

Screening of Cervical Cancer by Artificial Intelligence based Analysis of Digitized Papanicolaou-Smear Images

Rajeev Gupta¹, Abid Sarwar², Vinod Sharma³

ABSTRACT

Introduction: Cervical cancer is one of the main causes of female mortality due to cancer in the world. Researchers have suggested that by regular cervical cancer screening its incidence and mortality can be reduced by about 84%. Study aimed to apply multiple artificial intelligence based algorithms on the task of screening of cervical cancer and compare their working efficiency and to select the best performing algorithm and use that for developing intelligent software tool that can aid in the prognosis of cervical cancer.

Material and methods: Fifteen different machine learning algorithms were implemented for screening of cervical cancer by analysis of digitized Papanicolaou-smear images using a primary dataset for training. Slides of Pap smear test of about 200 clinical cases were collected from multiple health care institutions in north India and images were captured using a digital microscope. Images of 8091 cervical cells were obtained from these slides, which were quantitatively profiled and calibrated to prepare a primary database that could reflect the morphological features of the cells of cervix. All the algorithms were trained using this database and were tested using a subset of unseen cells. The generalizing ability of algorithms was tested and its performance was evaluated against the human pathologist using percentage of correctly classified cells and sensitively.

Results: Almost all the algorithms performed well for identifying the cells infected by cancer by classifying the cells according Bethesda system of classification. Among all the algorithms used, multiple backpropagation neural networks presented an efficiency of about 78.0%, whereas the efficiency of all other algorithms was in the range of 69% to 76%.

Conclusion: The results demonstrate that the artificial intelligence based techniques can effectively be used for developing tools for mass level screening of cervical cancer.

Keywords: Artificial Intelligence, Artificial Neural Networks, Cervical cancer, Medical Database, Pap-smear test.

to multiply with uncontrolled cell division and death. Since these cells are uncontrolled and unmanaged, they start multiplying enormously and get transformed into lump of mass called tumor. These cancerous cells can infect the other parts of body too by flowing through the blood stream. The infected cells are called as intraepithelial neoplasia or CIN. Usually the CIN remains constant and is purged by the response of the immune system, but in very small proportion of cases, if CIN is not timely treated, it grows to become a cancer. Scientific investigation has confirmed that CIN results from human papillomavirus (HPV) which is usually transmitted sexually. In the initial stages Cervical cancer can usually doesn't show-up any symptoms. As the cancer progresses to become an invasive cancer, symptoms start to appear in the form of abnormal vaginal discharge, abnormal vaginal bleeding, pain during vaginal intercourse. For women who have had menopause, new vaginal bleeding may occur. Papanicolaou test also called as Pap-Smear test has been a primary test used for screening of cervical cancer since past many decades and in many developed countries where Pap-Smear test is regularly done has also lead to steep lowering of the number of diagnosed cases.^{2,3} The main aim of this test is to report any unusual developments in that may transform into cervical dysplasia. In many developed countries regular screening through pap-smear test is recommended to women who have a history of having sex with several partners. After obtaining the sample, it is stained using Papanicolaou technique, which helps to differentiate the cervical cells from other bodily secretions. Research in past many years has witnessed an increasing use of Computer based intelligent techniques for solving complex problems in the domain of medical science such as Liver Pathology, Perinatology, Urology, Thyroid disorders, Oncology, Gynecology, Cardiology, etc.⁴⁻¹⁰ The main aim of using Artificial Intelligence techniques in medical science is development of such computer based intelligent systems as can support human experts in making more precise and timely diagnosis. Development of such diagnostic systems has also led to a considerable advancement in our understanding of clinical expertise, transformation of such expertise into abstract models, and further conversion of those abstract models into

