



SURVEY ARTICLE

CANCER RESEARCH THROUGH THE HELP OF SOFT COMPUTING TECHNIQUES: A SURVEY

Mr. Tanupriya Choudhury¹, Dr. Vivek Kumar², Dr. Darshika Nigam³

¹Research Scholar (CSE), Jagannath University, Jaipur, India

²Principal, Delhi College of Technology and Management, Palwal, India

³Department of BioChemistry, SLS, Dr. B.R. Ambedker University, Agra, India

¹ tanupriya86@gmail.com; ² profvivekkumar@gmail.com; ³ drdarshikanigam@rediffmail.com

Abstract— Soft Computing is a branch of artificial computational intelligence that employs a variety of statistical, probabilistic and optimization techniques that allows computers to “learn” from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets. This capability is particularly well-suited to medical applications, especially those that depend on complex proteomic and genomic measurements. As a result, computational intelligence is frequently used in cancer diagnosis and detection. More recently soft computing has been applied to cancer prophecy and prediction. A number of trends are there, including an increasing dependence on protein biomarkers and microarray data, a strong bias towards applications in prostate and breast cancer, and a heavy reliance on “older” technologies such artificial neural networks (ANNs) instead of more recently developed or more easily interpretable soft computing techniques. Among the better designed and validated studies it is clear that soft computing techniques can be used to substantially to improve the accuracy of predicting cancer susceptibility, recurrence and mortality. In addition to it provides a general idea for further improvement in this field.

I. INTRODUCTION

Bioinformatics [28] is the application of computer science and information technology to the field of biology and medicine. Bioinformatics deals with algorithms, databases and information systems, web technologies, artificial intelligence and soft computing, information and computation theory, software engineering, data mining, image processing, modelling and simulation, signal processing, discrete mathematics, control and system theory, circuit theory, and statistics. Bioinformatics generates new knowledge as well as the computational tools to create that knowledge.

Cancer research is a field of bio-informatics, where we can use classification, clustering algorithmic methods and soft computing techniques for better prediction and understand ability in an earlier stage.

In systems Biology [35] gene regulatory networks have an important role in advance prediction of cancer. By modelling understanding and analysis of these gene regulatory networks dynamics. It may shed light on the mechanism of diseases that occur when these cellular processes are deregulated .Accurate prediction of the behaviour of gene regulatory networks will also speed up in developing personalized medicines and earlier diagnosis.

II. SOFT COMPUTING METHODS

Classification & clustering [29][30] is a method in which Objects are characterized by one or more features
 Classification is a task which assigns objects to classes or groups on the basis of measurements made on the objects

- Have labels for some points
- Want a “rule” that will accurately assign labels to new points
- Supervised learning

Clustering is to group observations that are “similar” based on predefined criteria.

- No labels
- Group points into clusters based on how “near” they are to one another
- Identify structure in data
- Unsupervised learning

Table 1: Soft computing clustering and classification [40]

Various Important Clustering Methods	Various important Classifiers
Hierarchical Methods	Supervised Methods
<ul style="list-style-type: none"> • Agglomerative hierarchical clustering • Divisive hierarchical clustering • Single-link clustering • Complete-link clustering • Average-link clustering 	<ul style="list-style-type: none"> • Naïve Bayes Classifier • J48 Decision Trees • Support Vector Machines
Partitioning Methods	Unsupervised method
<ul style="list-style-type: none"> • Error Minimization Algorithms. • Graph-Theoretic Clustering 	<ul style="list-style-type: none"> • SenseClusters (an adaptation of the K-means clustering algorithm)
Density-based Methods	
Model-based Clustering Methods	Instance-based learning
<ul style="list-style-type: none"> • Decision Trees. • Neural Networks 	<ul style="list-style-type: none"> • Nearest neighbor classifier
Grid-based Methods	Perceptron-based techniques
Soft-computing Methods	<ul style="list-style-type: none"> • Single layered perceptrons • Multilayered perceptrons • Radial Basis Function (RBF) networks
<ul style="list-style-type: none"> • Fuzzy Clustering • Evolutionary Approaches for Clustering • Simulated Annealing for Clustering 	Statistical learning algorithms
	<ul style="list-style-type: none"> • Naive Bayes classifiers • Bayesian Networks • Instance-based learning

Soft computing [31][32] is a branch of computer science capable of analyzing complex medical data. Advances in the area of microarray based expression analysis have led to the promise of cancer diagnosis using new molecular based approaches. Many studies and methodologies have come up which analyzes the gene expression data by using the techniques in data mining such as feature selection, classification, clustering etc. embodying the soft computing methods for more accuracy.

