

# Intelligent Prediction of Lung Cancer Via MRI Images using Morphological Neural Network Analysis

T. Shravya<sup>1</sup>, T. Rajesh<sup>2</sup>

<sup>1</sup>PG Scholar, Dept. of CSE, G Narayanamma Institute of Technology & Science, Hyderabad, India.

<sup>2</sup>Asst. Professor, Dept. of CSE, G Narayanamma Institute of Technology & Science, Hyderabad, India.

\*\*\*

**Abstract:** The main objective of this project is to assist clinicians to efficiently identify the Lung Cancer via forecasting analysis by using Morphological Neural Network (MNN) Classification with Image Pruning Methodology. In image processing domain, the complex task to analyze and rectify is cancer estimation and prediction. When compared to tumor estimations, cancer estimations are more complex because these are purely cell-based paradigms and visually not clear to analyze, especially this system focus on lung cancer and its strategies. A new methodology is required to classify these kind of cancer cells and hence Morphological Neural Network Classifier with Image Pruning Scheme is introduced to efficiently classify the cancer cells and mark out the affected region more efficiently. Lung cancer is one of the prevalent diseases in human and can be diagnosed using several tests that include CT scans, MRI scans, biopsy and so on. Over the years, the use of machine learning and artificial intelligence techniques has transformed the process of diagnosing lung cancer. However, the accurate classification of cancer cells is still a medical challenge faced by researchers. Difficulties are routinely encountered in the search for sets of features that provide adequate distinctiveness required for classifying breast tissues into groups of normal and abnormal. Therefore, the aim of this approach is to prove that the MNN algorithm is more efficient for diagnosis, prognosis and prediction of lung abnormality, which is basically derived from two classical algorithms called Morphological Image Processing and Artificial Neural Network (ANN) with Image Pruning.

**Keywords:** lung cancer, image pruning, morphological neural network, image binarization and segmentation.

## 1. INTRODUCTION

Due to large prevalence of smoking and air pollution around the world, lung cancer has emerged as a common and deadly disease in recent decades. It often takes long time to develop and most people are diagnosed with the disease within the age bracket 55 to 65. Early identification and treatment is the best available option for the infected people. Reliable identification and classification of lung cancer requires pathological test, namely, needle biopsy specimen and analysis by experienced pathologists. However, because it involves human judgment of several factors and a blend of incidents, a decision support system is desirable in this case.

Recent developments in image processing, pattern recognition, dimensionality reduction and classification methods has paved the way for alternate identification and classification approaches for lung cancer. A number of machine learning (ML) algorithms including artificial neural network, support vector machine, discriminant analysis, decision trees and ensemble method tells about the classification of lung cancers through image processing and pathological identifiers.

In addition to these ML approaches, deep learning through restricted Boltzmann storage device in the form of autoencoders has shown promising success in classification tasks in different domain including acoustics, sentiment classification, and image and text recognition. Inspired by the positive outcome of deep learning in relevant fields, a deep learning-based classification method is investigated in this work. The endowment to this work is of two folds. Firstly, the proposed methods can surpass the existing works on small dataset for lung cancer classification.

Secondly, architecture of deep auto-encoder network for lung cancer classification is proposed which outperforms other methods and also show that the performance improvement is statistically significant. The most common cancers that lead to death are lung, prostate, breast and colon cancer. They represent 46% of all deaths due to cancer and pulmonary tumor is in charge for more than a quarter (27%) of all cancers. In developed countries, lung cancer patients have a 10 up to 16% chance of having a five-year survival rate. Nevertheless, early detection of lung cancer, through computed tomography (CT) can help in improving the patient's survival rate, insomuch that the five-year survival rate increases to 70%.

The medical specialist first identifies the pulmonary nodules from a CT or MRI scan, and then makes a possible prediction dependent on the nodule morphology assessment including the clinical context. However, he often has to analyze a huge number of nodules and make a prognosis quickly, and such tasks become burdensome under these circumstances. The author

used 33 cases, of which 14 were malignant and 19 were benign. The highest accuracy achieved was 90:91%. The Multi-Crop Convolutional Neural Network (MCCNN) model was introduced to categorize the pulmonary nodules. The authors have used eight hundred and eighty benign nodules and four hundred and ninety-five malignant nodules from the LIDC/IDRI dataset and obtained an accuracy of 87.14%.

Features retrieved from an auto-encoder were utilized to classify pulmonary nodules into either malignant or benign, using 4303 instances containing four thousand three hundred and twenty-three nodules from the dataset. This work focuses on classification of the nodules into malignancy levels using the feature extractor Structural Co-occurrence Matrix and, a well-known classifier, so as to support the radiologist in analyzing the nodules. The efficiency of the multilayer perception (MLP), the support vector machine and the k- nearest neighbors(k-NN) algorithms were compared to decide which classifier works better with the proposed SCM approach.