INTRODUCTION

Timely and precise diagnosis of cervical cancer is an important real-world medical problem. Cervical cancer has turned into one of the primary causes of mortality among women around the globe. This being the fourth most common women cancer in the world, researchers throughout the world are gravely concerned for finding a solution to it so that an early and timely diagnosis of this disease could be realized and as such the overall mortality rate could be mitigated. The total number of cases of cervical cancer in 2012 reported by World Health Organization (WHO) was 528000 and the number of deaths equal to 266000. Most of these cases about 84% occur in under developed and developing countries as against the developed countries; the reason being reduced access to proper screening, treatment and follow-up services.¹ Cervical cancer is a tumor, malignant in nature which occurs when the cervical cells begin

¹Assistant Professor; Department of Medicine, Guru Ramdas Medical College Amritsar, Punjab, ²Professor, Department of Computer Science and IT, ³Assistant Professor, Department of Computer Science and IT (BC), University of Jammu, J&K, India

Corresponding author: Rajeev Gupta, Assistant Professor; Department of Medicine, Guru Ramdas Medical College Amritsar, Punjab, India

How to cite this article: Rajeev Gupta, Abid Sarwar, Vinod Sharma. Screening of cervical cancer by artificial intelligence based analysis of digitized papanicolaou-smear images. International Journal of Contemporary Medical Research 2017;4(5):1108-1113.

result oriented diagnostics systems.¹¹⁻¹⁵ So present study aimed to apply multiple artificial intelligence based algorithms on the task of screening of cervical cancer and compare their working efficiency and to select the best performing algorithm and use that for developing intelligent software tool that can aid in the prognosis of cervical cancer.

MATERIAL AND METHODS

The performance of any machine learning system is largely dependent on the data with which it is trained. It is said that any human behavior can be copied into a machine provided that it is trained using a suitably precise database.¹⁶⁻¹⁸ The database used in this work to train the algorithms is made-up of 8,091 cervical cells which have been profiled on the basis of 40 morphological properties. Out of these 40 morphological attributes 19 belong to cytoplasm, 19 belong to nucleus, one attribute consists of the ratio of area of Nucleus and Cytoplasm, and one last attribute contains the cells label depicting as to which Bethesda class the cell belong to. The data for study i.e slides of Pap-smear test were obtained from three leading health care institutions from different geographical locations in northern India viz Government Medical College, Jammu, Acharaya Shree Chandra College of Medical Science and Hospital, Jammu and Nijjer Pathology Laboratory, Amritsar. These slides were observed under different levels of magnifications under a microscope with a high resolution digital camera attached to it (NIKON Nikon Eclipse E400 DS-F12). Images of the area of interest on the slides were captured at 40x magnification followed by multiple

rounds of image preprocessing subroutines to enhance visibility of individual discrete cells. Once the cells were distinctly visible, the individual cells were cropped off from cells clusters and unique-ids were assigned to them so as to identify each cell distinctly in the database. Once a considerable set of cervical cells were obtained, these cells were analyzed, inspected and tagged by multiple pathologists so as to ensure the accuracy and consistence in the database. Figure 1 show the block diagram of how the clinical data was transformed into a digital database.

For profiling of the cells so obtained, as open source software "CellProfiler" was used which evaluated the cells in the basis of 39 morphological features (figure-1). These profiled measurements of the cells were then arranged in a spreadsheet in such a way that each tuple represented a cervical cell. Various other properties of database were also integrated so that digital data could be easily correlated with the corresponding clinical case. A detailed analysis of morphological features is prerequisite for any automated cervical screening system as such the features set chosen for quantitative evaluation have to be such that they should be precisely measurable in specific time with good accuracy.¹⁹ Computer techniques based on various image processing algorithms provide simple and proficient formulization of morphological features as compared to the manual interpretation by cyto-pathologists which are subjective, laborious and many times prone to human errors.^{20,21}

Algorithms used and there implementation

The algorithms considered were Random subset space, Decorate, Ensemble of Nested dichotomies also called

