ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

Diagnosis Support System for Lung Cancer Detection Using Artificial Intelligence

K. V. Bawane[1], A. V. Shinde [2]

Student, Dr. Bhausaheb Nandurkar College of Engineering and Technology Yavatmal, India

Assistant Professor, Dr. Bhausaheb Nandurkar College of Engineering and Technology Yavatmal, India

ABSTRACT: In this paper a new classification algorithm is proposed for the Efficient Classification of Lung Tumor. In order to develop algorithm 150 CT scan images of patients have been considered consisting of Benign and Malignant Tumor Computed tomography (CT) Scan image. With a view to extract features from the CT scan images after image processing, an algorithm proposes WHT Transform domain coefficients. The Efficient classifiers based on Multilayer Perceptron (MLP) Neural Network. A separate Cross-Validation dataset is used for proper evaluation of the proposed classification algorithm with respect to important performance measures, such as MSE and classification accuracy. The Average Classification Accuracy of MLP Neural Network comprising of one hidden layers with 37 PE's organized in a typical topology is found to be superior (95.54 %) for Training . Finally, optimal algorithm has been developed on the basis of the best classifier performance. The algorithm will provide an effective alternative to traditional method of Lung Computed tomography (CT) scan image analysis for deciding the tumor in lung is Benign or Malignant.

KEYWORDS: Neural solution, MATLAB,CT scan images.

I. INTRODUCTION

The early detection of lung cancer is a most challenging problem to identify in human body, due to the different - different cells structure of cancer, where most of the cells are overlapped with each other. Cancer diagnosis is one of the most danger problems in the medical domain. Most of researchers have focused try to improve performance and possible to obtain Best and satisfactory results. The correct classification of cancer is an important real-world medical problem. Cancer has become one of the most harmful Death disease around the world and research into Lung cancer diagnosis and treatment has become an one of the most important issue all over the world.

In the modern age of computerized fully automated trend of living, the field of automated diagnostic systems plays an important and vital role. Automated diagnostic system designs in Medical Image processing are one such field where numerous systems are proposed and still many more under conceptual design due explosive growth of the technology today [1]. Lung cancer is considered to be the main cause of cancer death worldwide and it is difficult to detect in its early stages because symptoms appear only in the advanced stages causing the mortality rate to be the highest among all other types of cancer. More people die because of lung cancer than any other types of cancer such as breast, colon, and prostate cancers. There is significant evidence indicating that the early detection of lung cancer will decrease death rate [2].

The proposed algorithm provides Efficient Classification of Lung Tumor based on (M L P) Multi-layer Perceptron neural network approach and tested on the Lung 150 CT scan images comprising of features extracted using (DCT) discrete cosine Transform domain co-efficient .

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

II. RELATED WORK

NAME OF RESEARCHER: FARZAD VASHEGHANI FARAHANI
TOPIC: FUZZY RULE BASED EXPERT SYSTEM FOR DIAGNOSIS OF LUNG CANCER
Date of publication: 2015
Method: the approach is based on a genetic algorithm to tune both the fuzzy rules and fuzzy sets
Outcome of research: method is applied to make a proper decision for each patient.
NAME OF RESEARCHER: EMRE DANDIL
TOPIC: ARTIFICIAL NEURAL NETWORK-BASED CLASSIFICATION SYSTEM FOR LUNG NODULES ON COMPUTED TOMOGRAPHY SCANS
Date of publication: 2014
Method: the system provides classification between benign and malignant nodules with the help of neural networks model of self-organizing maps (som)
Outcome of research: 90.63% accuracy
NAME OF RESEARCHER: PREETI AGGARWAL AND RENU VIG
TOPIC: SEMANTIC AND CONTENT BASED MEDICAL IMAGE RETRIEVAL FOR LUNG CANCER DIAGNOSIS WITH THE INCLUSION OF EXPERT KNOWLEDGE AND PROVEN PATHOLOGY
Date of publication: 2013
Method: content based image retrieval (cbir) and algorithms for detection and classification of nodules.
Outcome of research: achieved an average precision of 88%.
NAME OF RESEARCHER: TADASHI KONDO, JUNJI UENO AND SHOICHIRO TAKAO
TOPIC: MEDICAL IMAGE DIAGNOSIS OF LUNG CANCER BY HYBRID MULTI-LAYERED GMDH-TYPE NEURAL NETWORK USING KNOWLEDGE BASE
Date of publication: 2012
Method: GMDH-type neural network were compared with those of the conventional sigmoid function neural network trained using the back propagation algorithm.
Outcome of research: GMDH-type neural network algorithm was accurate and a useful method for the medical image diagnosis of the lung cancer.

