Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, No. 3, December : 2023

CLASSIFICATION AND LOCALISATION OF LUNG CANCER USING AUTO ENCODER BASED DEEP LEARNING

Dr. K. Chaitanya, Assistant Professor, Dept. Of CSE, ANU College of Engineering and Technology, Acharya Nagarjuna University
M. Bheemeswara Swamy, M.Tech student, Dept. Of CSE, ANU College of Engineering and Technology, Acharya Nagarjuna University

Abstract
Lung cancer is a prevalent and deadly disease, necessitating early detection and precise staging to improve patient outcomes. While traditional diagnostic methods rely on medical imaging and expert interpretation, the advent of deep learning, particularly convolutional neural networks (CNNs) and auto encoders, offers an opportunity to revolutionize the diagnostic process. To effectively treat lung cancer, screening computed tomography (CT) images for pulmonary nodules is a crucial first step. In the past few years, machine learning has become increasingly popular for the detection, prediction, and classification of disease, particularly for difficult tasks like the identification and detection of lung cancer. The field of computer vision research has witnessed a significant transformation with the rise in popularity of deep learning (DL). *In this paper, we determine the improved classification accuracy over previous methods by employing a DL for lung cancer classification. In this paper, for detecting noncancerous and cancerous lung nodules using CT-based lung cancer image dataset consortiums (LIDC) deep learning based auto encoders are used.* We used MATLAB software to get improved results and better accuracy. Auto encoders more advance in deep learning as they extract features by own from images and form SoftMax layers. Performance analysis of auto encoder based deep learning model gives superior results by both subjective and objective analysis.

Keywords: Computed Tomography (CT) image, DL, Lung Cancer, deep learning based Auto encoder, softmax Layer.

Introduction
To predict and expose the various diseases in humans with the help of computational science, a lot of work has been done. The aim for Lung Cancer Detection & Stages Identification using Deep Learning is to develop a sophisticated and accurate model that can effectively detect the presence of lung cancer and identify the stage of the disease based on medical imaging data, such as X-rays or CT scans. The ultimate goal of this project is to significantly enhance the accuracy and efficiency of lung cancer diagnosis and staging, leading to earlier intervention and better treatment outcomes for individuals affected by lung cancer. Deep learning models have the potential to augment the capabilities of healthcare providers in the fight against this deadly disease. Lung cancer remains a pressing global health concern, with significant morbidity and mortality rates. Timely and precise diagnosis is a pivotal factor in enhancing patient outcomes and survival rates. Traditionally, the diagnosis of lung cancer has relied on medical imaging, such as X-rays and computed tomography (CT) scans, coupled with manual interpretation by expert radiologists and clinicians. Lung cancer continues to be a global health concern, representing a leading cause of cancer-related mortality worldwide. Early detection and accurate staging of lung cancer play pivotal roles in determining treatment strategies and significantly impacting patient outcomes. When compared to other cancers, lung cancer is known to cause an excessive number of deaths worldwide each year, making it one of the major causes of the rising death rate. This fatal disease is affecting both men and women. The significance of this paper lies in its potential to revolutionize the lung cancer diagnostic process. By reducing the scope for human error and subjectivity, we aspire to provide healthcare professionals with a robust tool that enables early intervention and improved patient outcomes. As we proceed through the paper, we will explore the methodology, data collection and pre-processing, model development,

model validation, ethical considerations, and the prospects of clinical validation and integration into existing healthcare systems. Deep learning algorithms, particularly convolutional neural networks (CNNs), have demonstrated remarkable potential in automating image analysis and diagnosis. This technological advancement has spurred research endeavors aimed at harnessing the power of deep learning to improve the detection and staging of lung cancer. In our application of lung cancer detection autoencoder based deep learning technique is introduced. Two layers of autoencoder are efficient to get fine features from lung CT scan image. The features then forms softmax layers with different feature layers. This paper significance lies in its potential to revolutionize the way lung cancer is diagnosed and staged, ultimately improving patient outcomes and increasing the efficiency of healthcare systems. In the paper, we will delve into the methodology, data collection and pre-processing, model development, validation, and ethical considerations. Furthermore, we will explore the prospects of clinical validation and the integration of the deep learning system into existing healthcare infrastructure.

