correspond to networks having different energy values (due

N.R. Pal and M. Sugeno (Eds.): AFSS 2002, LNAI 2275, pp. 444–449, 2002.

to different connectivity configuration). This may cause problems if energy value is considered as a fitness measure. For example, a network with more number of connections will provide more fitness value than the network with less number of connections even if they produce the same output image. This drawback can be overcome if the output status of the neurons in the converged state of each network is considered as a *fuzzy set* and a *fuzziness measure* of this set is taken as a measure of error/instability of the network which in turn reflects the fitness of the corresponding chromosome. In the present work, different types of fuzziness measures, namely, index of fuzziness, entropy of a fuzzy set and fuzzy correlation are used [8]. The working principle of object extraction using this network is similar to that used in [7]. The proposed technique is compared with its energy based counterpart [7]. It is found that in case of less corrupted images, to produce comparable segmented output, fuzziness measure based objective function evolves networks having less connectivity than the corresponding networks evoluted by energy based model.

2 Measures of Fuzziness of a Fuzzy Set

A fuzziness measure [8, 9] of a fuzzy set expresses the average amount of ambiguity in making a decision whether an element belongs to the set or not. Several authors have made attempts to define such measures. A few measures relevant to the present work are described here.

2.1 Index of Fuzziness

The index of fuzziness (*f1)* of a fuzzy set A having n supporting elements is defined as

$$\begin{aligned}\gamma_p(A) &= \frac{2}{n^{\frac{1}{p}}}\, l^p(A,\overline{A}) \\ &= \frac{2}{n^{\frac{1}{p}}}\left[\sum_{i=1}^{n}\{\min(\mu_A(x_i),\ 1-\mu_A(x_i))\}^p\right]^{\frac{1}{p}},\end{aligned} \tag{1}$$

when $l^P(A,\overline{A})$ denotes the distance between fuzzy set A and its nearest ordinary set $\overline{A}$. An ordinary set $\overline{A}$ nearest to the fuzzy set A is defined as:

$$\mu_{\overline{A}}(x) = \begin{cases} 0 \text{ if } & \mu_A(x) \le 0.5 \\ 1 \text{ if } & \mu_A(x) > 0.5. \end{cases} \tag{2}$$

We have considered quadratic index of fuzziness *i.e.*, $p = 2$.

2.2 Entropy

Entropy of a fuzzy set as defined by De Luca and Termini (*f2*) is given by

$$H(A) = \frac{1}{n \ln 2}\sum_{i=1}^{n}\{S_n(\mu_A(x_i))\}, \tag{3}$$

with

$$S_n(\mu_A(x_i)) = -\ \mu_A(x_i)\ \ln\{\mu_A(x_i)\} \\ -\ \{1-\mu_A(x_i)\}\ \ln\ \{1-\mu_A(x_i)\}; \tag{4}$$

and that of Pal and Pal (*f3*) is given by

$$H(A) = \frac{1}{n(\sqrt{e}-1)} \sum_{i=1}^{n} \{S_n(\mu_A(x_i)) - 1\} \tag{5}$$

with

$$S_n(\mu_A(x_i)) = \mu_A(x_i)\ e^{1-\mu_a(x_i)} + \{1-\mu_A(x_i)\}\ e^{\mu_A(x_i)}. \tag{6}$$

Another definition of entropy which involves the distance of the fuzzy sets from its furthest ordinary set is given by Bart Kosko (*f4*). It says

$$\begin{aligned} R_p(A) &= \frac{l^p(A,\overline{A})}{l^p(A,\underline{A})} \\ &= \frac{\left[\sum_{i=1}^{n} \{\min(\mu_A(x_i), 1-\mu_A(x_i))\}^p\right]^{\frac{1}{p}}}{\left[\sum_{i=1}^{n} \{\max(\mu_A(x_i), 1-\mu_A(x_i))\}^p\right]^{\frac{1}{p}}} \end{aligned} \tag{7}$$

where $\underline{A}$ is an ordinary set furthest to the fuzzy set A, defined by

$$\mu_{\underline{A}}(x) = \begin{cases} 0 \text{ if } & \mu_A(x) \geq 0.5 \\ 1 \text{ if } & \mu_A(x) \leq 0.5\ . \end{cases} \tag{8}$$

