

Evolving Digital Circuits for Early Stage Breast Cancer Detection Using Cartesian Genetic Programming

Zahra Khalid, Gul Muhammad Khan, Arbab Masood Ahmad

Abstract—Cartesian Genetic Programming (CGP) is explored to design an optimal circuit capable of early stage breast cancer detection. CGP is used to evolve simple multiplexer circuits for detection of malignancy in the Fine Needle Aspiration (FNA) samples of breast. The data set used is extracted from Wisconsin Breast Cancer Database (WBCD). A range of experiments were performed, each with different set of network parameters. The best evolved network detected malignancy with an accuracy of 99.14%, which is higher than that produced with most of the contemporary non-linear techniques that are computationally expensive than the proposed system. The evolved network comprises of simple multiplexers and can be implemented easily in hardware without any further complications or inaccuracy, being the digital circuit.

Keywords—Breast cancer detection, cartesian genetic programming, evolvable hardware, fine needle aspiration(FNA).

I. INTRODUCTION

BREAST Cancer has always remained the center of attention for active research. Cancer is a malignant tumor, an uncontrolled growth of cells that invade the surrounding tissues at early stages and spread to other areas of the body hence moving to complex levels. This work explore a computational technique to design an optimum circuit for early stage Breast cancer detection.

In the last century, massive development has been made in medical science and new treatment, remedial procedures have been introduced. For these to be effective, it is necessary to develop systems for reliable diagnostic decisions. Breast Cancer is the disease which accounts for a considerable percentage of mortality rate, of all cancer deaths, including both genders, but mostly women. Timely detection increases the chances of survival. There are several confounding factors which lead to late diagnosis of the disease. Conventionally accepted dogma for these delays is that most of the time people hesitate from undergoing painful diagnostic procedures. Failure to diagnose the disease is the second most common reason for diagnostic delays. For a pathologist, it is a routine and critically very important task to identify the presence or absence of cancer cells in patient's samples. The fatigue and low expertise of a pathologist both can lead to wrong diagnosis. There is no doubt that diagnosis of breast cancer is a challenging and a difficult job. At the same time, it is extremely laborious to identify few malignant cells among millions of normal cells, through a microscope. The work done in this research is to assist pathologists in making more

accurate decisions. Fine Needle Aspiration (FNA) data was taken from Wisconsin Breast Cancer Database (WBCD) [1]. Cartesian Genetic Programming is used to classify the data. The CGP network is first trained with a large number of FNA samples and then tested with equal number of test samples, to assess system's performance.

A. Evolvable Hardware

Evolvable Hardware is an area, where emphasis is placed on the utilization of evolutionary techniques to make specialized electronic circuits without needing traditional applied engineering. It is an amalgamation of reconfigurable hardware, artificial intelligence, problem tolerance and autonomous systems. It utilizes application of Evolutionary and Biologically Inspired Algorithms for the specific motive of creating novel designs of optimized physical circuits and systems. Simulators and reconfigurable hardware ensures the accuracy with respect to the final hardware designs, which might include device simulators [23], or actual devices [24]. Two main methods used to evolve the circuit, in Evolvable Hardware are: Extrinsic and Intrinsic evolution.

1) *Extrinsic Evolution*: Extrinsic Evolution is characterized by the assessment of electronic circuits through simulation instead real-time construction during training and testing phases, and later on the final design is evaluated on real hardware. Success of this technique manifestly depends upon two important components of the system, the simulation software program used and the kind of digital additives within the final hardware implementation. Thus when implemented and evaluated in hardware, it cannot be ensured that the extrinsically evolved circuit would work as predicted. The fundamental problem with extrinsic evolvable hardware is that if elementary components like simple AND, OR and NOT gates are used, simulation of the system becomes easy and time taken to develop the circuit is manageable depending on the computer. But as soon the problem grows to be more complicated and the additives end up more complex, the time required for the simulation increases considerably. After sufficient evolution cycles a satisfactory performing circuit is produced. The final design if digital usually shows exactly same behaviour as its simulation, for analogue systems it can be tricky.