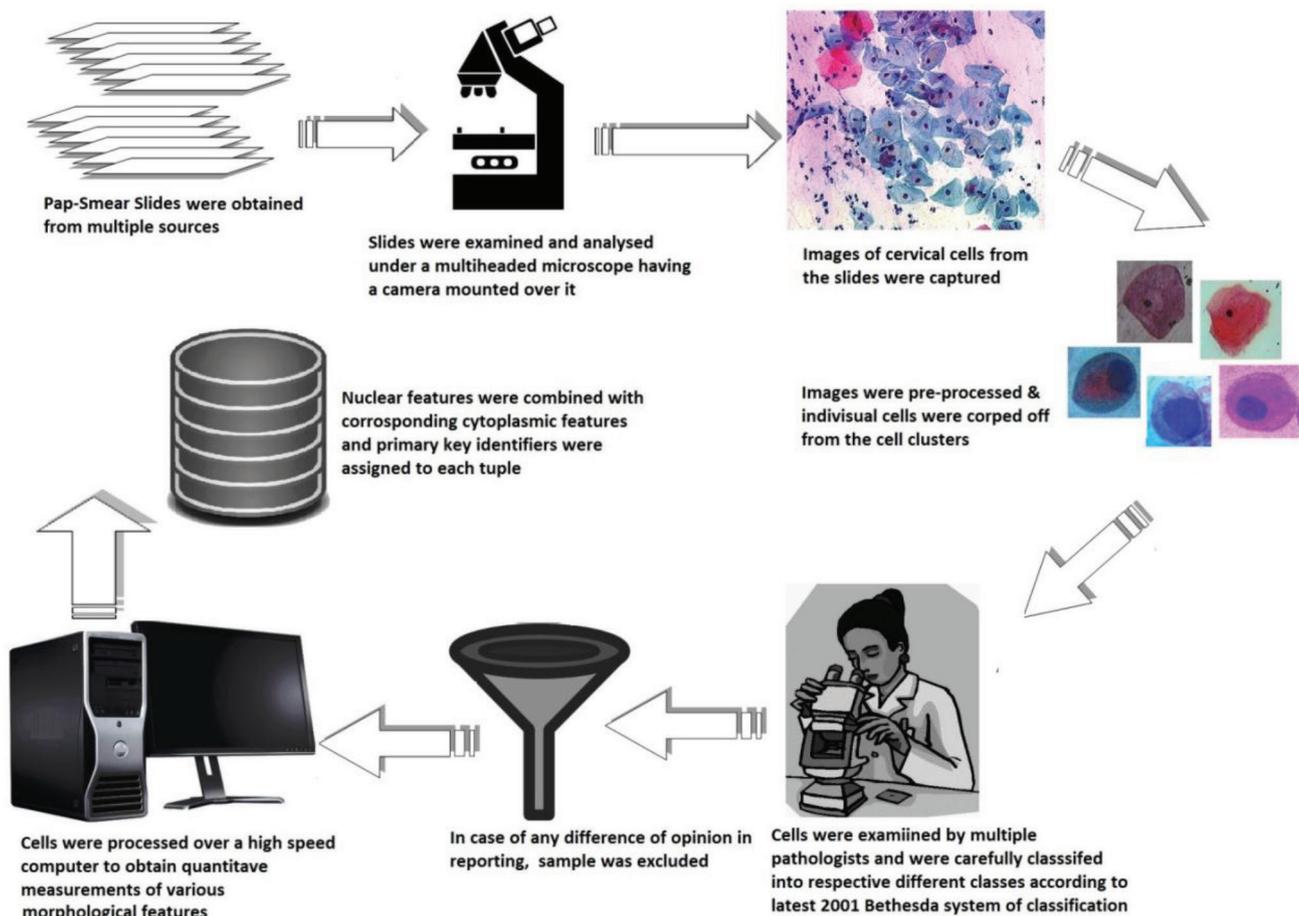


Figure-1: Block diagram of various steps showing how digital database was prepared

END, Back propagation artificial neural network, J48 graft, Radial basis function network (RBF), Random Committee, Decision Table, PART, Naïve Bayes, Multi-class classifier, Filtered Classifier (FC), Rotation Forest, Random forest and Bagging. These Algorithms were selected due to their reported excellent performance in recently published literature. Bagging partitions the dataset into several subsets which are used to train several models and the final output model is obtained by cumulative results from all the training models. In random forest, multiple decision trees are made and for each input tuple one path is traced in each of the decision trees constructed and this is repeated for all the training tuples. The final classification class for an instance is considered by taking the votes from all the decision trees.