The approach Soft Computing is helpful to be applied for classification, clustering and prediction of cancer as the data contains intangible parameters which are highly non linear and incomplete.

Table 2: Various soft computing techniques in diagnostics of diseases [36][37][38]

SI . No.	SC Techniques used	Diseases cure/detection/recognition
1	Fuzzy logic	Neural system disorder
2	Medical imaging (bio inspired soft computing)	Cancer, arteriosclerosis, epilepsy, alzheimer, parkinson
3	Object-oriented expert system	Diagnosis of fungal diseases of date palm

4	Decision support systems	Diagnosis of disease states and corresponding herbal prescriptions
5	Neural networks,image processing	Oral cysts
6	Artificial neural network	Neonatal diseasediagnosis
7	Decision support system	Congenital heartdisease diagnosis based on signs and symptoms
8	Fuzzy knowledge base	Glaucoma monitoring
9	Clustering techniques	To distinguish the data set to twoprimary clusters i.e. diseased and disease free
10	Classification techniques	To classify a sample at first asdiseased or free from disease and subsequently if diseased then particular type of the disease

Table 3: Survey of computational intelligent learning methods used in cancer prediction showing the types of cancer, clinical endpoints, choice of algorithm, performance and type of training data.

Sr .Nm	Cancer Type	Clinical Endpoint	Computational Intelligent Algorithm	Benchmark	Training Data
1	Bladder	Recurrence	Fuzzy Logic	Statistics	mixed
2	Bladder	Recurrence	ANN	N/A	Clinical
3	Bladder	Survivability	ANN	N/A	Clinical
4	Bladder	Recurrence	ANN	N/A	clinical
5	Brain	Survivability	ANN	Statistics	Genomic
6	Breast	Recurrence	Clustering	Statistics	Mixed
7	Breast	Survivability	Decision Tree	Statistics	Clinical
8	Breast	Susceptibility	SVM	Random	Genomic
9	Breast	Recurrence	ANN	N/A	Clinical
10	Breast	Recurrence	ANN	N/A	Mixed
11	Breast	Recurrence	ANN	Statistics	Clinical
12	Cervical	Survivability	ANN	N/A	Mixed
13	Colorectal	Recurrence	ANN	Statistics	Clinical
14	Colorectal	Survivability	ANN	Statistics	Clinical
15	Colorectal	Survivability	Clustering	N/A	Clinical
16	Esophageal	Treatment response	SVM	N/A	Proteomic
17	Esophageal	Survivability	ANN	Statistics	Clinical
18	Leukemia	Recurrence	Decision Tree	N/A	Proteomic
19	Liver	Recurrence	ANN	Statistics	Genomic
20	Liver	Recurrence	SVM	N/A	Genomic
21	Liver	Susceptibility	ANN	Statistics	Clinical
22	Liver	Survivability	ANN	N/A	Clinical
23	Lung	Survivability	ANN	N/A	Clinical
24	Lung	Survivability	ANN	Statistics	Mixed

The major broad areas of this paper is Soft Computing and Oncology(cancer Classification & Detection) which further embraces of following subjects/areas:

- **Cancer, its study, prediction, recognition techniques & analysis of its most common types:**
Cancer is an abnormal cell-growth occurring in human body and may originate from any of the areas or organs. The disorder can be very dangerous, or even fatal, if ignored for long. It develops in the form of tumors that have a typical tendency to metastasize. Such tumors metastasize or spread to various parts of the body via bloodstream [1].
- Cancer is characterized by out-of-control cell growth. [2,4]. Research requires detailed study of most common Cancer in men and women including Lung, Prostate, Breast and Oral Cancer and their recognition.