2) *Intrinsic Evolution*: Intrinsic Evolution is characterized by in-circuit evolution of the hardware instead of using simulation models. The circuit's fitness is evaluated at run-time

Zahra Khalid is with the University of Engineering and Technology, Pakistan (e-mail: khalid.xahra@gmail.com).

by comparing its output with target. Intrinsic evolvable hardware does not follow any pre-conditional guidelines which were required to layout an electronic circuit by extrinsic way, instead it uses device features in any way it chooses. It is very much similar to the simulation system (intrinsic Evolution), the final circuit may use characteristics that have wide tolerances, which change from system to system or even across a single system. These systems are also affected by the environmental conditions such as temperature or radiations.

II. RELATED WORK

Several computer-based methods have been applied to the diagnosis of breast malignancy, using the FNA technique. Hong et al. used Genetic Programming (GP) to extract parameters from data with the help of Fischer criterion [3]. This helps in analyzing data and interpreting its statistical values. Feature extraction techniques are used to remove redundancy and retain the useful parameters. Fischer criteria is applicable when the sample size is small. Fisher Linear Discriminant Analysis (FLDA) was thought to be the best method for statistical analysis and feature extraction, but using this method one can predict a limited number of features [4]. Later on Hong et al. proposed Modified Fisher Linear Discriminant Analysis (MFLDA). This method overcomes the limitations of Fisher Criterion. Features extracted using this method perform well for the target. MFLDA in conjunction with GP was used to generate a reduced feature set from the raw data. The reduced dataset was then put on a straightforward classifier called "Minimum Distance Classifier (MDC)". Tests show that (MDC) blend provides best results with increased accuracy as compared to other techniques. Once the classification requests were examined, a variety of methods were used to reach high classification accuracies. In [5] the authors used a hybrid method for diagnosis of Breast Cancer. The technique was based on Fuzzy-artificial and KNN algorithm. The results were 99.14% accurate. Quinlan attained 94.74% classification precision using 10-fold combination validation method with C4.5 decision tree method [6]. Hamilton et al. obtained 96% accuracy with RIAC method [7]. Linear Discriminant Analysis (LDA) resulted in 96.8% accuracy [8]. The result obtained from support vector machines(SVM) (5xCV) method was 97.2% accurate [9]. Neuro-fuzzy techniques produced 95.06% [10]. Pena Sipper and Rayes resulted in 97.36% accuracy using fuzzy-GA method [11]. Neuro-rule produced 98.1% accuracy [12]. Three different methods were applied to the nagging problem by Goodman et al. which resulted in following accuracies: Optimized-LVQ methods performance was 96.7%, Big-LVQ method reached 96.8%, and AIRS which the author proposed with respect to the artificial disease fighting capability, obtained 97.2% classification accuracy and reliability [13]. Nevertheless, a fuzzy clustering (SFC) was supervised by Abony and Szeifert who obtained 95.5% accuracy [14]. Table I shows a comparison between the results of all previous work and those achieved through CGP technique.

III. CARTESIAN GENETIC PROGRAMMING

Cartesian Genetic programming (CGP) is a form of Genetic Programming(GP) where nodes are arranged in Cartesian format. In Cartesian Genetic Programming, solutions are represented as a string of integers (genotype-of fixed length) that is mapped to a directed focused graph (phenotype). CGP can efficiently represent common computational structures. These include mathematical equations, computer programs, neural sites and general digital circuits. CGP encodes a prospect solution (typically a circuit or an application) using a wide range of programmable nodes [21]. Each node has fixed number of inputs commonly known as arity, and a single output. A node can perform only one predefined primitive function and can be linked, either to the outputs of nodes positioned in previous columns or to the program inputs. The primary feature of CGP is that not all nodes are necessarily connected in the final phenotype, thus giving a flexibility in terms of complexity in various phenotypes.