2.3 Fuzzy Correlation

A concept of correlation (*f5*), giving a measure of relationship between two representatives of an image was given by Pal and Ghosh. Correlation between a fuzzy property and its nearest two tone version represents the degree of closeness of the fuzzy set to the nearest ordinary set. In other words, correlation also provides a measure of information about the distance of a fuzzy set (represented by μ_1) from its nearest ordinary set (represented by μ_2) and is expressed as

$$C(\mu_1, \mu_2) = 1 - \frac{4}{X_1 + X_2} \sum_{i=1}^{n} \{\mu_1(i) - \mu_2(i)\}^2 \tag{9}$$

with

$$X_1 = \sum_{i=1}^{n} \{2\mu_1(i) - 1\}^2 \tag{10}$$

$$X_2 = \sum_{i=1}^{n} \{2\mu_2(i) - 1\}^2, \tag{11}$$

$0 \leq C(\mu_1, \mu_2) \leq 1.$

3 Optimum Architecture Evolution Using GAs

To use a Hopfield type neural network for object background classification, a neuron is assigned corresponding to every pixel. Each neuron can be connected to its neighbors (over a window) only. The connection can be full or can be partial. In a third order connectivity, the maximum number of connections of each neuron with its neighbors is 8. In practice, all these connections may not exist. The status updating rules are similar to those of Hopfield's model. The energy function of this model consists of two parts. In terms of images, the first part can be viewed as the impact of gray levels of the neighboring pixels and the second part can be attributed to the gray value of the pixel under consideration. The expression of energy for this problem takes the form : $E = -\sum_i \sum_j W_{ij} V_i V_j - \sum_i I_i V_i$, where V_i, V_j are the status of ith and jth neurons, respectively, W_{ij} represents the connection strength between these two neurons and I_i is the initial input bias (which is taken to be proportional to the actual gray level for the corresponding pixel). From a given initial state, the status of a neuron is modified iteratively to attain a stable state. The stable states of the network are made to correspond to the partitioning of a scene into compact regions.

In the present work, each chromosome of the GA represents a network architecture. For an $m \times n$ image, each pixel (neuron) being connected to at most k of its neighbors, the length of the chromosome is $m \times n \times k$ bits. If a neuron is connected to any of its neighbors, the corresponding bit of the chromosome is set to 1, else 0. The initial population is generated randomly. Each network is then allowed to run for object extraction as described in [4]. Each pixel of the extracted output image is considered as an element of a fuzzy set. The fuzziness measure of this set is taken as an index of fitness of the corresponding chromosome (minimum value corresponds to maximum fitness, except the case of using fuzzy correlation) for its selection for the next generation. Crossover and mutation operations are performed on these selected chromosomes to get new offspring (architectures). The whole process is continued for a number of generations until the GA converges. The best chromosome of the final population is considered as the optimum architecture for object extraction.

4 Implementation and Results

The effectiveness of the proposed technique has been demonstrated using some synthetic images of different types which are generated by adding N(0, σ_i^2) noise (σ =10, 20, 32) to each pixel of the binary (two-tone) image. The size of each image is 40×40.

A chromosome is represented as a binary string of length $40 \times 40 \times 8$ (we have taken 8 neighbors for a pixel). The population size is kept fixed at 30. Generational replacement technique and linear normalization selection procedure [10] are adopted. Number of copies produced by the ith chromosome with linear normalized fitness value f_i in a population of size n is taken as round(c_i); where

Table 1. *pcc* for the input images

Image with	*pcc* with fitness evaluation based on											
	Energy		fuzziness measure									
			$f1$		$f2$		$f3$		$f4$		$f5$	
	NE	E	NE	E	NE	E	NE	E	NE	E	NE	E
$\sigma = 10$	99.13	99.25	98.81	98.94	99.44	99.44	99.06	99.44	98.81	98.94	98.81	98.94
$\sigma = 20$	97.88	96.56	96.94	96.88	96.94	96.19	96.81	96.63	96.81	96.62	96.38	96.63
$\sigma = 32$	93.44	92.75	90.06	90.69	89.94	92.25	90.19	88.31	89.44	88.19	90.25	87.69