The below stated referred research works are classified as cancer detection methods and cancer classification methods. A comparative study is made between the detection methods and the classification methods separately.

Diagnosis[3][5] of any type of cancer in human being: Diagnosis is concerned with the development of algorithms and techniques that are able to determine whether the behavior of a system is correct. If the system is not functioning correctly, the algorithm should be able to determine, as accurately as possible, which part of the system is failing, and which kind of fault it is facing. The computation is based on observations, which provide information on the current behaviour.

Table 4. Comparisons of various cancer detection methods [14]

S. No	Authors	Cancer Type	Technique	Algorithms used	Results	Future Enhancement	Limitations
1	A Banumathi, Praylin Mallika, S Raju	Oral Cysts	Neural networks, Image Processing	Contrast stretching, Radial Basis function	Severity of cysts is measured. For each dental Image accuracy is calculated for classification of Cysts		
2	S murugavalli, V Rajamani	Brain Tumor	Neuro Fuzzy	Fuzzy c-means clustering algorithm	Detected brain tumour at an earlier stage		
3	Ghassan Hamameh, Artur Chodorowski	Oral cancer	Image Processing	Active contour model (snakes)	Segmentation of oral lesion is obtained in single band images from true color images	To further automate and improve segmentation	User assistance is required due to larger variability of objects
4	Varsha H Patil, Vaishali S Pawar	Breast Cancer	CAD, Image processing	Super resolution technique	Detected cancer at very early stage	To simulate the system	
5	H s Sheshadri, A Kandasamy	Breast Cancer	Image processing	Watershed segmentation	Detected cancer tumors at an early stage	A new methodology to extract various parameters which helps to view automatically identifies the suspect lesions	
6	Sibastian Steger, Marius Eddt, Gianfranco	Oral Cancer	Image Processing	Supervised segmentation, image Feature extraction	Oral cancer reoccurrence is predicted automatically	Incorporation of other source modalities like PET	
7	Ranjan Rashmi Paul et al	Oral cancer	Wavelet, neural networks	Multilayered feed forward neural network	The feature vectors are extracted from each contiguous 64*64 blocks by wavelet decomposition		
8	M Muthuramakrishnan, Chandan Chakraborty,	Oral cancer	Wavelet, Data mining, neural network	Bayesian classification, support vector machines	48 gabor wavelet features & 9 wavelet features are extracted	Improvement in accuracy 76.83% accuracy achieved Bayesian classification	

Expert diagnosis

The expert diagnosis (or diagnosis by expert system) is based on experience with the system. Using this experience, a mapping is built that efficiently associates the observations to the corresponding diagnoses. Model-based diagnosis is an example of abductive reasoning using a model of the system.

III. MICROARRAY TECHNOLOGY AND SOFT COMPUTING IN CANCER BIOLOGY

DNA microarray[41] technology has emerged as a boon to the scientific community in understanding the growth and development of life as well as in widening their knowledge in exploring the genetic causes of anomalies occurring in the working of the human body. microarray technology makes biologists be capable of monitoring expression of thousands of genes in a single experiment on a small chip. Extracting useful knowledge and info from these microarray has attracted the attention of many biologists and computer scientists. Soft computing is a branch of computer science capable of analyzing complex medical data. Advances in the area of microarray –based expression analysis have led to the promise of cancer diagnosis using new molecular based approaches. Many studies and methodologies have come up which analyses the gene expression data by using the techniques in data mining such as feature selection, classification, clustering etc. embedding the soft computing methods for more accuracy.

DNA Microarray Technology:

One intense area of microarray[41] research at the NIH is the study of cancer. In the past, scientists have classified different types of cancer based on the organs in which the tumors develop. With the help of microarray technology, however, they will be able to further classify these types of cancer based on the patterns of gene activity in the tumor cells and will then be able to design treatment strategies targeted directly to each specific type of cancer. Additionally, by examining the differences in gene activity between untreated and treated—radiated or oxygen -starved, for example—tumor cells, scientists can better understand how different types of cancer therapies affect tumors and can develop more effective treatments.