III. PROPOSED ALGORITHM

A. Research methodology :

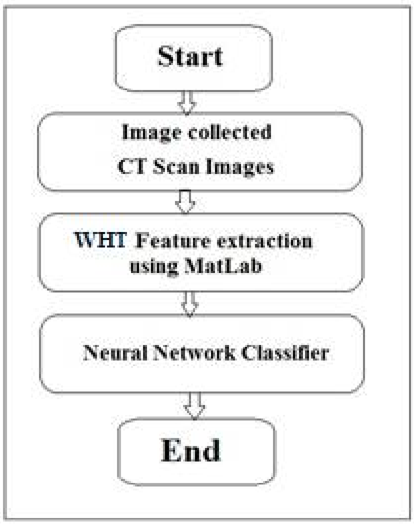

Fig1. Methodology of work

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

It is proposed to study Efficient Classification of Lung Tumor using Neural Classifier. Data acquisition for the proposed classifier designed for the diagnosis of Lung Cancer shall be in the form of CT Scanned images. Image data will be Collected from the different- different hospitals of the country .The most important un correlated features as well as coefficient from the images will be extracted .In order to extract features, statistical techniques, image processing techniques, transformed domain will be used.

1) Images Collected

We have collected the 150 CT-Scan images of lung cancer from the private hospital. By using this CT scan images an algorithm is developed which proposes two-dimensional WHT Transform domain coefficients in addition to Average, Standard Deviation, Entropy, Contrast, Correlation, Energy, Homogeneity total coefficients I get in excel sheet by using MATLAB Code .

2) Feature Extraction

Collected Lung tumor CT Scan images are in .jpg format. By using CT scan images processing & cropping the region of Tumor the 128 features are extracted .

Fig2. Few Samples of input processed CT scan images of lung tumor.
(Above lung images are of Benign, Malignant and Normal types)

Each Lung CT image is represented by a feature vector, F; which is comprised of 128 different parameters. The dataset contains 150 instances (exemplars) for three different classification The classifier based on neural network is trained from the training dataset, where a feature vector is mapped on to a particular class or name of the Lung disease. The neural network learns from data (training exemplars) and the connection weights and biases are estimated as a result of this learning.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

3) Neural Networks

Following Neural Networks are tested:

a) Multilayer perceptron (MLP)
The most common neural network model is the multi layer perceptron (MLP). This type of neural network is known as a supervised network because it requires a desired output in order to learn. The goal of this type of network is to create a model that correctly maps the input to the output using historical data so that the model can be used to produce the output when the desired output is unknown. A graphical representation of an MLP is shown below:

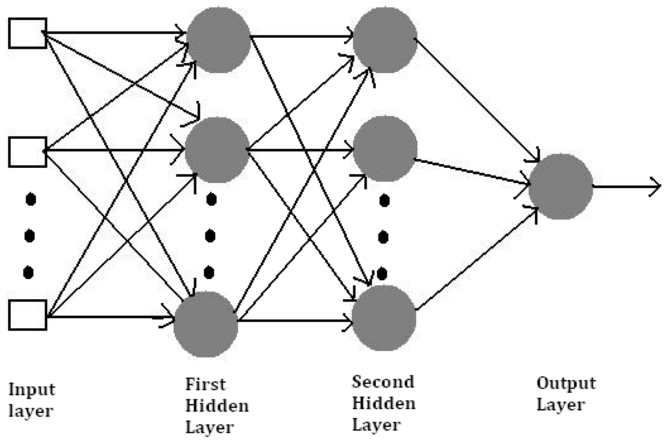

Fig3. The structure of neural network model multi layer perceptron (MLP)