Random subspace is a general form of random forest algorithm, in random forest, decision trees are used as base classifiers whereas in random subspace any type of classifier can be used as base classifier. The system of nested dichotomies recursively transforms the multiclass categorization problem into a relatively easy two class categorization problem which can easily be represented using binary trees. The resultant classification is obtained by aggregating the results obtained by the individual binary trees. In random committee, an assembly of individual multiple tree classifiers is made, each of which is trained and tested with a common dataset but using a different randomness generating seed value. The DECORATE algorithm iteratively adds a new learning classifier to the current ensemble at each iteration. The data used for training is built by adding some artificial data to the original data. A new classifier is added only if on addition of new classifier, the total error is reduced. Rotation forest works by creating an ensemble of classifiers each trained using a different subset of attributes from the main dataset. J48-Graft creates a special decision tree which is grafted by a technique that adds fresh nodes on the places of either the old leaf nodes or in between the leaf nodes such that the accuracy in prediction is improved. Naïve Bayes classifier is based on the theory of conditional independence, which assumes that each attribute has its independent roll in the prediction of the final outcome and is also independent of the presence or absence of other attributes. The multiclass classifier works on the principle of divide and conquer technique and transforms the complex problem into simple two class problem which can be solved using one versus all approach where m -binary classifiers are built to identify m -classes. PART stands for partial decision tree algorithm and is a merger of RIPPER rule and C4.5 algorithm. The algorithm works by producing a set of rules called decision lists. A data item to be tested is evaluated with each rule in the list and the test item is classified to the class of first matching rule. Multiple back propagation algorithm (MBP) is a generalization of back propagation (BP) algorithm where there are two sub networks i.e. main and space network. The space network determines factors for the activation of neurons in the main network; as such the main network neurons get activated only for a set of training patterns rather than for all as in multilayer perceptron. The RBF stands for Radial basis function network that uses a non-linear activation function in the hidden layers and linear activation function in the output layer. The training of RBF network is relatively faster than other neural network variants. Decision table converts the training dataset into a decision table

and the data is classified into one of the classes by tracing the rows of the table.

The realization of algorithms was carried out using WEKA workbench which is a Java coded machine learning data analyzer. Before feeding the data to the system it was transformed to Attribute-Relation File Format (ARFF) format as required by WEKA workbench. The total training data was divided in the ratio of 70:30, in such a way that 70% (5,664 cell images) of data was used for training while as 30% (2,427 cell images) of the data was used for testing of the algorithms. The experimentation performed on the benchmark database using neural networks was of two kinds. In the first, we tested the ability of neural network to handle the two-class diagnostic problem i.e. to identify the case to be a normal (classes 1, 2, 3 and 4) or abnormal (5, 6 and 7) and in the second, we tested the ability of the network for 7-class problem i.e. to identify and segregate the cells into their respective class among the seven classes as done by an expert human cyto-pathologist. All these experiments were carried out with 70:30 training-testing approach, along with 10 fold cross validation. Each of the algorithms was trained with the same data derived out of the benchmark database built in this work. The overall performance of algorithms was evaluated using 10 fold cross validation.

RESULTS

In general, all the algorithms performed better for 2-class problem as compared to the 7-class problem. For 2-class problem algorithms showed efficiency in the range of 93% to 95%, while as for 7-class problem, efficiency in the range of 69% to 78%. The reason for this is attributed to the increase in the complexity of the classification problem for classifying the cell instances into corresponding class among 7 classes as compared to classifying into just 2 classes. Among all the algorithms used, multiple backpropagation neural networks presented an efficiency of about 78.0%, whereas the efficiency of all other algorithms was in the range of 69% to 76%. Detailed results regarding the accuracy obtained by each intelligent methodology on both the 2-class and the 7-class problem are presented in Table 1.