DNA microarrays[33] are also generally known as gene-chip or DNA chip. In which it is a group of microscopic DNA spots attached to a solid surface. Scientists utilize DNA microarrays to determine the expression levels of huge numbers of genes concurrently. Significant information can be extracted from these data by the use of data analysis techniques.

In short the usefulness of dna technology can be listed as

1. Can follow the activity of MANY genes at the same time.
2. Can get a lot of results fast
3. Can COMPARE the activity of many genes in diseased and healthy cells
4. Can categorize diseases into subgroups.

Table 5: Use Of Microarray Technology With Soft Computing Cancer Research

Techniques	Main-objectives	Developers
2-way Clustering	Both genes & tumors were clustered	H. Midelfart, A. Læg Reid, and J. Komorowski, Classification of Gene Expression Data in an Ontology, vol. 2199. Lecture Notes in Computer Science, Berlin, Germany: Springer-Verlag, 2001, pp. 186–194
Hierarchical clustering	Found different groups of Breast cancer	M. Banerjee, S. Mitra, and H. Banka, "Evolutionary-rough feature selection in gene expression data," IEEE Trans. Syst., Man, Cybern. C, Appl.
Nearest Shrunken centroid method (PAM)	Limit on the number of genes necessary to prediction.	-
ANN & DCT	Very high success rate for classification of tumor and non tumors	Ahmad M. Sarhan, "Cancer Classification Based on Micro array Gene Expression Data Using DCT and ANN", Journal of Theoretical and Applied Information Technology, Vol. 6, No. 2, pp. 208-216, 2009
Supervised machine learning	High accuracy with only two genes	Bharathi and Natarajan, "Cancer Classification of Bioinformatics data using ANOVA",

		International Journal of Computer Theory and Engineering, Vol. 2, No. 3, pp. 369-373, June 2010
Manifold learning method	Efficient discriminant feature extraction and gene expression data classification.	Bo Li, Chun-Hou Zheng, De-Shuang Huang, Lei Zhang and Kyungsook Han, "Gene expression data classification using locally linear discriminant embedding", Computers in Biology and Medicine, Vol. 40, pp. 802–810, 2010
Rough sets ,feature selection	Superior in applicability and robustness.	Xiaosheng Wang and Osamu Gotoh, "A Robust Gene Selection Method for Micro array-based Cancer Classification", Journal of Cancer Informatics, Vol. 9, pp. 15-30, 2010
Gene ranking and gene subset ranking	Improved classification performance	Mallika and Saravanan, "An SVM based Classification Method for Cancer Data using Minimum Micro array Gene Expressions", World Academy of Science, Engineering and Technology, Vol. 62, No. 99, pp. 543-547, 2010
ANN,classification	Simultaneous pattern extraction,Leukemia classification	S. B. Cho and J. Ryu, "Classifying gene expression data of cancer using classifier ensemble with mutually exclusive features," Proc. IEEE, vol. 90, no. 11, pp. 1744–1753, Nov. 2002. S. Bicchato, M. Pandin, G. Didon`e, and C. DiBello, "Pattern identification and classification in gene expression data using an autoassociative neural network model," Biotechnol. Bioeng., vol. 81, pp. 594–606, 2003.
GA,classification	reliable and accurate classification based on their expression levels,minimization of gene subset size	K. Deb and A. Raji Reddy, "Reliable classification of two-class cancer data using evolutionary algorithms," BioSystems, vol. 72, pp. 111–129, 2003.
NF,feature selection	Feature selection	K. Deb and A. Raji Reddy, "Reliable classification of two-class cancer data using evolutionary algorithms," BioSystems, vol. 72, pp. 111–129, 2003.
Fuzzy NN(dynamic structure growing),feature selection, ANN,classifiers	Colon classification,Classification of acute leukemia, having highly similar appearance in gene expression data	S. Bicchato, M. Pandin, G. Didon`e, and C. DiBello, "Pattern identification and classification in gene expression data using an autoassociative neural network model," Biotechnol. Bioeng., vol. 81, pp. 594–606, 2003. K. Deb and A. Raji Reddy, "Reliable classification of two-class cancer data using evolutionary algorithms," BioSystems, vol. 72, pp. 111–129, 2003.
RS+GA,clustering	effectiveness of the algorithm is demonstrated on three cancer datasets, viz., colon, lymphoma, and leukemia.	S. Mitra, "An evolutionary rough partitive clustering," Pattern Recognit. Lett., vol. 25, pp. 1439– 1449, 2004