Genotype is usually represented as an array of integers as shown below:

$$\{I_1 I_2 F_0, I_1 I_2 F_1, I_1 I_2 F_2, O_1 O_2 O_3 \}$$

In this case, there are a total of three nodes, each with arity two and three system outputs. In CGP encoding, although the size of genotype is fixed but size of the phenotype varies, since not all nodes are in path from input to output. These unused nodes are known as inactive nodes, hence they are not evaluated during the process. This ultimately makes CGP computationally faster than other genetic algorithms. Let us consider a genotype.

$$\{1 \ 2 \ 3, 3 \ 3 \ 4, 4 \ 2 \ 5, 4 \}$$

Fig. 1 shows Phenotype of the above mentioned genotype. It clearly explains how various genes are linked in CGP. This chromosome shows an acyclic graph made up of three system inputs, three nodes and one result (output). The arity of every node is two. All nodes do not necessarily contribute to the end result of the graph (node 5 is not used in this case). Nodes which do not contribute to the system result are reported to be inactive. Not all of the inputs need to be used when determining the outputs (in this case 0 is not used).

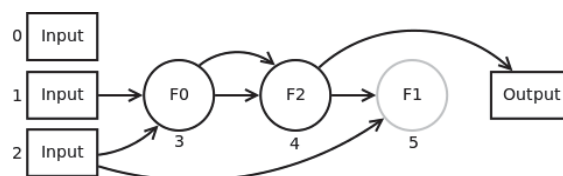


Fig. 1 ·Phenotype

A. Evolution Strategy

$1 + \lambda$ evolution strategy is introduced in this work. Where λ represents the number of offspring in a population. Hence $(1 + \lambda)$ suggests that the evolution strategy have one parent, which is then mutated to produce λ offspring. In all the experiments λ is set to 9. Thus the evolution strategy becomes "1+9". The following steps are performed to execute the whole process:

TABLE I
COMPARISON TABLE FOR PREVIOUS WORK AND RELATED WORK

| S.no | Method | Accuracy percentage | References |
|------|------------------------------------|------------------------------|--------------|
| 1 | EPNet | 97.254 | [15] |
| 2 | MPANN | 98.1 | [16] |
| 3 | BPA | 98.14 | [17] |
| 4 | LDA | 96.8 | - |
| 5 | RIAC | 96 | - |
| 6 | SVM | 97.2 | - |
| 7 | Fuzz-artificial and k-nn algorithm | 99.14 | - |
| 8 | C4.5 decision tree method | 94.74 | - |
| 9 | Neuro-fuzzy techniques | 95.06 | - |
| 10 | fuzzy-GA | 97.36 | - |
| 11 | Neuro-rule | 98.1 | - |
| 12 | SFC | 95.5 | - |
| 13 | NLC with GA | 97.0 | [18] |
| 14 | NegBoost with $\lambda = 0$ | 98.3 | [19] |
| 15 | CGPANN | 99.09 Type 1 99.90 Type 2 | [20] |
| 16 | CGP | 99.14 | Present Work |

- Population of 10 chromosomes is generated randomly and then each of them is evaluated for its fitness.
- The fittest individual is selected as the parent for next generation.
- The parent along with its 9 mutated genotypes (λ i.e 9 offspring) form the new population.
- If fitness of an offspring is greater than or equal to the parent, child is promoted as a parent, otherwise parent remains parent for next generation too.

IV. APPLICATION OF CARTESIAN GENETIC PROGRAMMING FOR BREAST CANCER DIAGNOSIS

This section explores the aspects of Fine Needle Aspiration data set, the working strategy, results and analysis.