## 1.1 SYSTEM CONTRIBUTIONS

The main objective of the proposed system is to identify the pulmonary tumor from the MRI images. The main intend is to analysis of quantitative procedures and to retrieve the features of image using statistical and geometrical natures. In the past many approaches all researchers are using Bayesian Network Algorithm. In the proposed scenario we are using the Morphological Neural Network Classifier with Image Pruning Scheme, which improves the performance of the system.

## 2. LITERATURE REVIEW

In software development process, literature survey is the most necessary step in order to determine the economy, time factor, and company strength before developing the tool. The required needs for developing the project are found out completely only after the survey.

### 2.1 CT screening for lung cancer: five-year prospective experience -S. Swensen, et al,2005

**Purpose:** To report results of a 5-year prospective low-dose CT scan study for lung cancer.

**Materials and Methods:** After informed written consent was obtained, 1520 individuals were enrolled. Participants were aged 50 years and older and had smoked for more than 20 pack-years.

**Results:** In 788 (52%) men and 732 (48%) women, 61% (927 of 1520) were current smokers, and 39% were former smokers. Forty-eight participants died of various causes since enrollment.

**Conclusion:** CT allows detection of early-stage lung cancers. Benign nodule detection rate is high. Results suggest no stage shift.

### 2.2 Computed tomography screening and lung cancer outcomes - P. Bach,et al., 2007

**Purpose:** The main aim is screening and monitoring the pulmonary tumor outcomes.

**Materials and Methods:** Screening using low-dose computed tomography (CT) represents a unique development in the struggle in order to enhance the outcomes for people with lung cancer.

**Results:** One of the most important issues confronting those who wish to consider implementation of LDCT screening in high-risk populations is the issue with high rate of positive examinations, primarily pulmonary nodules.

**Conclusion:** This review talks about the logic behind screening, the results of on-going trials, potential harms of screening and current knowledge gaps.

### 2.3 Prognostic implications of cell cycle, apoptosis, and angiogenesis biomarkers in non-small cell lung cancer: a review - Singhal S1, Vachani A, Antin-Ozerkis D, 2005

**Purpose:** The main purpose is to identify the cell cycle affection over cancer attack patients.

**Materials and Methods:** Hundreds of papers have appeared over the past several decades proposing a variety of molecular markers or proteins that may have prognostic significance in non-small cell lung cancer. Although no single marker has yet been shown to be apt with the prediction of a patient's outcome, a profile on basis of the best of these markers may prove

useful in directing patient therapy.

**Conclusion:** This review analyzes the huge and most accurate of these studies with the aim of compiling the most important prognostic markers in early stage non-small cell lung cancer. In this review, the key focus is on biomarkers particularly involved in one of three major pathways: cell cycle regulation, apoptosis, and angiogenesis.

#### 2.4 Screening for epidermal growth factor in lung cancer-R. Rosell, et al., 2009

**Background:** We evaluated the feasibility of large-scale screening for EGFR mutations in patients with advanced non-small-cell lung cancer and analyzed the association between the mutations and the outcome of erlotinib treatment.

**Methods:** From April 2005 through November 2008, lung cancers from 2105 patients in 129 institutions in Spain were screened for EGFR mutations and the analysis was carried out in a central laboratory. Patients with tumors holding EGFR mutations were eligible for erlotinib treatment.

**Results:** EGFR mutations were identified in 350 of 2105 patients (16.6%). Mutations were observed repeatedly in women (69.7%), in patients who have never smoked (66.6%), and in those with adenocarcinomas (80.9%) ( $P < 0.001$  for all comparisons).

**Conclusion:** Large-scale screening of patients with lung cancer for EGFR mutations is achievable and can play a role in decision making about treatment.

### 3. EXISTING SYSTEM

In the past system researchers faced lots of struggle during the evolution of a unified approach that can be put to application to all types of medical images and applications such as lung cancer estimation and tumor analysis. A difficult problem here is the selection of a suitable technique for a specific kind of medical image. Thus, there is no universal accepted method for medical image segmentation process. Hence, it remains as a demanding and testing problem in image processing industry and computer vision fields. Fuzzy C-means, which is the most desired and popular method to estimate tumor and cancer cells has been implemented but the predictions falls into certain prescribed level not for a global mean, hence leading to a conclusion that the approach is yet a problem to developers. Therefore, researchers/developers decide to combine two different algorithms to work out the result for cancer cell estimation and predictions. Also, the existing system used SVM for classification.