The MLP and many other neural networks learn using an algorithm called back propagation. With back propagation, the input data is repeatedly presented to the neural network. With each presentation the output of the neural network is compared to the desired output and an error is computed. This error is then fed back (back propagated) to the neural network and used to adjust the weights such that the error decreases with each iteration and the neural model gets closer and closer to producing the desired output. This process is known as "training". [10]

- **Learning Rules used**:

  - **Momentum(MOM)**

Momentum simply adds a fraction m of the previous weight update to the current one. The momentum parameter is used to prevent the system from converging to a local minimum or saddle point. A high momentum parameter can also help to increase the speed of convergence of the system. However, setting the momentum parameter too high can create a risk of overshooting the minimum, which can cause the system to become unstable. A momentum coefficient that is too low cannot reliably avoid local minima, and can also slow down the training of the system.

  - **Conjugate Gradient(CG)**

CG is the most popular iterative method for solving large systems of linear equations. CG is effective for systems of the form A=xb-A (1) where x is an unknown vector, b is a known vector, and A is a known, square, symmetric, positive-definite (or positive-indefinite) matrix. These systems arise in many important settings, such as finite difference and finite element methods for solving partial differential equations, structural analysis and circuit analysis.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

- **Quick propagation(QP)**

Quick propagation (Quickprop) [1] is one of the most effective and widely used adaptive learning rules. There is only one global parameter making a significant contribution to the result, the e-parameter. Quick-propagation uses a set of heuristics to optimise Back-propagation, the condition where e is used is when the sign for the current slope and previous slope for the weight is the same.

- **Delta bar Delta(DBD)**

Developed by Widrow and Hoff, the delta rule, also called the Least Mean Square (LMS) method, is one of the most commonly used learning rules. For a given input vector, the output vector is compared to the correct answer. If the difference is zero, no learning takes place; otherwise, the weights are adjusted to reduce this difference. The change in weight from ui to uj is given by: dwij = r* ai * ej, where r is the learning rate, ai represents the activation of ui and ej is the difference between the expected output and the actual output of uj. If the set of input patterns form a linearly independent set then arbitrary associations can be learned using the delta rule.
It has been shown that for networks with linear activation functions and with no hidden units (hidden units are found in networks with more than two layers), the error squared vs. the weight graph is a paraboloid in n-space. Since the proportionality constant is negative, the graph of such a function is concave upward and has a minimum value. The vertex of this paraboloid represents the point where the error is minimized. The weight vector corresponding to this point is then the ideal weight vector. [10]

IV. SIMULATION RESULTS

1) Computer Simulation

The MLP neural network has been simulated for 150 tumor CT Scan images out of which 121 (80% of total images) were used for training purpose and 29 (20% of total images) were used for cross validation.

The simulation of best classifier along with the confusion matrix is shown below :

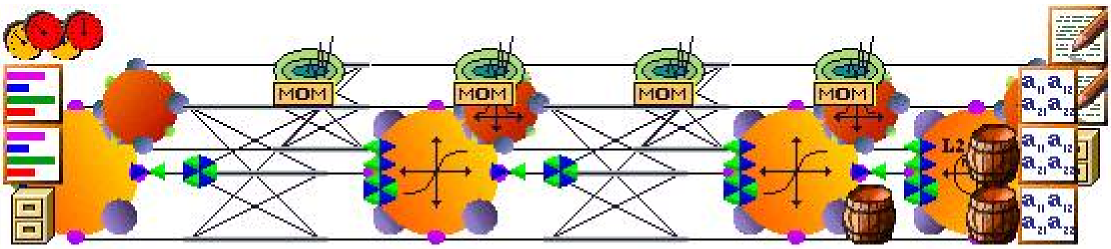

Fig.4 MLP neural network trained with MOM learning rule

2) Results

Output / Desired	NAME(B)	NAME(M)
NAME(B)	9	1
NAME(M)	1	18

Table I. Confusion matrix on CV data set

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

Output / Desired	NAME(B)	NAME(M)
NAME(B)	42	2
NAME(M)	0	77

Table II. Confusion matrix on Training data set

Here Table I and Table II Contend the C.V as well as Training data set. B stand for Benign And M stand For Malignant tumor identification.