A. Diagnostic Procedure and Extraction of Data Set

Fine Needle Aspiration (FNA) of breast is a diagnostic technique which involves the extraction of fluid from the breast lump, with the help of a thin hollow needle. After staining the specimen, it is inspected under a microscope. The case may be declared malignant or benign, based on shape features of the cells. For research purpose there are two popular datasets formed using Fine Needle Aspiration (FNA) methodology. The original Wisconsin Breast Cancer Database that consists of nine parameters [1] and the one that consists of thirty parameters [2]. The two data set are basically different in their detailed diagnostic mechanism, although the FNA technique is used in both strategies. In nine parameter database, the data consists of nine crude parameters of the cell, obtained after microscopic examination. Nine parameters comprises of the following:

- 1) Clump Thickness.
- 2) Uniformity of Cell Size.
- 3) Uniformity of Cell Shape.
- 4) Marginal Adhesion.
- 5) Single Epithelial Cell Size.
- 6) Bare Nuclei.
- 7) Bland Chromatin.
- 8) Normal Nucleoli.
- 9) Mitosis.

The second database consists of cellular features of the sample, which are collected and analyzed on the basis of a digital scan, using Xcyt software. Fluid is excised from the suspected region of breast through a fine needle. In order to highlight the cell nuclei stained slide is examined under microscope. Well differentiated cells are magnified and captured with an electronic camera and frame grabber board. Individual nuclei of the cell are isolated through Xcyt. Across each nucleus an approximate boundary is drawn with a mouse pointer and the precise boundaries of the nuclei are extracted from the computer perspective strategy known as snake. Fig. 2 shows the Xcyt program interface. The dotted lines are the approximate boundaries of the nuclei, initialized by the operator, using mouse pointer. On the other hand Xcyt uses a curve-fitting program to draw the exact boundaries shown by solid lines. This process takes five minutes per slide. The following ten features of a sample are computed through Xcyt:

- 1) Radius.
- 2) Perimeter.
- 3) Area.
- 4) Compactness.
- 5) Smoothness.
- 6) Concavity.
- 7) Concave points.
- 8) Symmetry.
- 9) Fractal Dimension.
- 10) Texture.

A total of thirty real-valued parameters are obtained by calculating the mean value, worst (mean of the three largest values) and standard error of the ten characteristics, mentioned above, for each sample.

B. Experimental Setup

The CGP network is trained on the 350 samples in case of FNA nine parameters data set. The trained network is then tested for same number of samples (other than training samples) of WBCD data set. The following parameters are defined by the user: number of inputs, maximum number of nodes, arity, number of functions and number of system outputs.

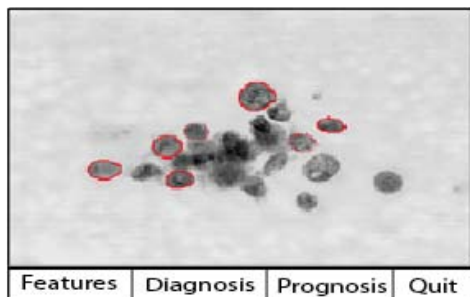


Fig. 2 Xcvt Program Interface

An individual genotype is evaluated for its fitness, by applying the training data to its developed phenotype. The nine parameters of each sample are assigned as input to the system. The input values are first normalized using the following equation

$$I_N = (I - Min) \frac{newMax - newMin}{Max - Min} + newMin \quad (1)$$

The normalized values are rounded off to the nearest number. These round-off numbers are converted into four bit binary. Thereby a total of thirty-six bits are obtained, each of them is applied to the CGP algorithm as an individual input. Thus there are a total of thirty six system inputs.

The effectiveness of CGP is dictated by the numerous parameters, one being the function set utilized. In this case the function set comprises of multiplexers, which are atomic in nature and can implement any logic circuit [22].

The function set used is:

$$\begin{aligned} F0 &= a.c' + b.c, \\ F1 &= a'.c' + b.c, \\ F2 &= a.c' + b'.c, \\ F3 &= a'.c' + b'.c. \end{aligned}$$

After evaluation, eighteen system outputs are obtained, as defined by user. Since single output is required to compare the system output with target value, therefore these 18 outputs are XORed to get an equally distributed single result. If the system output is equal to the target output, the fitness is incremented otherwise it remains the same. The system classifies the input sample as benign if the output (obtained after applying XOR) is 0; and malignant if the output is 1. Fig. 3 shows the way the input parameters are applied to CGP network (genotype); and its outputs XORed to obtain a single output. The genotype with maximum fitness is selected as a parent to create a new population (with the same number of members) through mutation. This process continues until a preset level i.e target fitness is achieved or the generation count reaches its preset limit.