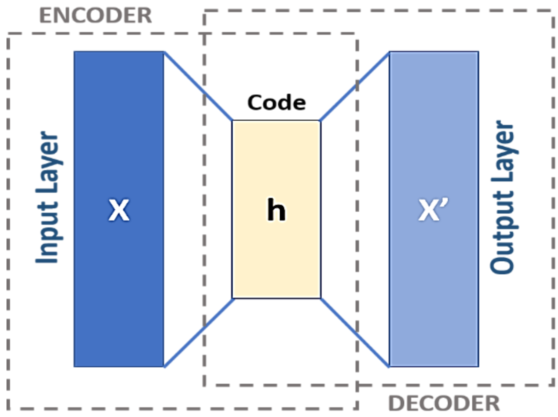

Fig.1.1 Deep Learning structure with encoder and decoder

Deep learning structure has input layer, encoder , h (channel) or structure , decoder and output layer. Autoencoder are advance branch in deep learning. Autoencoders are used with softmax layer to save features in stack format. Then prediction is applied which compares the stack data with test features. Softmax layer is the stack format of features.

Literature

For almost twenty years, researchers have been working in this field of lung cancer extensively. However, because they may need additional image processing modules, like feature extraction, lung nodule segmentation, and computed tomography image transformation, to build a whole CADe system, earlier computer-aided detection (CADe) schemes were time-consuming and intricate. These schemes find it challenging to handle and evaluate huge amounts of data when the number of medical images keeps rising. Furthermore, some cutting-edge deep learning schemes might have strict database standards. This work suggests a useful method for detecting lung nodules based on multi group patches extracted from lung images and improved by the Frangi filter. A four-channel convolution neural network model is designed to learn radiologists' knowledge for four-level nodule detection by combining two sets of images. The outcomes show how effective the multi group patch-based learning system is at enhancing lung nodule detection performance and significantly lowering false positives across a large volume of image data. [1]

In lung disease, the pulmonary nodule is common. The extensive use of CT technology has greatly improved doctors' diagnostic efficiency. This paper proposes a lung segmentation method based on morphology and image area size statistics, effectively eliminating the influence of the trachea on pulmonary parenchyma image segmentation. We propose a technique for extracting regions of interest (ROIs) based on circular filters and morphology that aims to preserve the integrity of the ROI form while lowering the number of false positives. Using a convolutional neural network, we have successfully developed a dependable compute-assisted diagnosis application for lung nodules on CT images. [2]

In helical computed tomography (CT) scans of the thorax we developed a fully automated computerised method for the detection of lung nodules. On the basis of these methods, 3D and 2D analyses can acquire image data at the time of diagnostic CT scans. To create a segmented lung volume for additional analysis, lung segmentation is carried out section by section. Contiguous three-dimensional structures within each thresholded lung volume are identified using an 18-point connectivity scheme, and those which satisfy a volume criterion are chosen as initial lung nodule candidates. For every nodule candidate, grey-level and morphological features are calculated. [3]

In clinical practice, the traditional chest radiograph is still widely used and likely will be for some time to come. Interpretation is notoriously challenging for these. This explains why CAD for chest radiography continues to be of interest. This survey aims to classify and provide a brief overview of the over 150 papers that have been published in the last 30 years on computer analysis of chest images. Future research directions are provided, along with a list of outstanding challenges. [4]

An essential component of image processing is image segmentation. In order to diagnose unusual diseases, segmentation is widely used in the field of medical imaging. One can manually segment the same medical images. However, when compared to manual segmentation, the accuracy of image segmentation using segmentation algorithms is higher. A wide range of imaging modalities, including computed tomography (CT), radiography, and magnetic resonance imaging (MRI), are currently available for use in medical diagnosis. For the majority of subsequent image analysis tasks, medical image segmentation is a necessary first step. While the original FCM algorithm does a good job of segmenting images free of noise, it is not able to segment images that have been tainted by noise, outliers, or other imaging artefacts. This paper presents a method for segmenting images using the Modified Fuzzy C-Means (FCM) algorithm and the Fuzzy Possibility C-Means (FPCM) algorithm. [5]