IV. COMPARATIVE STUDY ON CANCER (IMAGE DIAGNOSIS) USING SOFT COMPUTING TECHNIQUES USING NEURAL NETWORK AND FUZZY TECHNIQUES

Computer-aided diagnosis system (CAD) can be very helpful for radiologist in detection and diagnosing abnormalities earlier and faster than traditional screening programs. CAD as such employs several techniques to accomplish this task. Here is a comparative study of two classification methods: One in which we utilize the texture features extracted from the images by directly feeding to the Neural Network based classifier stage to classify the images into benign or malign and in the other hybrid method, those texture features are made to undergo fuzzy discretization before feeding to the Neural Network classifier for the classification. The studies so far conducted using both the systems show that the hybrid system is far superior to the first method in its accuracy. Backward Propagation Network (BPN) algorithm is used in the training stage[42].

The field of medicine has its own computer aided, manual as well as automated, tools for various activities. Though diagnosis is easy and simple for many diseases there are few diseases that includes cancer which requires much caution because, the fatal diseases are required to be detected and confirmed in real time or at very early stages, since the available treatment methods call for it. Computer Aided Diagnosis is an automated system, which utilizes techniques available in the areas of data mining, digital image processing and radiology..

Image classifiers established using neural network architecture achieved accuracy to a great extent. A Back Propagation Feed forward network is an interconnected network in which computing elements are arranged in multilayer. The weight associated with each connection modifies the input before they are fed into threshold element. In training the neural network structure, weights are adjusted using the Back Propagation algorithm. Once the learning phase is over, the network is used to perform the image classification. The proposed neural network classifiers make use of spatial information of the image known as features for input.

Fuzzy logic deals with uncertainty and impreciseness in various domains. A hybrid neuro-fuzzy system improves the accuracy and speed of the system. When fuzzified data set is given to neural structure the classification accuracy is improved. Fuzzy discretization is a process that characterizes sub ranges of a continuous variable. Fuzzy version of the crisp data set is defined by the degree of membership of crisp attribute to a fuzzy set by a membership function. In the proposed work trapezoidal membership function is used for fuzzifying the feature data set of images before classifying

Table 6: Comparison of Relative Works

Sr Nm	Author Name	Soft Computing Technique Used	Description
1	Jesmin Nahar	Combining Microarray and Image Data	how image data classification plays a vital role in detecting cancer
2	Haralick	classification of images using texture features	Texture feature shows the difference in the intensity level which could easily identify the cancer images.
3	M.Vasantha	decision tree ID3 algorithm	Since the feature set was not discretized it may affect the accuracy of the classifier. BPN based classifier can provide more accuracy than decision tree algorithm.
4	Qurat-ul-ain	ANN Classification	Features are directly used for classification that leads to inaccuracy. Number of features may increase the computation complexity and minimizes prediction accuracy.
5	Jenn- Lung Su	BPN	compared various data analysis techniques and discussed that BPN network classify the images with high accuracy.

6	Brijesh Verma and John Zakos	Neural network + fuzzy techniques	Classifier is tested with different set of features for accuracy. With all features only 72.2% of accuracy is obtained. Number of iterations needed in training phase is also large. Fuzzification could improve the accuracy and reduce the number of iterations.
---	------------------------------	-----------------------------------	--

V. PREDICTING AND CLASSIFICATION OF CANCER IN DIGITAL MAMMOGRAPHY USING SOFT- COMPUTING TECHNIQUES

Here propose an automatic procedure for digital mammography based on soft-computing technique, for image interpretation, with increased accuracy and a feature subset selection algorithm that selects the most important features, used by Multilayer Perceptron neural networks to classify the digital mammography. In order that the structure of the system can be automatically modified, and evolutionary algorithm is introduced.