Performance	**NAME(B)**	**NAME(M)**
MSE	0.0846608	0.07711079
NMSE	0.374735434	0.341316708
MAE	0.136856549	0.120684907
Min Abs Error	0.005757587	0.00010542
Max Abs Error	1.006658276	1.036843761
R	0.811864928	0.830262335
Percent Correct	90	94.73684211

Table III. Accuracy of the network on CV data set

Performance	**NAME(B)**	**NAME(M)**
MSE	0.010561957	0.01139234
NMSE	0.046605671	0.050269814
MAE	0.049089145	0.055537282
Min Abs Error	0.000103241	0.000101029
Max Abs Error	0.648412545	0.655977876
R	0.977991443	0.976359332
Percent Correct	100	97.46835443

Table IV. Accuracy of the network on training data set

Here Table III and Table IV Contain the Training and C.V result. Table III show the CV result or identify the Benign 90%and Malignant 94.73% and Table IV show the training result or identify the Benign 100 %and Malignant 97.46%.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

Learing Rule	Processing Elements	Training		Cross Validation		Accuracy
		Benign	Malignant.	Benign	Malignant.	
MOM	37	100%	97.46%	90%	94.73%	95.54%

Table V. Multilayer Perceptron Neural Network with Transform Domain : WHT Transform

V. CONCLUSION AND FUTURE WORK

The MLP classifier with MOM learning rule gives best performance of 95.54% .Using our Algorithm, Doctor can Classified lung cancer with enough confidence. Moreover, our Algorithm can also be used by the experts in order to confirm their decision.

VI. ACKNOWLEDGMENT

We are very grateful to our Dr. Bhausaheb Nandurkar College of Engineering and Technology Yavatmal to support and other faculty and associates of ENTC department who are directly & indirectly helped me for these paper.

REFERENCES

[1] Farzad Vasheghani Farahani, M.H. Fazel Zarandi, A. Ahmadi Department of Industrial Engineering Amirkabir University of Technology Tehran, Iran, Fuzzy Rule based Expert System for Diagnosis of Lung Cancer, 978-1-4673-7248-0/15/$31.00 ©2015 IEEE

[2] Joey Mark Diaz, Raymond Christopher Pinon,“ Lung Cancer Classification Using Genetic Algorithm to Optimize Prediction Models”,2014 IEEE.

[3] Emre DandÕl1,3, Murat ÇakÕro÷lu2, Ziya Ekúi3, Murat Özkan3,4, Özlem Kar Kurt5, Arul Canan6, “Artificial Neural Network-Based Classification System for Lung Nodules on Computed Tomography Scans”, International Conference of Soft Computing and Pattern Recognition 2014 IEEE.

[4] Preeti Aggarwal and Renu Vig, “ Semantic and Content-Based Medical Image Retrieval for Lung Cancer Diagnosis with the Inclusion of Expert Knowledge and Proven Pathology”, Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013).

[5] Naoufel Werghi, Fatma Taher and Hussain Al-Ahmad, Bayesian Classification and Artificial Neural Network Methods for Lung Cancer Early Diagnosis, 978-1-4673-1260-8/12/$31.00 ©2012 IEEE.

[6] Tadashi Kondo, Junji Ueno and Shoichiro Takao, “Medical Image Diagnosis of Lung Cancer by Hybrid Multilayered GMDH-type Neural Network Using Knowledge Base”, Proceedings of 2012 ICME International Conference on Complex Medical Engineering July I - 4, Kobe, Japan

[7] Mhd Saeed Sharif, Maysam Abbod, PET Volume Analysis Based On Committee Machine For Tumour Detection And Quantification, 978-0-7695-4593-6/11 $26.00 © 2011 IEEE DOI 10.1109/DeSE

[8] Fatma Taher and Rachid Sammouda, “Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods”, 2011 IEEE GCC Conference and Exhibition (GCC), February 19-22, 2011, Dubai, United

[9] Rachid Sammouda, Fatma Taher, Morphology analysis of sputum color imahes for early lung cancer diagnosis, 978-1-4244-7167-6/10/$26.00 ©2010 IEEE

[10] Jianwei Ma, Lu Liu, Bin Zhang, and Junli Yan, “Lung Cancer Metastases and Non-Metastases Tumid Lymph Nodes Classification in CT Image with Multi-resolution Histogram”, International Conference on Intelligent Control and Information Processing August 13-15, 2010 - Dalian, China. Arab Emirates.