For developing the semi-automated cervical cancer screening tool we used multiple back propagation algorithm, it being the best predictor among all the algorithms under consideration. The architecture used was standard backpropagation-based multilayer perceptron (MLP). In this neural network the neurons are organized in layer based graph with directed edges where each layer is fully connected to the next layer in feed-forward manner. The randomly initialized weights are updated and adjusted using the error values between network outputs and actual outputs. Backpropagation network training is a systematic way of training the network using extended gradient-descent technique based on delta-learning rule in order to determine and adjust the weights in the direction of minimizing error function. For processing the inputs at the artificial neurons, sigmoid function was used as the activation function on the network nodes both at the hidden layer and output layer. For realization of desired ANN, we have developed a self designed decision support tool in Matlab. The network was trained and tested for different number of hidden neurons and the weights corresponding to the best trained network were saved. The network exhibited the best results for 21 neurons in the hidden layer so the numbers

S No.	Algorithms	Binary classification problem		Multiclass classification problem	
		Correctly classified (%)	Sensitivity	Correctly classified (%)	Sensitivity
1.	Bagging	95.2536%	0.953	76.4629%	0.765
2	PART	94.4083%	0.944	73.5587%	0.736
3	Decision Table	94.5384%	0.945	74.0355%	0.74
4	Decorate	95.0585%	0.951	76.4629%	0.769
5	END	93.9965%	0.940	75.8778%	0.759
6	Filtered Classifier	94.7334%	0.947	74.5340%	0.745
7	J48 Graft	94.4083%	0.944	73.6238%	0.736
8	Multi-Class Classifier	94.9502%	0.950	76.2029%	0.762
9	Rotation Forest	95.4270%	0.954	75.1409%	0.751
10	Naïve Bayes	93.6064%	0.936	69.7226%	0.697
11	Radial Basis Function	94.1482%	0.941	74.7941%	0.748
12	Random Committee	95.3836%	0.954	76.2029%	0.762
13	Random Forest	95.2536%	0.953	75.3576%	0.754
14	Random Sub Space	95.2536%	0.953	76.9181%	0.766
15	Multiple Backpropagation ANN	95.6220%	0.956	78.0668%	0.781

Table 1: Results of 10 folds cross validation for different algorithms

of neurons in the input, hidden, and output layers are 40, 21, and 1 respectively. The stopping criteria of the training process was 500 epochs or mean absolute error value less than or equal to 0.01. The weight matrices that stored the weights of the connections between input-hidden nodes and that of between hidden-output nodes were initialized by a randomized function that returned values between -1 and +1. The training and testing datasets were normalized before feeding to the network using Mean (μ) and Standard deviation (σ) of the corresponding datasets. Once the network was trained and presented good generalization accuracy, multiple test instance previously unseen by the network were fed to the system for validating the effectiveness of the tool. The semi-automated tool is developed in two modules, one module is used for training the network and the other module is used for querying the trained network. The training and querying module of the Matlab program are shown in the Figure 2(a) and Figure 2(c) respectively.

DISCUSSION

Cervical cancer being one of the most common women cancers in the world is a major cause of female mortality. In developed countries there is a proper system of annual regularly screening all the females that are susceptible to cervical cancer, unfortunately in low income countries and countries having a large boom of population like India, such mass level screening programs are very difficult to implement owing to various economic and social constraints. India is contributing to more than 25% of the whole world's burden of cervical cancer which accounts for death of eight patients every hour. India being a developing nation with world's second largest share of population, implementation of regular screening for cervical cancer like that of pulse polio is next to impossible. In such a scenario computer based intelligent screening techniques can prove to be very useful in reduction of burden of cervical cancer on a nation in particular and on the whole world in general. Such tools also can help to identify the cases that are more prone to this disease and as such can aid in timely identification of cases having premalignant transformations that may later on develop into cancers.