For complex system design, evolvable hardware technique is preferred. The best solution (fittest genotype) obtained through CGP algorithm represents an evolved circuit for breast cancer detection.

Specifications of the best genotype obtained and transformed to phenotype as a logical system, which is shown in Fig. 4. Total thirty-six inputs from zero to thirty-five are provided to system, represented by system inputs label.

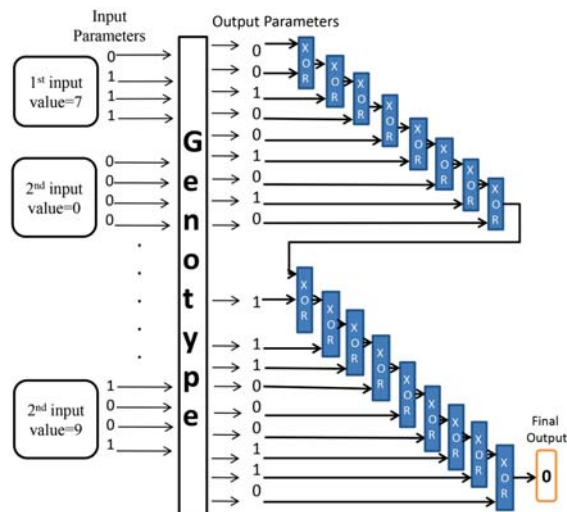


Fig. 3 Evaluation Process for a Single Genotype

In this case there are thirty-eight nodes, number of nodes can be greater or smaller depending on genotype size.

Parameters of a single node i.e. arity, node function and node output, in phenotype are clearly shown in Fig. 5. Node input and node function is randomly selected and then node's output is generated. Node output could be used by the proceeding nodes as their input. The process continues and total eighteen outputs are produced by the system. Four types of node functions are used 0, 1, 2 and 3.

The logic circuit of each function is shown in Fig. 6, where a, b and c represent inputs of a node. Case:0, Case:1, Case:2 and Case:3 represent the four system functions F0, F1, F2, and F3 respectively. The logic gates used in circuitry of each function are And, OR and Inverter. The connectivity of these logic gates is different for each function and is clearly shown in figure.

C. Result and Analysis

Experiments performed on the proposed algorithm are listed in Table II. Three sets of experiments are carried out, each set consisting of five experiments. Following are the common specification in all the three cases.

- Number of inputs to the system = 36
- Number of outputs = 18
- Number of functions = 4
- Input per node (arity) = 3
- Evolution Strategy = 1+9
- Mutation Rate = 10

Different number of nodes are specified for each set of experiments. Experiments in each set differ from other sets in terms of seed used, which is randomly selected. Table II shows the accuracy values of first, second and third set of experiments for which the number of nodes were set to 100, 150 and 200 respectively. Each network is applied to the parameters extracted from all the 350 samples and its fitness determined.