### 4. PROPOSED SYSTEM

The proposed approach comprises three main stages: preprocessing, feature extraction and classification. The preprocessing step includes methods that help in noise removal, enhancement and segmentation steps of the MRI scanned images. The feature extraction stage entails five steps: wavelet decomposition, wavelet coefficient extraction, normalization, energy computation, and coefficient reduction. Lastly, the classification stage involves the combinatorial use of Morphological Image Processing and Artificial Neural Network, called 'Morphological Neural Network with Image Pruning', which is used to classify lung tissue into normal and abnormal.

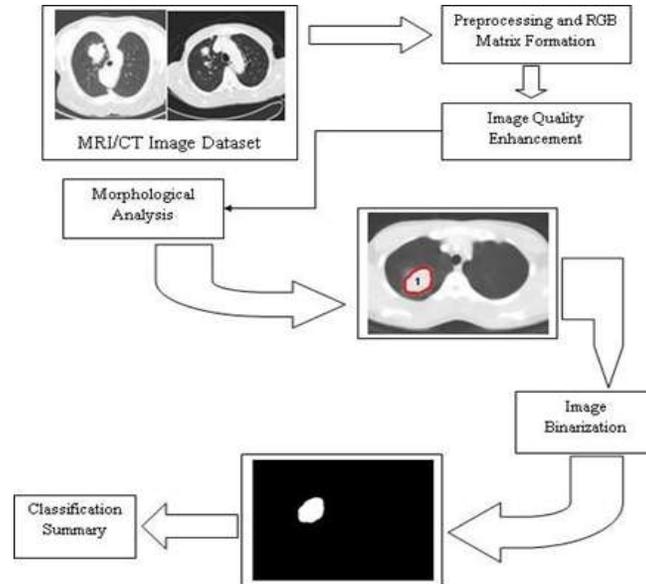


Fig-1: System Architecture

Computer programs or software created based on the human intellect could be helpful in assisting the doctors in decision making without conferring with specialists directly. The software was not developed to substitute the specialist or doctor, but to aid in the diagnosis and prediction of patient’s condition from specific regulations or "experience". Patients with high-risk factors or symptoms or predisposed to specific diseases or any illness, can be preferred to see a specialist for more treatment. Utilizing the technology particularly Artificial Intelligence (AI) techniques in medical applications could lower the cost, time involved, human proficiency and medical inaccuracies. For all our proposed approach aims to provide the complete solution to the medical image processing strategy with the help lung cancer estimation and prediction scenarios.

### 5. METHODOLOGY

The proposed methodology consists of following modules:

1. Image Acquisition
2. Image Processing
3. Image Binarization and Morphological Analysis
4. ROI Selection and Segmentation
5. Feature Extraction
6. Classification using Neural Network

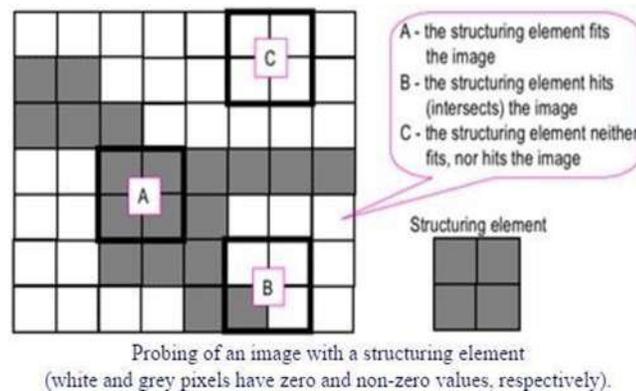
**Image Acquisition:** The major intent of this module is to obtain the image and convert it into matrix format. For any vision system, the image acquisition stage would be the first step. In image processing, image acquisition could be widely defined as the activity of extracting an image from any source, specifically a hardware-based source, such that it could be passed through the processes that might occur later on. The ideal goals of this process is to have a source of input that would operate within a controlled and measured guidelines that the same image can, if necessary, be nearly perfectly reproduced under the same conditions so that anomalous factors are easier to locate and eliminate.

**Image Preprocessing:** We need an image processing model in order to enhance the image quality. This step can remarkably improve the reliability of an optical examination. Various filter operations that escalate or minimize few definite image details enable an easier or faster assessment. Edge filters, soft focus, selective focus, user - defined filter, image plane

separation, static or dynamic binarization and binning are few of the image pre-processing methods. This module comprises many internal works that include various functions for image processing like contrast increase by static or dynamic binarization, resolution reduction via binning, look-up tables or image plane separation, image rotation and conversion of color images to gray value images.