Authors also compared the prognostic ability of the cervical

cancer screener proposed in this work with the machine learning based diagnostic systems developed by other researchers. Olivier Regnier-Coudert et.al in 2012 have done a comparison of 3 different machine learning models for prognosis of Prostate cancer.²² The models used by them were artificial neural networks, bayesian networks and logistic regression which expressed the true positive rates of 0.656, 0.679 and 0.675 respectively. Compared with the prognostic results of these automated system build for prostate cancer, the novel cervical cancer screener shows much better sensitivity of 0.781, which is very close to the best possible classification potential by any classifier. Sokouti et al., 2014, have proposed also proposed a model (Levenberg–Marquardt feedforward MLP neural network-LMFFNN) for classification of cervical cell images.²³ The system proposed by them is trained with data obtained from 100 patients and classifies the cervical cells in normal and cancerous. The system proposed in this work not only classifies the cells into normal and abnormal, but also categorizes them according to Bethesda system of classification which is universally accepted system of reporting cervical smears. Chiu et al., also studied the application of artificial neural networks in predicting the skeletal metastasis in patients suffering from prostate cancer.²⁴ They analyzed whole body bone scintigraphies of patients with prostate cancer who underwent the technetium-99m methylene diphosphate (tc-99m MDP). For evaluating the performance of ANN, Receiver Operating Characteristics (ROC) analysis was used. The area under the ROC curve (0.88 ± 0.07) revealed excellent discriminatory power ($p < 0.001$) with the best simultaneous sensitivity (87.5%) and specificity (83.3%). Authors concluded that that an ANN, which is based on limited clinical parameters, to be a promising method in forecasting of the skeletal metastasis. The system proposed in this work is comparable with their automatic diagnostic system. Artificial intelligence based diagnostic systems have several advantages over conventional methods, as these programs have a potential to greatly reduce the number of diagnostic possibilities by using the patho-physiological reasoning.¹⁰⁻¹² As such they can create a logical model of a specific patient's medical condition. Such models can also model the complexities generated by overlapping symptoms of multiple diseases.