The following performance metrics were determined for the

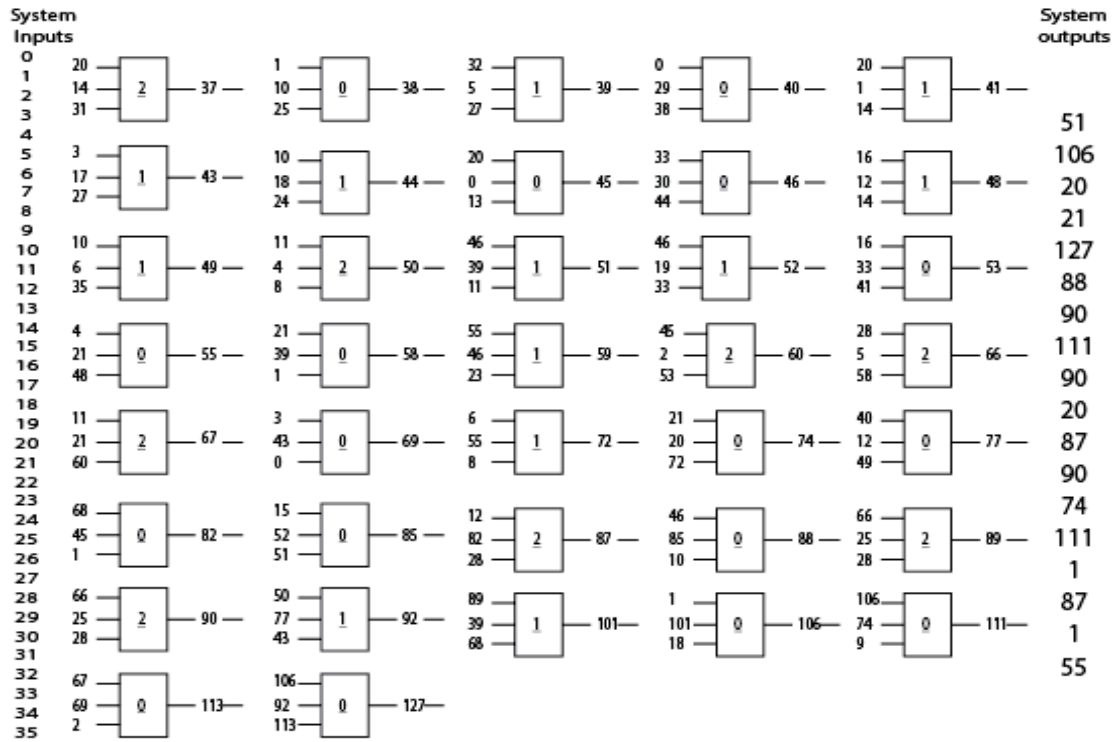


Fig. 4 Phenotype of the circuit obtained through CGP

TABLE II
 RESULTS

| Experiment No. | No. of Nodes | Trainig Fitness | Testing Fitness | Percentage Accuracy | Average Percentage Acc. |
|----------------|--------------|-----------------|-----------------|---------------------|-------------------------|
| 1 | | 338 | 330 | 94.28 | |
| 2 | | 342 | 340 | 97.14 | |
| 3 | 100 | 304 | 273 | 78.00 | 91.89 |
| 4 | | 339 | 325 | 92.28 | |
| 5 | | 342 | 342 | 97.77 | |
| 6 | | 326 | 329 | 94.00 | |
| 7 | | 301 | 295 | 84.28 | |
| 8 | 150 | 301 | 395 | 87.14 | 90.51 |
| 9 | | 326 | 317 | 90.57 | |
| 10 | | 332 | 338 | 96.57 | |
| 11 | | 298 | 304 | 86.86 | |
| 12 | | 282 | 287 | 82.00 | |
| 13 | 200 | 318 | 321 | 91.70 | 89.20 |
| 14 | | 344 | 347 | 99.14 | |
| 15 | | 297 | 302 | 86.28 | |

