

Breast Cancer Diagnostic System using Symbiotic Adaptive Neuro-evolution (SANE)

R.R.Janghel¹, Anupam Shukla², Ritu Tiwari³, Rahul Kala⁴
ABV-IIITM, Gwalior, India
rrj.iiitm@gmail.com¹, dranupamshukla@gmail.com², rt_twr@yahoo.co.in³,
rahulkalaiiitm@yahoo.co.in⁴

Citation: R. R. Janghel, A. Shukla, R. Tiwari, R. Kala (2010) Breast Cancer Diagnostic System using Symbiotic Adaptive Neuro-evolution (SANE). *Proceedings of the 2010 IEEE International Conference of Soft Computing and Pattern Recognition*, Cergy Pontoise/Paris, France, pp 326-329.

Final Version Available At: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5686161

© 2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Abstract: Breast cancer is the second leading cause of cancer deaths in women worldwide and occurs in nearly one out of eight women. In this paper we develop a hybrid intelligent system for diagnosis, prognosis and prediction for breast cancer using SANE (Symbiotic, Adaptive Neuro-evolution) and compare with ensemble ANN, modular neural network, fixed architecture evolutionary neural network (F-ENN) and Variable Architecture evolutionary neural network (V-ENN). While the monolithic neural and fuzzy systems have been extensively used for diagnosis, the individual limitations of the various models put a great threshold on prediction accuracies, which may be overcome with the use of SANE. The SANE system coevolves a population of neurons that cooperate to form a functioning neural network. Breast cancer database from the University of Wisconsin available at UCI Machine Learning Repository is used for conducting experimental work.

Key words: Cancer, SANE (Symbiotic, Adaptive Neuro-evolution), ensemble, modular neural network, fixed architecture evolutionary neural network, variable architecture evolutionary neural network.

1. Introduction

Breast cancer is the second leading cause of cancer deaths in women worldwide and occurs in nearly one out of eight women. It is known that the best prevention method is the precocious diagnosis,

which lessens mortality and improves the treatment. Mammography, Biopsy and Fine Needle aspiration are three commonly used techniques for detection and diagnosis of breast cancer, but only one method applied in one system [1]. We develop a breast cancer diagnostic system using soft computing techniques which can be used for all type of the diagnosis and prognosis in one expert system.

The monolithic neural and fuzzy systems have been extensively researched in literature and numerous comparisons have been done between their various models. The application of these systems effectively solves problems, but their weakness result in sub optimality that needs to be addressed. The problems especially become important when the training data size is very large or the system is very complex. In most of these scenarios, the weaknesses of these systems enlarge and hence the problem cannot be optimally solved by using one system alone. Hence we make use of a combination of systems with an aim to combine the positive aspects of the individual systems and removing their negative aspects. The negative aspects of one system are removed by the positive aspect of the other system. The Hybrid approaches form a very exciting field of work and research. In these systems we combine Neural Networks, Evolutionary Algorithms, Fuzzy Logic and heuristics in numerous ways to make a much more efficient system [2]. A good collection of methods and applications can be found in the books [3, 4]. Evolutionary algorithms provide a general training tool in which few assumptions about the domain are

necessary. Since evolutionary algorithms only require a single fitness evaluation over the entire (possibly multi-step) task, they are able to learn in domains with very sparse reinforcement, which makes them particularly well-suited for evaluating performance in sequential decision tasks. No examples of correct behavior are necessary. The evolutionary algorithm searches for the most productive decision strategies using only the infrequent rewards returned by the underlying system. Together evolutionary algorithms and neural networks offer a promising approach for learning and applying effective decision strategies in many different situations.

Evolution of problems with highly dimensional fitness landscapes is extremely difficult task, which is largely due to the expanse of the search domain. This further highlighted by the presence of large modalities into the search domain, which makes it impossible to timely locate the global minima and converge towards it. CO-evolutionary algorithms are great problem solving tools in such scenarios, where the complete problem of optimization of a highly dimensional fitness landscape may be broken down into multiple sub-problems that together constitute the main problem. The sub-problems are simpler to solve or optimize, and hence aid in providing effective components for the solution of the main problem. It is however important for the different sub-problems to co-operate with each

other, in order to enable effective evolution that ultimately results in locating global minima.