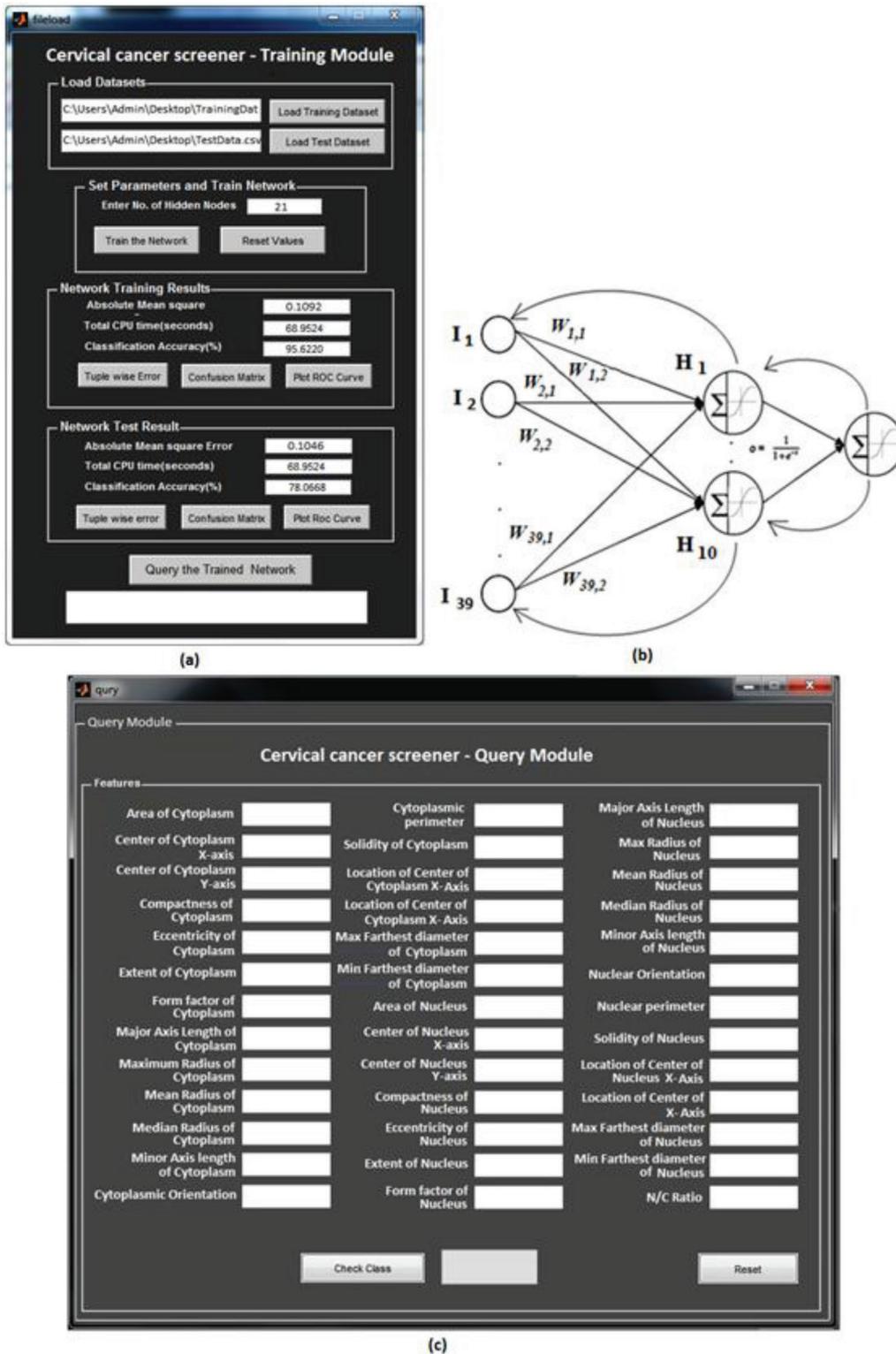


Figure-2: (a) Training module of cervical cancer screener showing various sub modules in execution. (b) Schematic representation of topology of neural network used for developing cervical cancer screener tool. (c) Query module of cervical cancer screener tool

Although only a few such programs are currently available, the evidence suggests that the continued development of artificial intelligence techniques will eventually give the computer a major role as an expert consultant to the physician.

CONCLUSION

The results demonstrate that the artificial intelligence based techniques can effectively be used for developing computer

assisted intelligent tools for mass level screening of cervical cancer.

REFERENCES

1. Kent, A. HPV Vaccination and Testing. Reviews in obstetrics and gynecology. 2010;3:33–34.
2. Vinod B Shidham et al. p16 INK4a immunocytochemistry on cell blocks as an adjunct to cervical cytology: Potential