What is mammogram: [44]

A mammogram is an x-ray picture of the breast. Screening mammograms are used to check for breast cancer in women who have no signs or symptoms of the disease. Diagnostic mammograms are used to check for breast cancer after

a lump or other sign or symptom of the disease has been found.

Results from randomized clinical trials and other studies show that screening mammography can help reduce the number of deaths from breast cancer among women ages 40 to 74.

Main Points:

- Breast cancer is a malignant tumor that develops when cells in the breast tissue divide and grow without the normal controls on cell death and cell division.
- Although scientists do not know the exact causes of most breast cancer, they do know some of the risk factors that increase the likelihood of a woman developing breast cancer.
- Treatments for breast cancer are separated into two main types, local and systematic. Surgery and radiation are examples of local treatments whereas chemotherapy and hormone therapy are examples of systematic therapies.
- The main goal of breast cancer detection methods is the best possible selection of patients at risk, this means, the selection of the smallest group with the highest risk developing breast cancer.
- Recent advances in multimedia and image processing techniques can be utilized to assist pathologists in this respect propose an automatic procedure for digital mammography based on soft-computing technique, for image interpretation, with increased accuracy.

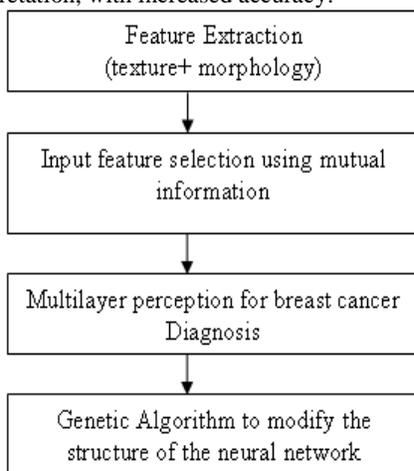


Fig.1: Flowchart of the working procedure

In the above stated figure all the processing steps according to the research paper have been summarized. The last step should be furtherly modified into Proposed algorithm, Solution representation, Selection function and Fitness function. In the experimented result The Classification step occurs after feature extraction and selection have been applied. In our experiments the MIAS MiniMammographic Database was used. Each acquired image has a spatial resolution of 1024x1024 pixels. The various types of breast abnormalities, which are visible in mammograms, include calcification, well-defined/circumscribed masses, spiculated masses, ill-defined masses, architectural distortion and asymmetry. Masses and clustered microcalcification often characterize early breast cancer. In the MIAS database there is also a column indicating the severity of abnormality: Benign or Malignant.

VI. MODELLING BIOLOGICAL NETWORKS USING SOFT COMPUTING TECHNIQUES

In systems Biology gene regulatory networks have an important role in advance prediction of cancer. By modelling understanding and analysis of these gene regulatory networks dynamics. We may shed light on the mechanism of diseases that occur when these cellular processes are deregulated .Accurate prediction of the behaviour of gene regulatory networks will also speed up in developing personalized medicines.

Motivation for modeling GRN:

- To present a synthetic network view of the currently available biological knowledge and to structure it in such a way that it brings to sight relevant properties which may remain hidden without the appropriate model.
- To predict dynamic behaviour of the network. These predictions are compared with experimental results. It may allow either confirmation of the model’s accuracy or recommend correction of the model.
- The complexity of molecular and cellular interactions requires modeling tools that can be used to properly design and interpret biological experiments.