[11] Sandrine Tomei, Simon Marache-Francisco, Christophe Odet, Carole Lartizien, “Automatic detection of activenodules in 3D PET oncology imaging using the Hotelling Observer and the Support Vector Machines: a comparison study”, 2008 IEEE Nuclear Science Symposium Conference Record.

[12] JIA Tong, “Computer-Aided Lung Nodule Detection Based on CT Images”, 2007 IEEE/ICME International Conference on Complex Medical En2ineerin2.

[13] Kenji Suzuki, “Computer-Aided Diagnostic Scheme for Distinction Between Benign and Malignant Nodules in Thoracic Low-Dose CT by Use of Massive Training Artificial Neural Network”, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 24, NO. 9, SEPTEMBER 2005.

[14] Manuel G. Penedo,* Mar´ıa J. Carreira, Antonio Mosquera, and Diego Cabello, Associate Member, IEEE, Computer-Aided Diagnosis: A Neural-Network-Based Approach to Lung Nodule Detection, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 17, NO. 6, DECEMBER 1998

[15] Jyh-Shyan Lin,* Member, IEEE, Shih-Chung B. Lo, Member, IEEE, Akira Hasegawa, Matthew T. Freedman, and Seong K. Mun, Member, IEEE, Reduction of False Positives in LungNodule Detection Using a Two-Level Neural Classification, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 15, NO. 2, APRIL 1996

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

[16] Guruprasad Bhat, Vidyadevi G Biradar , H Sarojadevi Nalini, " Artificial Neural Network based Cancer Cell Classification (ANN – C3)", Computer Engineering and Intelligent Systems, Vol 3, No.2, 2012.
[17] Almas Pathan, Bairu.K.saptalkar, "Detection and Classification of Lung Cancer Using Artificial Neural Network", International Journal on Advanced Computer Engineering and Communication Technology Vol-1 Issue:1.
[18] V.Krishnaiah, Dr.G.Narsimha, Dr.N.Subhash Chandra;" Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques"; Vol. 4 (1), 2013, 39 –45.
[19] Ismail SARITAS, Novruz ALLAHVERDI and Ibrahim Unal SERT;" A Fuzzy Expert System Design for Diagnosis of Prostate Cancer"; International Conference on Computer Systems and Technologies - CompSysTech'2003.
[20] K.Balachandran, Dr. R.Anitha;" Supervised Learning Processing Techniques for Pre-Diagnosis of Lung Cancer Disease"; International Journal of Computer Applications (0975 – 8887) Volume 1 No. 4- 2011.
[21] Fix, E. and Hodges, J. L. (1951). Discriminatory analysis—"Nonparametric discrimination": Consistency properties. Technical Report 4, Project no. 21-29-004, USAF School of Aviation Medicine, Randolph Field, Texas.
[22]T.M. Cover, and P.E. Hart, "Nearest neighbor pattern classification", IEEE Transactions on Information Theory, 13, pp. 21–27, 1967.
[23] Manish Sarkar and Tze-Yun Leong;"Application of K- Nearest Neighbors Algorithm on Breast Cancer Diagnosis Problem". International Journal on Advanced Computer Engineering and Communication Technology Vol-1 Issue:1
[24] A. Marcano-Cedeno, J. Quintanilla-Dominguez, D. Andina;" WBCD breast cancer database classification applying artificial metaplasticity neural network"; Expert Systems with Applications 38 (2011) 9573–9579.
[25] Hu Y.H., Hwang Jenq-Neng, Handbook of Neural Networks Signal Processing, New York, CRC Press, 2002.