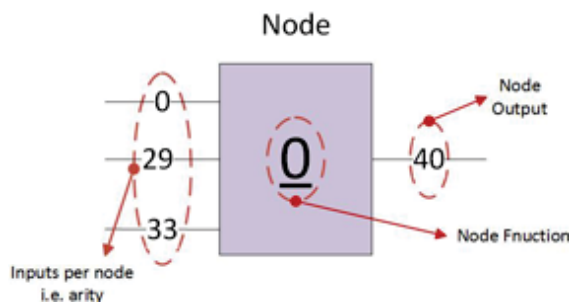


Fig. 5 Parameters of Single Node

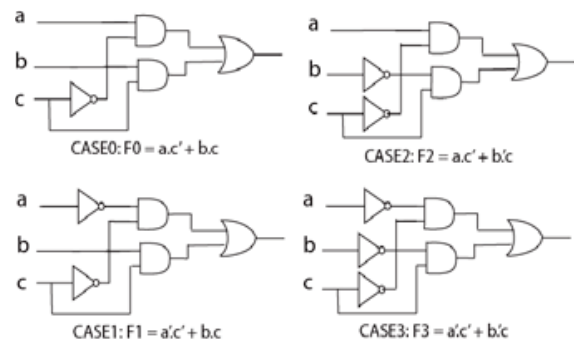


Fig. 6 Parameters of Single Node

best network and shown in Table III:

$$Accuracy = \frac{TP + TN}{N} \quad (2)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

where

True Positive (TP): A=1, T=1.

False Positive (FP): A=1, T=0.

True Negative (TN): A=0, T=1.

False Negative (FN): A=0, T=0.

Total Number of Samples (N)=350

A=Actual result,

T=Target result

and MCC stands for Matthews Correlation Coefficient.

TABLE III
 RESULT OF BEST NETWORK

| Fitness of the Best Network | Genotype |
|-----------------------------|----------|
| Accuracy | 99.14% |
| Sensitivity | 97.59% |
| Specificity | 99.62% |
| Precision | 98.78% |
| MCC | 97.62% |

V. DISCUSSION

To evaluate the effectiveness of the optimally evolved circuit, experiments were performed on WBCD database as mentioned above and results compared with previously used methods. A comparative performance analysis based on percentage accuracy of various classifiers is shown in Table I. The performance measures emphasize that Cartesian Genetic Programming network has greater potential to be used as breast cancer classifier and is generally better in comparison with other classifiers. The results show that the best network is found to be 99.14% accurate and that the CGP approach generates a sensitivity of 97.59% and specificity of 99.62%. Thus classifier introduced in this research work has a greater potential to be applied in automatic diagnosis of breast cancer. The CGP network classifier is superior to the other networks based on the fact that the circuit is made up of simple logic AND, OR and NOR gates, which makes it more suitable for efficiently detecting breast cancer.

VI. CONCLUSION

In this project an extrinsically evolved system for breast cancer detection has been presented. The system consists of a digital logic circuit that was evolved using Cartesian Genetic Programming. The evolution involves presenting training samples to the network, together with the target outputs. These samples consist of cell features calculated from the microscopic images of breast FNA. The data is available at the popular Wisconsin's Breast Cancer Database (WBCD). In each generation of evolution a population of possible networks are produced from the best network of the previous generation, through mutation. The best network is chosen based on how

well the actual and target outputs match. The system developed in this project is the first step towards a complete solution for breast cancer detection. The next step in this work would be to load the design into a Field Programmable Gate Array (FPGA). Advanced version of this project would include cell segmentation and feature extraction from an FNA cytology slide image, before applying those features to the system developed in this project.

ACKNOWLEDGMENT

We are grateful to M. Lichman, University of California, Irvine, School of Information and Computer Sciences for the Machine Learning Repository from which we downloaded the Wisconsin's Breast Cancer Dataset [25].