- reflex testing on specially prepared cell blocks from residual liquid-based cytology specimens. *CytoJournal*. 2011;8:1-4.
3. Saslow, D, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer. *Journal of Lower Genital Tract Disease*. 2012;8:23-27.
 4. Lisboa PJ, Taktak AF. The use of artificial neural networks in decision support in cancer: a systematic review. *Elsevier Neural Networks*. 2006;19:408–415.
 5. James W. F. Catto, Derek A. Linkens, Maysam F. Abbod, Minyou Chen, Julian L. Burton, Kenneth M. Feeley, and Freddie C. Hamdy. Artificial Intelligence in Predicting Bladder Cancer Outcome: A Comparison of Neuro-Fuzzy Modeling and Artificial Neural Networks. *Clin Cancer Res*. 2003;9:4172-79.
 6. Anagnostou T, Remzi M, Lykourinas M, Djavan B. Artificial neural networks for decision-making in Urologic Oncology. *European Urology*. 43(6), pp.596–603.
 7. I. Bratko and I. Kononenko. *Learning Rules from Incomplete and Noisy Data. Interactions in Artificial Intelligence and Statistical Methods*, Hampshire, Technical Press, Hampshire, England, 1987.
 8. L. Lesmo, L. Saitta, P. Torasso. Fuzzy Production Rules: a Learning Methodology. *Advances in Fuzzy Sets, Possibility Theory and Applications*, pp. 181-198.
 9. J. Catlett. On changing continuous attributes into ordered discrete attributes. *Proceedings of European Working Session on Learning-91, Portugal, March 4-6 1991*, pp. 164-178.
 10. P. Clark and R. Boswell. Rule Induction with CN2: Some Recent Improvements. *Proceedings of European Working Session on Learning-91, Portugal, March, 1991*, pp.151-163.
 11. Nunez M. Decision Tree Induction Using Domain Knowledge. *Current Trends in Knowledge Acquisition*, Amsterdam, IOS Press, 1990.
 12. Hojker S, Kononenko I, Juka A, Fidler V. and Porenta. M, Expert System's Development in Management of Thyroid Disease, *proc. European Congress for nuclear medicine*, Milano, Sept., 1988
 13. Horn K.A, Compton P.Lazarus L., Quinlan J.R, an Expert System for Interpretation of Thyroid Assays in Clinical Laboratory. *The Australian Computer Journal*. 1985;17:7-11.
 14. J. Kern, Dezelic G., Tezak-Bencic M., Durriegl T. Medical Decision Making Using Inductive Learning Program. *Proceedings of 1st Congress on Yugoslav Medical Informatics, Beogard*. 1990, pp.221-228.
 15. Abid Sarwar and Vinod Sharma. *Comparative analysis of machine learning techniques in prognosis of type II diabetes*. AI and Society, Springer Verlag London, April 2013
 16. Babak Sokouti et.al. A framework for diagnosing cervical cancer disease based on feedforward MLP neural network and ThinPrep histopathological cell image features. *Neural Computing and Applications*. 2012;6:23-29.
 17. Nor Ashidi Mat-Isa, Mohd Yusoff Mashor and Nor Hayati Othman, "An automated cervical pre-cancerous diagnostic system", *Artificial Intelligence in Medicine*. 2008;42:1-11.
 18. Chun-Lang Chang and Ming-Yuan Hsu. The study that applies artificial intelligence and logistic regression for assistance in differential diagnostic of pancreatic cancer. *Expert Systems with Applications*. 2009;36:10663-10672.
 19. V. Nandakumar, L. Kelbauskas, R. Johnson, and D. Meldrum. Quantitative characterization of preneoplastic progression using single-cell computed tomography and three dimensional karyometry. *Cytometry Part A*, 2011;79:25–34.
 20. P. Bamford and B. Lovell. A water immersion algorithm for cytological image segmentation, in *Proc. APRS Image Segmentation Workshop, Sydney, Australia, 1996*, pp. 75–79.
 21. O. Lezoray and H. Cardot, "Cooperation of color pixel classification schemes and color watershed: A study for microscopic images," *IEEE Trans. Image Process*. 2002; 11:783–789.
 22. Olivier Regnier-Coudert et.al. Machine learning for improved pathological staging of prostate cancer: A performance comparison on a range of classifiers, *Artificial intelligence in Medicine*, Elsevier. 2012;55:25–35.
 23. Sokouti et al. A framework for diagnosing cervical cancer disease based on feedforward MLP neural network and Thin Prep histopathological cell image features. *Neural Computing and Applications*. 2014;24:221-232.
 24. Chiu, Jaiin-Shiun, et al. Artificial neural network to predict skeletal metastasis in patients with prostate cancer. *Journal of medical systems*. 2009;33:91-95.

Source of Support: Nil; **Conflict of Interest:** None

Submitted: 25-02-2017; **Accepted:** 02-05-2017; **Published:** 06-06-2017