Lung cancer is a serious disease that affects humans. This study proposes a method for detecting abnormal lung tissue growth using artificial neural networks. Manual result interpretation is incapable of avoiding misdiagnoses. In this study, we can determine whether the lung images are healthy or cancerous. For the different CT scanning system views, data bases have also been created, such as sagittal, coronal, and axial. For the classification of normal images, a neural network based on the textural characteristics of the images allows for distinguishing them from malignant ones. For cancer detection to address this issue, Google Net and CNN deep learning algorithms have been proposed. The VGG-16 architecture serves as the foundation for both the classifier network and the region proposal network. In detection and classification, the algorithm achieves a precision of 98%. [6]

From above references it is observed that in existing different machine learning techniques such as support vector machine, k-nearest neighbour, random forest classifiers are used and other thresholding techniques are used for prediction of lung cancer from lung CT images. The techniques used in existing are not efficient and not giving higher accuracy so in proposed model there is implementation of autoencoder based deep learning technique for lung cancer identification and stages classification.

Proposed Method

Proposed method is having more superior performance than state of art techniques. The flowchart of the proposed method for lung cancer identification is shown in the figure. There are six steps in methodology.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, No. 3, December : 2023

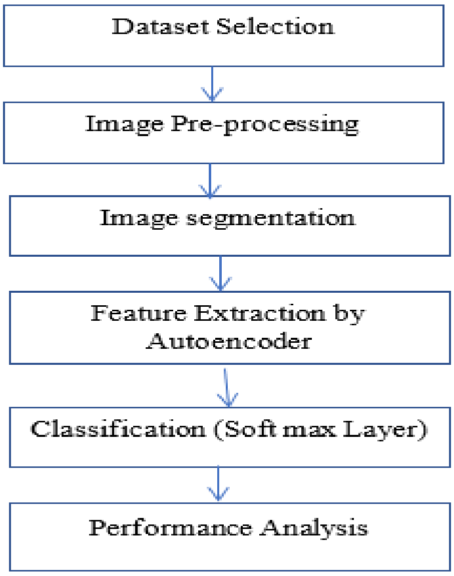

Fig.1. Block Diagram for Proposed Method

Following are the steps for proposed autoencoder based deep learning method for lung cancer detection from CT scan images.

1. Load the sample dataset of lung CT images
2. Apply auto encoder for feature extraction
3. Second auto encoder is used for deep feature extraction.
4. Save the autoencoder features in the SoftMax layer
5. Stack of soft max layer to form a deep network which is further used for prediction of lung cancer
6. Performance analysis of proposed model

Above steps are explained in detail below,

Load the sample dataset of Lung CT images

We collected a comprehensive dataset of medical images, including CT scans, coupled with patient data. This dataset was meticulously pre-processed, ensuring consistency and high-quality input for our deep learning model. In proposed model we have selected CT scan of lung as CT scan has more advantages over other modalities for lung scan. It gives clearer picture of lung and different fluids.

Train an auto encoder with features of Lung CT image

In MATLAB, we designed, trained, and fine-tuned a powerful deep learning model, specifically an autoencoder. The model was optimized for image analysis, enabling it to recognize the subtle patterns and features indicative of lung cancer. There are two autoencoders are used in proposed model for feature extraction of the lung CT image.

Softmax Layer

The SoftMax layer is stack of autoencoder features. The features with differences will form a new layer. In proposed model we have normal and abnormal images so two layers will get form. One is for normal lung Ct scan image features and another us for abnormal lung Ct scan image features.

Train the deep network

Our deep learning model was extended to not only detect lung cancer but also classify it into various stages. This important aspect of the project has the potential to significantly impact treatment decisions and patient care. Automating and enhancing the diagnostic process, we have the potential to improve the lives of individuals affected by lung cancer and make a positive impact on the field of medical imaging and healthcare.

Using softmax layer and autoencoder features combinly formed the deep learning and it is used for classifying the new query image into normal or cancerous.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, No. 3, December : 2023

Performance Analysis

Here the accuracy of prediction is calculated by testing multiple images. Accuracy of proposed model is higher compare to state of art techniques.

Result

The results of our efforts in this project are promising. The deep learning model demonstrated a high degree of accuracy and efficiency in detecting lung cancer and classifying it into its different stages. This innovation has the potential to substantially reduce human error and subjectivity in the diagnostic process, ultimately leading to earlier intervention and better treatment outcomes for patients.