As summarized by Yao [5] the design of artificial neural networks (ANNs) using evolutionary techniques can be roughly classified into three levels: evolving connection weights; evolving architectures and evolving learning rules. In the past decade, more and more approaches have focused on simultaneously evolving connection weights and architectures of ANNs. Both standard evolutionary techniques and co-evolutionary techniques have been introduced into ANNs design.

A novel neuro-evolution mechanism called SANE (Symbiotic, Adaptive Neuro-evolution) is presented that is specifically designed for efficient sequential decision learning. Unlike most approaches, which operate on a population of neural networks, SANE applies genetic operators to a population of neurons. Each neuron's task involves establishing connections with other neurons in the population to form a functioning neural network. Since no one neuron can perform well alone, they must specialize or optimize one aspect of the neural network and connect with other neurons that optimize other aspects. SANE thus

decomposes the search space, which creates a much more efficient genetic search. Moreover, because of the inherent diversity in the neuron population, SANE can quickly revise its decision policy in response to shifts in the domain.

In this paper, we first use the SANE algorithm to effectively solve the problem of Breast Cancer. In this problem SANE is a representative of the class of co-evolutionary neural networks. We then compare the performance of SANE to a number of hybrid algorithms, namely neural network ensembles, modular neural network (MNN), fixed architecture evolutionary neural network (F-ENN) and variable architecture evolutionary neural network (V-ENN).

2. Literature Review

A number of methods have been adopted to solve the problem of breast cancer diagnosis. These methods revolve around development of effective systems for extraction of useful features that result in good diagnosis, and development of effective machine learning systems for recognition. Yuanjiao et al. proposed a technology to extract micro-calcifications clusters with accurate edge effects to obtain much more hidden information which can't be detected by the naked eye on mammograms in order to help the doctors in diagnosing early breast cancer [6]. Computerized microcalcification detection based on fuzzy logic, vibro-acoustography and probabilistic neural network on mammograms for breast cancer diagnosis has been carried out by Heng-Da et al. [7]. Image feature extraction was utilized to retrospectively analyze screening mammograms taken prior to the detection of a malignant mass for early detection of breast cancer in [8]. Statistical texture features for breast cancer detection using Support Vector Machine (SVM) and other machine learning methods like LDA, NDA, PCA, and ANN was done in [9]. SVM was able to achieve better classification accuracy. Early detection of breast cancer is the key to improve survival rate. Thermogram is a promising front-line screening tool as it is able to warn women of cancer up to 10 years in advance [10].

Evolutionary Computing techniques are search algorithms based on the mechanisms of natural selection and genetics. That is, they apply Darwin's principle of the survival of the fittest [11] among computational structures with the stochastic processes of gene mutation, recombination, etc. Central to all evolutionary computing techniques is the idea of searching a problem space by evolving an initially random population of solutions such that better -

"fitter" - solutions are generated over time. These techniques have been applied to a wide variety of domains such as optimization, design, classification, control, economics, biological modelling, and many others. A full review of evolutionary computation is beyond the scope of this article, but recent introductions can be found in [12-15].

Miller et al. [16] proposed that evolutionary computation is a very good candidate to be used to search the space of topologies because the fitness function associated with that space is complex, noisy, nondifferentiable, multimodal, and deceptive. Almost all the current models try to develop a global architecture, which is a very complex problem. Although, some attempts have been made in developing modular networks [17], [18], in most cases the modules are combined only after the evolutionary process has finished and not following a cooperative coevolutionary model. Smalz and Conrad [19] developed a cooperative model that had some similarities with COVNET. In this model there are two populations: a population of nodes, divided into clusters and a population of networks that are combinations of neurons, one from each cluster. Both populations are evolved separately.