Table 2. Total number of connections (*noc*) of the evolved networks

Image with	*noc* with fitness evaluation based on											
	Energy		fuzziness measure									
			$f1$		$f2$		$f3$		$f4$		$f5$	
	NE	E	NE	E	NE	E	NE	E	NE	E	NE	E
$\sigma = 10$	9155	8910	7351	7848	7433	7441	7412	7480	7360	7477	7436	7795
$\sigma = 20$	8871	8592	7384	7804	7495	7738	7344	7754	7392	7722	7372	7837
$\sigma = 32$	8387	8488	7425	7716	7487	7903	7452	7834	7424	7943	7460	8049

$c_i = \frac{n \times f_i}{\sum_{i=1}^{n} f_i}$. Both the elitist model (E) and the standard GA (non-elitist model, NE) are implemented. The crossover and mutation probabilities are taken as 0.8 and 0.002, respectively. The algorithm has been run for 200 iterations. It is seen that the average value of the percentage of correct classification of pixels (*pcc*) over the population does not change much after this. The experiment is performed on a SPARC Classic Workstation (frequency 59 MHz). To execute each generation, CPU time requires 2 minutes (approximately).

The percentage of correct classification of pixels of the best chromosome in the last generation using energy and fuzziness measures based fitness evaluations for different noisy versions of the synthetic image is depicted in Table 1. The number of connections (*noc*) of these evolved architectures is put in Table 2. It is seen from Table 1 that the energy based fitness measure performs better than the corresponding fuzziness based ones when the inputs are highly corrupted by noise, and in other cases the results are comparable. On the other hand, it is found from Table 2 that for all the cases, fuzziness measure based evaluations evolve networks with less connectivity (which implies less cost) than those evolved by energy based one.

5 Conclusion

An investigation is carried out to design Hopfield type networks for object extraction using fuzziness measure as fitness criterion. It is found that when the inputs are less corrupted this technique has been able to evolve networks whose connectivity is less compared to the corresponding energy based one in order to produce comparable segmented output. Hence the proposed technique evolves better networks than those evolved using energy based evaluation.

References

1. Mitchell, M.: An Introduction to Genetic Algorithms. The MIT Press, Cambridge, Massachusetts (1996)
2. Whitley, D., Starkweather, T., Bogart, C.: Genetic algorithms and neural networks: Optimizing connections and connectivity. Parallel Computing. **14** (1990) 347–361
3. Pal, S. K., Bhandari, D.:Selection of optimal set of weights in a layered network using genetic algorithms. Information Sciences. **80** (1994) 213–234
4. De, S., Ghosh, A., Pal, S. K.:An application of genetic algorithms to evolve Hopfield type optimum network architectures for object extraction. Proceedings 1995 IEEE International Conference on Evolutionary Computation, ICEC'95, Perth, Western Australia. 1 (1995) 504–508
5. Hopfield, J. J.: Neural network and physical systems with emergent collective computational abilities. Proceedings National Academy of Science, USA. **79** (1982) 2554–2558
6. Ghosh, A., Pal, N. R., Pal, S. K.: Object background classification using Hopfield type neural network. International Journal of Pattern recognition and Artificial Intelligence. **6** (1992) 989–1008
7. Pal, S. K., De, S., Ghosh, A.: Designing Hopfield type networks using genetic algorithms and its comparison with simulated annealing. International Journal of Pattern Recognition and Artificial Intelligence. 11 (1997) 447–461
8. Ghosh, A.: Use of fuzziness measures in layered networks for object extraction : a generalization. Fuzzy Sets and Systems. **72** (1995) 331–348
9. Klir, G. J., Folger, T.: Fuzzy Sets, Uncertainty and Information. Prentice-Hall, Englewood, Cliffs (1988)
10. Davis, L. (ed.): Handbook of Genetic Algorithms. Van Nostrand Reinhold, New York (1991)