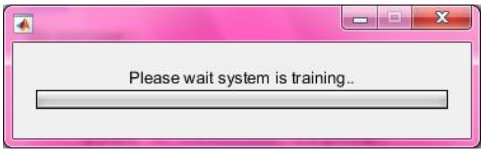

Fig.4.1 Train Deep learning for Normal or Abnormal CT lung images

There are two folders in dataset one include ‘Normal’ images and another ‘Abnormal’ images. Both are given as input to the autoencoder based deep learning.

There are two layers of autoencoders are used

a) Autoencoder level1

b) Autoencoder level2

The softmax layer uses these autoencoder features to form a stack of autoencoder features. The autoencoder features and softmax layer forms a deep learning model. Obtained deep learning model is further used for prediction of lung CT image into normal and abnormal. Abnormal lung CT image again classified into two categories as benign and malignant. Benign is the starting stage and malignant is the higher stage.

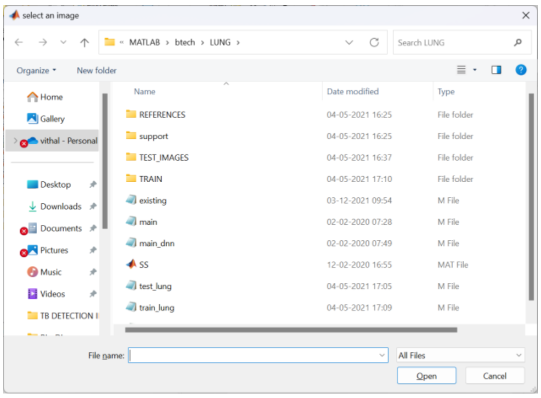

Fig.4.2 Selection of New Test Image for Prediction

In above window, user can select any test image for prediction. The algorithm detects the image as either normal or abnormal and further classified into stages.

For Testing

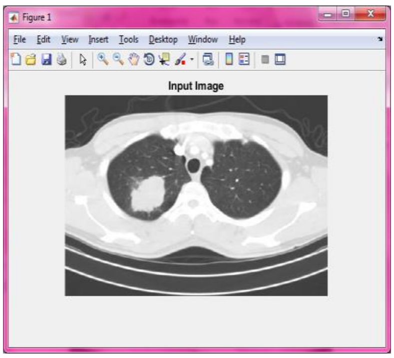

Fig.4.3 Input Testing Image

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, No. 3, December : 2023

Above lung CT scan image is used for testing or query for prediction of nromal or abnormal. The selcted image resized to standard size.

Input lung CT image may contain noise as well as contrast issues which can be solved using preprocessing techniques. Preprocessing techniques helps to get lung CT image in standard format. Lung CT image obtained from preprocesisng gives better performnace for detection as well as prediction of lung CT scan.

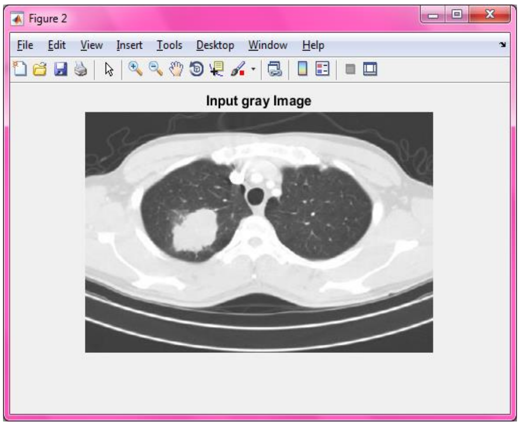

Fig.4.3 Input Testing Image

Fig. Image is converted to gray scale to reduce complexity of an image. In gray scale single pixel uses only 8 bits while operation while in color image it uses 24 bits while operation.

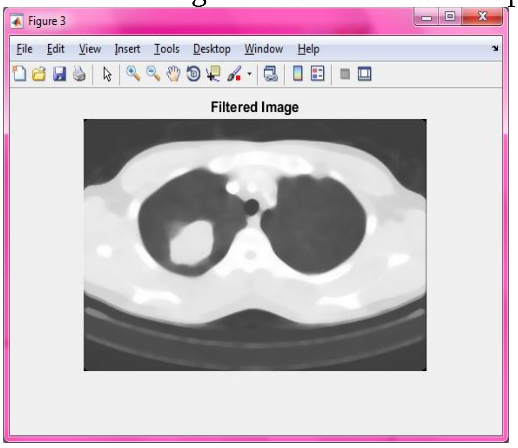

Fig.4.5 Filtered Image

The medical images may have different types of noises like gaussian noise , salt and pepper noise, white noise etc. but most common noise is salt and pepper noise which can be removed using median filter. Above obtained image is median filtered image.