Soft Computing Techniques:

- Fuzzy logic
- artificial neural networks
- evolutionary algorithms (genetic algorithm, genetic programming and evolutionary strategies)
- simulated annealing
- swarm optimization and probabilistic reasoning are fundamental computing constituents of soft computing

Among the various soft computing constituents, fuzzy logic (FL), artificial neural networks (ANNs) and evolutionary algorithms (EAs) are considered as the core methodologies of soft computing

Table 7: Hybradized models for Modeling GRNs [43]

Modeling Techniques	Results Obtained	References
Neuro Fuzzy	Reconstruction of partial GRN of yeast	Liu et. Al,2011
Neuro Fuzzy	Extract regulatory relationship & construct GRN	Vineetha et. Al,2010
RNN+Fuzzy	Extracted GRN from Yeast	Maraziotis,et. Al.,2010
Clustering+PSO+RNN	Inferred GRN	Zhang,et.al.,2009
RNN+Fuzzy	Determine regulatory interaction from genes	Datta et. Al.,2009

RNN+GA	Extracted GRN modules	Chiang & chao,2007
Neuro Fuzzy	Reconstructed GRN from microarray data	Jung & Cho,2007
RNN+PSO	Extracted GRN from gene expression profiles	Xu Rui et. Al. 2007

VII. CONCLUSION AND FUTURE ENHANCEMENT

Cancer classification, prediction and diagnosis is an emerging research area in the field of Bio-informatics. In this survey various soft-computing methods, data-mining and machine learning based algorithms for gene selection and cancer classification were discussed in detail. And also we have attempted to explain compare and performing of soft-computing methods which are using of cancer classification , prediction and prognosis to detect it in a earlier stage. specifically in a personalized way we identified a number of trends with respect to the types of computational intelligent methods being used and the types of training data being incorporated ,the kinds of endpoint predictions being made ,the types of cancers being studied, and the overall performance of these methods to predict cancer.

In future better neural network techniques can be incorporated with the present research work for less complexity and better learning adaptability. Moreover ,better neuro fuzzy techniques could also be used to improve the classification rate and accuracy. Better microarray techniques can also be adapted as well as multi objective ant-colony approach and genetic algorithmic approach for better prediction of cancer and to enhance the previous techniques.

REFERENCES

- [1] <http://www.cancer8.com/>
- [2] http://www.wcrf.org/cancer_facts/5-most-common-cancers.php
- [3] <http://www.cancer.gov/cancertopics/cancerlibrary/what-is-cancer>
- [4] <http://www.medicalnewstoday.com/info/cancer-oncology/>
- [5] http://www.wcrf.org/cancer_facts/5-most-common-cancers.php
- [6] <http://www.scribd.com/doc/9522057/an-overview-of-soft-computing-with-its-techniques-and-application>
- [7] <http://dl.acm.org/citation.cfm?id=1539227>
- [8] <http://www.iconip2012.org/s02.pdf>
- [9] <http://www.biomedcentral.com/1471-2105/6/148>
- [10] <http://news.harvard.edu/gazette/tag/cluster-of-cancer-cells/>
- [11] <http://www.healthcommunities.com/cancer-treatment-and-care/cancer-staging.shtml>
- [12] http://en.wikipedia.org/wiki/Oncology#Progress_and_research
- [13] <http://www.cancer.net/patient/All+About+Cancer/Cancer.Net+Feature+Articles/Treatments%2C+Tests%2C+and+Procedures/Understanding+Cancer+Research+Studies%2C+Part+II>
- [14] K. Anuradha and K.Sankaranarayanan "International Journal of Advances in Engineering & Technology, March 2012" Vol. 3, Issue 1, pp. 84-91
- [15] M. E. Futschik, A. Reeve, and N. Kasabov, "Evolving connectionist systems for knowledge discovery from gene expression data of cancer tissue," *Artif. Intell. Med.*, vol. 28, pp. 165–189, 2003
- [16] H. Midelfart, A. Læg Reid, and J. Komorowski, *Classification of Gene Expression Data in an Ontology*, vol. 2199. Lecture Notes in Computer Science, Berlin, Germany: Springer-Verlag, 2001, pp. 186–194
- [17] M. Banerjee, S. Mitra, and H. Banka, "Evolutionary-rough feature selection in gene expression data," *IEEE Trans. Syst., Man, Cybern. C, Appl.*
- [18] Ahmad M. Sarhan, "Cancer Classification Based on Micro array Gene Expression Data Using DCT and ANN", *Journal of Theoretical and Applied Information Technology*, Vol. 6, No. 2, pp. 208-216, 2009
- [19] Bharathi and Natarajan, "Cancer Classification of Bioinformatics data using ANOVA", *International Journal of Computer Theory and Engineering*, Vol. 2, No. 3, pp. 369-373, June 2010
- [20] Bo Li, Chun-Hou Zheng, De-Shuang Huang, Lei Zhang and Kyungsook Han, "Gene expression data classification using locally linear discriminant embedding", *Computers in Biology and Medicine*, Vol. 40, pp. 802–810, 2010