**Image Binarization and Morphological Analysis:** Binarization is a method that converts a pixel image to a binary image. A major contribution of this module is to convert the gray image into binary and remove the unwanted contents of lung part.

The images were initially classified into normal and abnormal. Based on the theory that abnormal cases involve microcalcifications and masses whether malignant or benign, 207 images were identified to be normal cases while the remaining 115 images have uncommonness. The neural network was trained using 80% of both types and tested with the remaining 20%. This strategy has been performed 4 times repeatedly for each level of decomposition, from level 2 to level 5.



**Fig-2: Image Binarization**

For each level, features extracted using the reduction procedure were utilized where binary images might have innumerable uncommonness. Morphological image processing aims at removing these imperfections by taking the structure and form of the image into account.

**ROI Selection and Segmentation:** The images were initially classified into normal and abnormal. Lung image has two parts i.e., the left lung and right lung respectively. Image preprocessing techniques are essential for discovering the orientation of the mammogram, noise removal and quality improvement of the image. Prior to the implementation of image processing algorithm on mammogram, preprocessing steps are very significant to focus the identification of uncommonness without needless influence from background of the mammogram. For this study, the first phase of preprocessing entailed two procedures, noise elimination and image contrast enhancement. The second phase involves segmentation, which is executed to remove the background area (high intensity rectangular label, tape artifact, and noise). The third deals with the application of phase reduction and global gray level thresholding to extract region of interest (ROI).

**Feature Extraction:** The major goal of feature extraction module is to retrieve the key data from the image that is the statistical and geometric features like mean, brightness, skewness, kurtosis, area, perimeter. The main role of feature extraction is to acquire new variables from the matrix of the image to create distinct classes. The feature extraction procedure includes the use of the wavelet decomposition process. These features are then classified in the next stage. Features extraction comprises five processing steps by which the features in the system are extracted from the wavelet coefficients.

**Classification using Neural Network:** After obtaining the features, we need to classify with the trained data. An artificial neuron network (ANN) is a computational model developed on basis of the arrangement and functions of biological neural networks. ANN is a well-ordered system of basic processing elements, nodes or units, whose functions are to some extent derived from the animal neural cell. The processing capability of the network is controlled by the inter unit connection weights or strengths achieved based on learning from, or adaptation to a set of training patterns. The most essential elements of ANN's are modeled after the configuration of the human lungs.

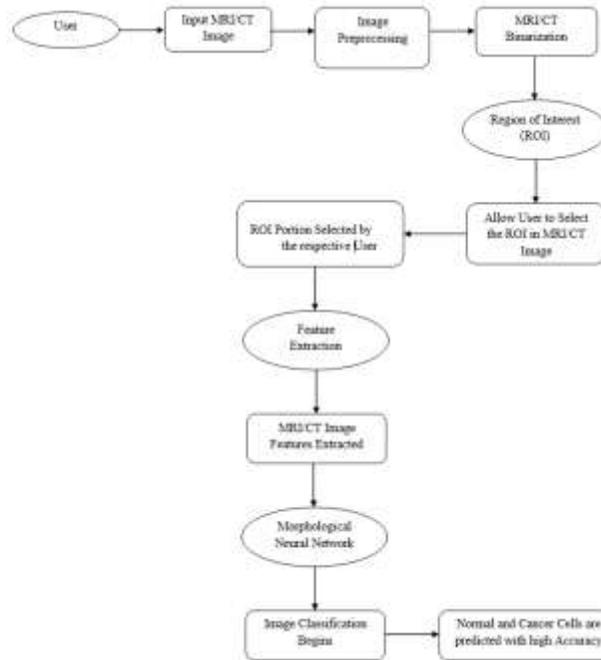


Fig-3: Modules Flow

## 6. RESULTS

In the existing scheme, 8 combinations of SCM and classifier were tested, each one using a filter and a specific input image type. Mean, Gaussian, Laplacian and Sobel filters were applied with GLCM, LBP, CM and SM extractors. Sobel filters were applied with GLCM, LBP, CM and SM extractors. The system proposed not only allows the user to give the input but also lets them choose the region of interest where the nodules are to be checked by going through various preprocessing steps and filtrations. The morphological neural network not only determines if the cancer is malignant or benign but also provides the accuracy rate on the basis of cell vector distance.



Fig-4: User Interface for lung cancer assessment

**Table-1:** Comparison with existing scheme

| Extractor   | Classifier | Accuracy |
|-------------|------------|----------|
| SCM Mean GU | MLP        | 95.4%    |
|             | SVM        | 95.7%    |
|             | k-NN       | 95.3%    |
| SCM Mean GU | MNN        | 98.2%    |

The highest accuracy achieved here is 98% when compared to other implementations so far. Also, morphological neural network with combination of any other deep learning methodology can be implemented in future in order to achieve an accuracy rate of 99.5%.