Whitehead and Choate [20] developed a cooperative-competitive genetic model for radial-basis function (RBF) neural networks. In this work there is a population of genetically encoded neurons that evolves both the centers and the widths of the RBFs. There is just one network that is formed by the whole population of RBFs. The major problem, as in our approach, is to assign the fitness to each node of the population, as the only performance measure available is for the whole network. Cho and Shimohara [21] developed a modular neural network evolved by means of genetic programming. Each network is a complex structure formed by different modules which are codified by a tree structure.

Based on the pioneering work of Hillis et al [22-23] competitively coevolved the weights (evolutionary learning) of ANNs of fixed topology for a synthetic classification task. The population consists of a fixed number of training patterns. Potter et al [24] proposed Cooperative Coevolutionary Algorithms (CCGAs) for the evolution of ANNs of cascade network topology. Barbosa et al [25] employs a competitive coevolutionary GA for structural optimization problems using a game theoretic approach.

3. Methodology

Symbiotic Adaptive Neuro-Evolution (SANE) [26] evolves three layer networks, where the number of hidden nodes of the network is predefined and fixed, but the network connection topology is not. There are two populations evolved in SANE, a population of neurons and a population of network blueprints. Each individual in the neuron population represents the connection paths and weights of a hidden neuron from the input layer and to the output layer. Each gene of an individual contains two parts: one part specifies the connection path (that is which neuron to connect to) and the other specifies the weight of that connection. All the hidden neurons have the same number of connections, but could have different connection paths from the input layer and to the output layer. Networks are constructed by combining selected individuals from the neuron population. The information of combination is saved in the blueprint population. Each individual of the blueprint population represents a combination of selected individuals from the neuron population. At the beginning of blueprint evolution, combinations are created randomly. Effective combinations can be maintained and new combination forms can be explored by evolving the blueprint population. A well-contributing neuron does not always cooperate well with any other neurons. Therefore, through maintaining a blueprint population, well-contributing neurons are protected from being eliminated due to ineffective cooperation with some other neurons. SANE is an intra-population neurocoevolutionary algorithm that is all the cooperative neurons come from the same population. Each individual of neuron population represents a partial solution instead of a complete network. A complete network is formed by a collection of cooperative neuron individuals. The fitness of a neuron individual is not evaluated independently of other individuals in the neuron populations, but is based on its cooperation. After evaluating all the networks built by blueprints, besides assigning the fitness to each blueprint, each neuron also obtains a fitness value that equals the fitness sum of the best five networks in which the neuron participates. Cooperation only happens when we evaluate individuals, while two populations perform recombination and mutation process independently. The basic steps in the evaluation phase are as follows

Algorithm: The basic steps in one generation of SANE.

1. Clear the fitness value of each neuron and blueprint.
2. Select ζ neurons from the population using a blueprint.
3. Create a neural network from the selected neurons.
4. Evaluate the network in the given task.
5. Assign the blueprint the evaluation of the network as its fitness.
6. Repeat steps 2 to 4 for each individual in the blueprint population.

7. Assign each neuron the evaluation of the best five networks in which it participated.
8. Perform crossover and mutation operations on both populations.

During the evaluation stage, each blueprint is used to select neuron subpopulations of size ζ to form a neural network. Each blueprint receives the fitness evaluation of the resulting network, and each neuron receives the summed fitness evaluations of the best five networks in which it participated. Calculating fitness from the best five networks, as opposed to all the neuron's networks, discourages selection against neurons that are crucial in the best networks but ineffective in poor networks.

4. Simulation Results

The paper proposes the use of SANE for the diagnosis of Breast cancer. The Breast Cancer database is taken from the UCI Machine Learning Repository. Here the system is required to diagnosis the type of breast cancer. It is hence a classificatory problem with two output classes as Malignant and Benign. This database consists of 29 real valued inputs. These correspond to the following features for each cell nucleus: radius (mean of distances from center to points on the perimeter), texture (standard deviation of gray-scale values), perimeter, area, smoothness (local variation in radius lengths), compactness (perimeter² / area - 1.0), concavity (severity of concave portions of the contour), concave points (number of concave portions of the contour), symmetry, fractal dimension ("coastline approximation" - 1). The entire data set consists of a total of 357 benign and 212 malignant cases, totaling to 569 instances in the database.