REFERENCES

- [1] Breast Cancer Wisconsin (Diagnostic) Data Set. [http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)).
- [2] <http://mlr.cs.umass.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>.
- [3] Hong Guo and Asoke K. Nandi. Breast cancer diagnosis using genetic programming generated feature. *Pattern Recognition*, 39(5):980-987, 2006.
- [4] J. Han and N. Cercone. *ule Viz: a model for visualizing knowledge discovery process*. In *KDD*, pages 244-253, 2000.
- [5] Seral Sahan, Kamal polat, Halife Kodaz and Salih Gunes. Hybrid method based on fuzz-immune system and k-nn algorithm for Breast Cancer Diagnosis, 37(2007) 415-423.
- [6] J. R. Quinlan, Improved use of continuous attributes in C4.5, *J. Artif. Intell. Res* 4(1996) 77-90.
- [7] H. J. Hamilton, N. Shan, N.Cerone. RIAC: a rule induction algorithm based on approximate classification, Technical Report CS 96-06, University of regina, 1996.
- [8] B. Ster, A. Dobinka, Neural networks in medical diagnosis: comparison with other methods, in: *Proceedings of International Conference on Engineering Applications of Neural Networks (EANN 96)*, 1996, pp. 427-430.
- [9] K. P. Bennet, J. A. Blue, A support vector machine approach to decision trees, *Math Report*, vols. 97-100, Rensselaer Polytechnic Institute.
- [10] D. Nauck, R. Kruse, obtaining interpretable fuzzy classification rules from medical data *Artif. Intell-Med.* 16(1999) 149-169.
- [11] C. A. Pena-Reyes, M.Sipper, A fuzzy-genetic approach to breast Cancer diagnosis, *Artif. Intell. Med.* 17(1999). 131-155.
- [12] R. Setiono, Generating concise and accurate classification rules for Breast Cancer Diagnosis, *Artif. Intell. Med.*18(2000) 205-219.
- [13] D. E. Goodman, L.Bogges, A. Watkins, Artificial immune System Classification of multipleclass problems, in: *Proceedings of artificial Neural Networks in Engineering ANNIE (2002)*, 2002, pp.179-183
- [14] J. Abonyi, F. Szeifert, Supervised Fuzzy clustering for the identification of fuzzy classifiers, *Pattern Recognition Lett.* 24(2003) 2195-2207.
- [15] W. H. Land Jr, L. Albertelli, Y. Titkov, P. Kaltsatis, and G. Seburyano. Evolution of neural networks for the detection of breast cancer. In *Proc. IEEE. Int. Joint Symposia on Intelligence and Systems, INTSYS '98*, pages 34-, 1998.
- [16] Hussein A. Abbass, An evolutionary artificial neural networks approach for breast cancer diagnosis. *Artificial Intelligence in Medicine*, 25:265-281, 2002.
- [17] R. Janghel, Anupam Shukla, Ritu Tiwari, and Rahul Kala. Intelligent Decision Support System for Breast Cancer. In *Advances in Swarm Intelligence*, volume 6146, chapter 46, pages 51-358. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [18] Hai H. Dam, Hussein Abbass, Chris Lokan, Xin Yao, et al. Neural-based learning classifier Systems. *Knowledge and Data Engineering, IEEE Transactions on*, 20 (1):26-39, 2008.
- [19] Md Monirul Islam, Xin Yao, S. M. Shahriar Nirjon, Muhammad Asiful Islam, and Kazuyuki Murase. Bagging and boosting negatively correlated neural networks. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 38 (3):771-784, 2008
- [20] Arbab Masood Ahmad, Gul Muhammad Khan, and Sahibzada Ali Mahmud. Classification of mammograms using cartesian genetic programming evolved artificial neural networks. In *Artificial Intelligence Applications and Innovations*, pages 203-213. Springer, 2014

- [21] Khan, Maryam Mashal and Ahmed, Arbab Masood and Khan, Gul Muhammad and Miller, Julian. Fast learning Neural Networks using Cartesian Genetic Programming. Neurocomputing, Volume 121, pages 274-289, Elsevier, 2013.
- [22] Chen, X., and Hurst, S. (1982). A comparison of universal-logic-module realizations and their application in the synthesis of combinatorial and sequential logic networks. IEEE Transactions on Computers, 31, 140-147.
- [23] T. Higuchi et al., Real-World Applications of Analog and Digital Evolvable Hardware, IEEE Transactions on Evolutionary Computation, vol 3, no 3, pp 220-235, Sept 99.
- [24] A. Thompson, On the Automatic Design of Robust Electronics Through Artificial Evolution in Proceedings of the International Conference on Evolvable Systems: from Biology to Hardware, pp13-24, 1998.
- [25] Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.