- [21] Xiaosheng Wang and Osamu Gotoh, "A Robust Gene Selection Method for Micro array-based Cancer Classification", *Journal of Cancer Informatics*, Vol. 9, pp. 15-30, 2010
- [22] Mallika and Saravanan, "An SVM based Classification Method for Cancer Data using Minimum Micro array Gene Expressions", *World Academy of Science, Engineering and Technology*, Vol. 62, No. 99, pp. 543-547, 2010
- [23] K. Deb and A. Raji Reddy, "Reliable classification of two-class cancer data using evolutionary algorithms," *BioSystems*, vol. 72, pp. 111–129, 2003.
- [24] H. Midelfart, J. Komorowski, K. Nørsett, F. Yadetie, A. K. Sandvik, and A. Læg Reid, "Learning rough set classifiers from gene expression and clinical data," *Fundamenta Inf.*, vol. 53, pp. 155– 183, 2002.
- [25] F. Chu, W. Xie, and L. Wang, "Gene selection and cancer classification using a fuzzy neural network," in *Proc. 2004 Annu. Meet. North Amer. Fuzzy Information Processing Soc. (NAFIPS)*, vol. 2, pp. 555–559
- [26] S. Mitra, "An evolutionary rough partitive clustering," *Pattern Recognit. Lett.*, vol. 25, pp. 1439– 1449, 2004
- [27] S. Biccato, M. Pandin, G. Didone, and C. DiBello, "Pattern identification and classification in gene expression data using an autoassociative neural network model," *Biotechnol. Bioeng.*, vol. 81, pp. 594–606, 2003.
- [28] <http://en.wikipedia.org/wiki/Bioinformatics>
- [29] www.broadinstitute.org/...2006/.../Classification_and_Clustering.ppt
- [30] <http://www.ims.nus.edu.sg/Programs/microarray/files/SingDiscV2.pdf>
- [31] <http://cscjournals.org/csc/manuscript/Journals/IJBB/volume5/Issue4/IJBB-133.pdf>
- [32] An Extensive Survey on Artificial Neural Network Based Cancer Prediction Using SOFTCOMPUTING Approach. Manaswini Pradhan, Dr. Ranjit Kumar Sahu
- [33] <http://ijcta.com/documents/volumes/vol2issue5/ijcta2011020550.pdf>
- [34] <http://www.iula.upf.edu/materials/040701wanner.pdf>
- [35] Sami Halawani and Rafat Parveen, *International Journal of Research and Reviews in Computer Science (IJRRCS)*, Vol. 3, No. 1, February 2012, ISSN: 2079-2557
- [36] <http://www.softcomputing.es/metaspaces/portal/7/274>
- [37] <http://www.medwelljournals.com/abstract/?doi=ijscmp.2009.201.207>
- [38] http://www.scribd.com/chris_cao_11/d/82513248-A-Proposed-Model-of-SoftComputing-in-Diagnosing-Diseases-and-Prescribing-Herbal-Prescriptions-by-Oriental-Medicine
- [39] www.la-press.com/redirect_file.php?...Soft-Computing-Cancer-Classification
- [40] Megha Gupta, Naveen Agarwal "Classification Techniques Analysis", NCCI 2010 -National Conference on Computational Instrumentation
- [41] <http://www.mrc-lmb.cam.ac.uk/genomes/madanm/microarray/chapter-final.pdf>
- [42] <http://www.ijcaonline.org/volume19/number5/pxc3873087.pdf>
- [43] www.sciacademypublisher.com/journals/index.php/IJRRCS/.../581
- [44] <http://metronu.ulb.ac.be/imacs/papers/T3-I-114-0597.pdf>