In this section we experimentally see the behavior of each of the discussed systems for solving the problem. We compare the various methods with respect to their ability to train, learn and generalize the data. The ultimate aim is to have a larger generalizing capability that would mean an effective detection of the disease in the final system. In the next sub-sections we present and compare the results of the various algorithms and hybrid systems that we use.

The SANE algorithm is implemented for the Breast Cancer database. The entire expert system of diagnosis is written in JAVA. The software uses a data module, which consists of the Breast Cancer data. The data obtained from UCI library as text file is loaded into this module at start. The data is divided into training and testing data sets. The SALE algorithm is used for the training of the data, which is later tested for the testing data set.

The various parameters used for the execution of the algorithm include maximum number of connections per neuron as 24, evolutionary population size of 1000, maximum neurons in hidden layer as 12, SANE elite value of 200, mutation rate of 0.2, and number of generations as 100.

In order to analyze the ability of the system to quickly learn from the training data set and generalize the results over the testing data set, we experiment different percentage breakups of training and testing data sets. The results in terms of diagnostic accuracy for different breakups of training and testing data sets is shown in table 1. It may be easily seen that the best results have a training accuracy of 97.88% and a testing accuracy of 97.54.

The effectiveness of the SANE algorithm in solving the problem is studied by comparing it with other approaches taken from literature. These include ensemble neural networks, modular neural networks (MNN), fixed architecture evolutionary neural network (F-ANN) and variable architecture evolutionary neural network (V-ANN). MATLAB was used as an implementation platform for all these.

The first method against which the performance of the SANE algorithm was tested was multi layer perceptron neural network (MLP) trained with the help of Back Propagation Algorithm. The testing was done using MATLAB as the developmental platform. The network had 1 hidden layer with 18 neurons. Learning rate was fixed to be 0.05. Momentum was fixed to 0.7. The network was trained for 3500 epochs. The resultant network gave a training accuracy of 97.01% and a testing accuracy of 94.61%.

Next we compare SANE with ensemble of neural networks. This makes use of multiple neural networks to solve the same problem. The results of all neural networks are later integrated by an integrator. Here we use 4 neural networks. All 4 ANNs were first trained independently one after the other. Then all these ANNs were used to make an integrator which gave the inputs to each of these ANNs, collected their outputs and then gave the final result of the system. The 4 ANNs were more or less similar with small changes. The resulting system had a total performance of 97.98% on the training data and 95.95% on the testing data.

Next we study the use of Modular Neural Networks for the same problem. In this approach we tried to divide the entire input space into a set of clusters. Each cluster served as a separate problem domain, where the task of diagnosis needs to be carried out. Each cluster is trained and tested by a separate neural network, which forms one of the modules of the complete MNN. Here each neural network used a single hidden layer. The ANN structure used for this problem was [30 x 18 x 1]. After training and testing, the system gave an accuracy of 96.49% on the training data set and 95.08% on the testing data set.

The other approach used to compare with SANE was Fixed Architecture Evolutionary Neural Network. Here we made use of Genetic Algorithm for the task of training of the neural network. The training involved the setting of weights and biases, so as to maximize the performance on training data. F-ENN used a single hidden layer ANN for this problem. The ANN structure used for this problem was [30 x 18 x 1]. We then applied

GA for the parameter optimization. The weight matrix consisted of 30*18 weights between input and hidden layer, 18*1 between the hidden and the output layer and a total of 18 hidden layer biases and 1 output layer bias. This made the total number of variables for the GA as 577. The total number of individuals in the population was 10. A uniform creation function, rank based scaling function and stochastic uniform selection methods were used. The elite count was 2. Single point crossover was used. The program was executed till 15 generations. The crossover rate was 0.8. The genetic algorithm used back propagation algorithm as the local search strategy. The BPA had a learning rate of 0.05, and momentum of 1. Training was carried for 30 epochs. The algorithm had a training accuracy of 93.92% and testing accuracy of 95.40%.