## 7. CONCLUSION

This work presented a new approach to classify pulmonary nodules into benign and malignant nodules. This approach applied the proposed extractor combined with a classifier and so as to identify the best SCM approach, 8 combinations of SCM and classifier were tested, each one using a filter and a specific input image type. The filters used were: Mean, Gaussian, Laplacian and Sobel, while the input images used were Grayscale units. After analyzing all the combinations on LIDC/IDRI dataset, the best results for nodule classification into malignant and benign were acquired by the proposed extractor composed with the Mean filter and applied to the images combined with the MNN classifier. The proposed extractor was also used in the classification of nodule malignancy levels for the 5-class classification problem so as to determine the capability of our proposed approach. The obtained results validated our proposed extractor approach in handling the malignant classification of pulmonary nodules, and therefore can be helpful in assistance for making medical diagnoses more precise and faster in CAD systems. The future works should be directed at improving the proposed extractor in order to classify the nodules into malignancy levels through observations by implementing various other alternatives.

## REFERENCES

- [1] F. Han et al., "A texture feature analysis for diagnosis of pulmonary nodules using LIDC-IDRI database," in Proc. IEEE Int. Conf. Med. Imag. Phys. Eng. (ICMIPE), Oct. 2013, pp.1418.
- [2] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in Proc. 24th Int. Conf. Inf. Process. Med. Imag., Jun. / Jul. 2015, pp.588599.
- [3] P.P. Rebouças Filho, R. M. Sarmiento, G. B. Holanda, and D. de AlencarLima, "New approach to detect and classify stroke in skull CT images via analysis of brain tissue densities," Comput. Methods Programs Biomed., vol. 148, pp. 2743, Sep. 2017.
- [4] H. Krewer et al., "Effect of texture features in computer aided diagnosis of pulmonary nodules in low-dose computed tomography," in Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC), Oct. 2013, pp.38873891.
- [5] G.L.B. Ramalho, D.S. Ferreira, P.P.R.Filho, and F.N.S.deMedeiros, "Rotation-invariant feature extraction using a structural co-occurrence matrix," Measurement, vol. 94, pp. 406415, Dec. 2016.
- [6] P. P. Rebouças Filho, P. C. Cortez, A. C. da Silva Barros, V. H. C. Albuquerque, and J. M. R. Tavares, "Novel and powerful 3D adaptivecrispactive contour method applied in the segmentation of CT lung images," Med. Image Anal., vol. 35, pp. 503516, Jan. 2017.
- [7] P. P. Rebouças Filho et al., "Automated recognition of lung diseases in CT images based on the optimum-path forest classier," in Neural Computing and Applications. London, U.K.: Springer, 2017, pp. 114. [Online]. Available: <https://link.springer.com/article/10.1007/s00521-017-3048-y#citeas>
- [8] P. P. Rebouc, R. M. Sarmiento, P. C. Cortez, and V. H. C. De, "Adaptive crisp active contour method for segmentation and reconstruction of 3D lung structures," Int. J. Comput. Appl., vol. 111, no. 4, pp. 18,2015.

- [9] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," IEEE Trans. Syst., Man, Cybern., vol. SMC-3, no. 6, pp. 610621, Nov.1973.
- [10] T. Ojala, M. Pietikäinen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," in Proc. IAPR Int. Conf. Pattern Recognit. (ICPR), vol. 1. 1994, pp. 582585.
- [11] L. Wang and D.-C. He, "Texture classification using texture spectrum," Pattern Recognit., vol. 23, no. 8, pp. 905910, 1990.
- [12] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971987, Jul.2002.
- [13] M.-K. Hu, "Visual pattern recognition by moment invariants," IRE Trans. Inf. Theory, vol. 8, no. 2, pp. 179187, Feb.1962.
- [14] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification. Hoboken, NY, USA: Wiley, 2012.
- [15] S. S. Haykin, Neural Networks and Learning Machines. EnglewoodCliffs, NJ, USA: Prentice-Hall, 2008, vol.3.

### **BIOGRAPHIES**

T. Shravya is pursuing her M. Tech in Department of computer science and engineering in G. Narayanamma institute of technology and science, Hyderabad, Telangana and has completed her B. Tech from Stanley college of engineering and technology for women, Abids, Hyderabad.

T. Rajesh is current working as Asst. Professor Department of computer science and engineering in G Narayanamma institute of technology and science, Hyderabad, Telangana. His interest includes Data Mining.