The last method of interest was the variable architecture evolutionary neural network. Here we tried to optimize the neural network architecture as well as the weights and biases with the help of Genetic Algorithm. We follow a connectionist approach. The neural network is assumed to be consisting of one hidden layer. In place of an all-connected architecture, we assume that only some connections are allowed from the input layer to hidden layer and hidden layer to output layer. The information regarding the connections is stored into the genetic individual. The parameters of the GA were a maximum number of 30 neurons, 10 as the population size with an elite count of 2. The creation function was uniform and double vector representation was chosen. Rank based fitness selection was used. Stochastic Uniform selection method was used. Crossover ratio was 0.8. The algorithm was run for 15 generations. Extra connections were penalized by assigning a penalty of 0.01 per connection. The accuracy in this case was 97.01% for the training data and 95.21% for testing data.

Table 1. Performance of SANE.

Database Training/ Testing	Generation	Training accuracy	Testing accuracy
10 / 90	999	98.31933	95.46828
30 / 70	999	97.54386	97.88733
40 / 60	999	94.405594	93.28622
50 / 50	999	98.81657	94.3723
60 / 40	999	96.81159	95.98214
70 / 30	999	97.74436	96.47059
90 / 10	999	96.33028	96.99248

Table 2. Comparative Analysis of SANE.

S. No.	Methods	Training accuracy %	Testing accuracy %
1.	BPA	97.01%	94.61%
2.	Ensemble	97.98%	95.95%
3.	MNN	96.49%	95.08%
4.	F-ENN	93.92%	95.40%
5.	V-ENN	97.01%	95.21%
6.	SANE	97.54%	97.88%

The comparative analysis of SANE with the other approaches is given in table 2. It may be easily seen that the performance of SANE is the highest as compared to the other approaches for the testing data. This naturally proves a good generalizing capability of this co-evolutionary neural network model.

5. Conclusion

In this paper we presented the application of SANE algorithm to the problem of diagnosis of breast cancer. Attainment of high diagnostic accuracy is a natural demand in any expert system that is applied into the bio-medical domain. This largely results in recording of a large number of parameters that ultimately lead to an excessive increase in the problem complexity. The complex problems are very difficult for the neural network to model and get effectively trained. The possibility of sub-optimality due to the limitations of the human model designer or the training algorithm are further responsible for limitation in use of neural networks that further necessitate the use of neural network ensembles or the evolutionary neural networks for solving the problem. Each of these in their own modeling mechanisms result in possibility of better systems for medical diagnosis. The effectiveness is measured by the diagnostic accuracy in the testing data that is completely unknown to the expert system. The immensely large complexity of the search space for any genetic algorithm that tries to evolve the neural network, further results in inefficient exploration of the fitness landscape and hence convergence to a local minima. This further calls for the use co-evolutionary class of neural networks for effectively solving the problem.

In this paper we first attempted to solve the problem of breast cancer diagnosis using SANE algorithm. The algorithm gave a high degree of diagnosis that could promise an effective diagnosis system. It is however necessary to study the effectiveness in comparisons to the other modular, evolutionary and ensemble systems. Experimental analysis validate the supremacy of the co-evolutionary model SANE over the other models.

Further experimentation may be carried out in the use of other datasets of Breast Cancer, use of different co-evolutionary models, and use of these models in conjunction with effective feature selection techniques. The large training time is a limitation in the use of evolutionary systems, which

further gets increased in the presence of massive sized datasets. Time-effective network training may hence be another point of focus.

Acknowledgements

The authors sincerely acknowledge the Director ABV-IIITM, Gwalior, India for providing facilities to carry out this research work.

6. References

- [1.] <http://www.breastcancer.org>.
- [2.] Srinivasaa, K.G., Venugopala, K.R. & Patnaikb, L.M. (2007). A self-adaptive migration model genetic algorithm for data mining applications, *Information Sciences*, 177(20) 4295-4313
- [3.] Melin, P & Castillo, O (2005). *Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing*, Springer
- [4.] Bunke, H & Kandel A (Eds)(2002). *Hybrid Methods in Pattern Recognition*, World Scientific.
- [5.] Yao, X., *Evolving Artificial Neural Networks*, Proceedings of the IEEE, vol. 87, pp. 1423-1477, 1999.
- [6.] Yuanjiao MA, Ziwu WANG, Jeffrey Lian LU, Gang WANG, Peng LI Tianxin MA, Yinfu XIE ,Zhijie ZHENG “Extracting Micro-calcification Clusters on Mammograms for Early Breast Cancer Detection” Proceedings of the 2006 IEEE International Conference on Information Acquisition August 20 - 23, 2006, Weihai, Shandong, China .pp499-504.
- [7.] Heng-Da Cheng, Yui Man Lui, and Rita I. Freimanis “A Novel Approach to Microcalcification Detection Using Fuzzy Logic Technique” *IEEE transactions on medical imaging*, vol. 17, no. 3, June 1998, pp442-450.
- [8.] Mohammad Sameti, Rabab Kreidieh Ward, Jacqueline Morgan-Parkes and Branko Palcic “Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer” *IEEE journal of selected topics in signal processing*, vol. 3, no. 1, February 2009, pp 46-52.
- [9.] Al Mutaz M, Abdalla , Safaai Deris, Nazar Zaki and Doaa M. Ghoneim “ Breast Cancer Detection Based on Statistical Textural Features Classification“ 2008 IEEE , pp 728-730.
- [10.] T.Z. Tan , C. Quek , G.S. Ng a, E.Y.K. Ng “A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure” *Expert Systems with Applications* 33 (2007) 652–666.
- [11.] Darwin C R (1859)(ed.), *The Origin of Species*, John Murray.
- [12.] Koza J R (1992)(ed.), *Genetic Programming*, MIT Press.
- [13.] Fogel D B (1995)(ed.) *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, IEEE.
- [14.] Mitchell M, Hraber P, & Crutchfield J (1993), "Revisiting the Edge of Chaos: Evolving Cellular Automata to Perform Computations", *Complex Systems* 7:89-130.
- [15.] Baeck T (1996)(ed.), *Evolutionary Algorithms in Theory and Practice*, Oxford University Press
- [16.] G. F. Miller, P.M. Todd, and S. U. Hedge, “Designing neural networks,” *Neural Networks*, vol. 4, pp. 53–60, 1991.
- [17.] Y. Liu and X. Yao, “Ensemble learning via negative correlation,” *Neural Networks*, vol. 12, no. 10, pp. 1399–1404, Dec. 1999.
- [18.] B. E. Rosen, “Ensemble learning using decorrelated neural networks,” *Connection Sci.*, vol. 8, no. 3, pp. 373–384, Dec. 1996.
- [19.] R. Smalz and M. Conrad, “Combining evolution with credit apportionment: A new learning algorithm for neural nets,” *Neural Networks*, vol. 7, no. 2, pp. 341–351, 1994.
- [20.] B. A. Whitehead and T. D. Choate, “Cooperative-competitive genetic evolution of radial basis function centers and widths for time series prediction,” *IEEE Trans. Neural Networks*, vol. 7, July 1996.
- [21.] S.-B. Cho and K. Shimohara, “Evolutionary learning of modular neural networks with genetic programming,” *Appl. Intell.*, vol. 9, pp. 191–200, 1998.
- [22.] W. D. Hillis. Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D*, 42:228 {234, 1990.
- [23.] Jan Paredis. Steps towards Co evolutionary Classification Neural Networks. In R. Brooks and P. Maes, editors, *Proceedings Artificial Life IV*, pages 545{552. MIT Press / Bradford Books, 1994.
- [24.] Mitchell A. Potter and Kenneth A. De Jong. *Evolving Neural Networks with Collaborative Species*. In *Proceedings of the 1995 Summer Computer Simulation Conference*, 1995.
- [25.] Helio J. C. Barbosa. A coevolutionary genetic algorithm for a game approach to structural optimization. In Thomas Back, editor, *Proceedings of the Seventh International Conference on Genetic Algorithms*, pages 545-552, San Francisco, California, 1997. Morgan Kaufmann Publishers, Inc.
- [26.] Moriarty, D. E. and Miikkulainen, R., *Forming Neural Networks through Efficient and Adaptive Coevolution*, *Evolutionary Computation*, vol. 5, pp. 373-399, 1997.