Salt and pepper noise means mostly we can see black and white dots. Such black and white dots can be removed by taking median of neighboring pixels to the salt or pepper. (that is 0 or 1).

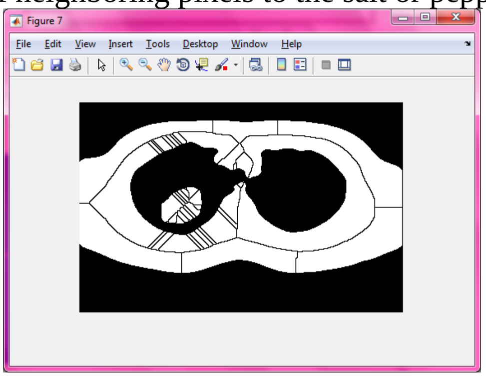

Fig.4.6 Lung Image Segmented

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 12, No. 3, December : 2023

Using segmentation techniques lung image is segmented. It shows both lung cancer as well as lung outlier.

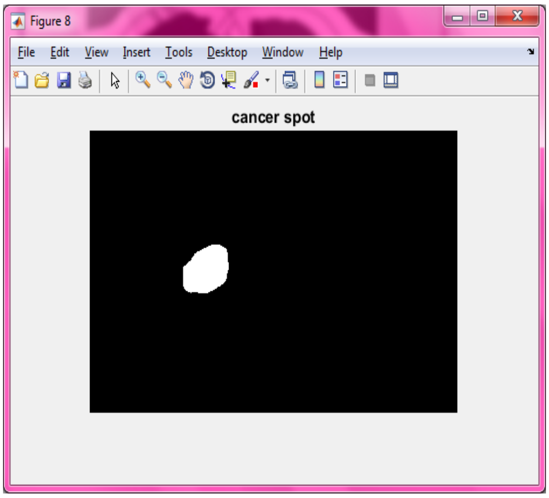

Fig.4.7 Detected Cancer spot using Morphological operations

Proposed method uses morphological operations such as dilation , erorsion , oprning and closing for segmnetation of perfect lung cancer part.

The detected cancer spot is convted to bloack and white image which is segmentation results.

- Black Region: Background
- White Region : Cancerous Part

Thenafter features of white region are extracted such as are , perimerer , these features are further used for classification.

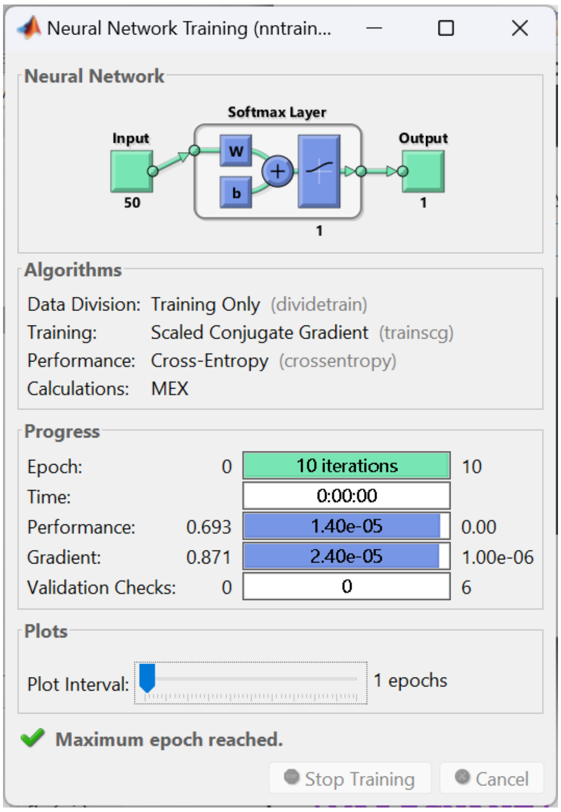

Fig. The Deep learning model is predicting with 10 iterations.

There are 10 iterations used for prediction. Deep learning model with different autoencoder layers and softmax layer are used for further prediction.

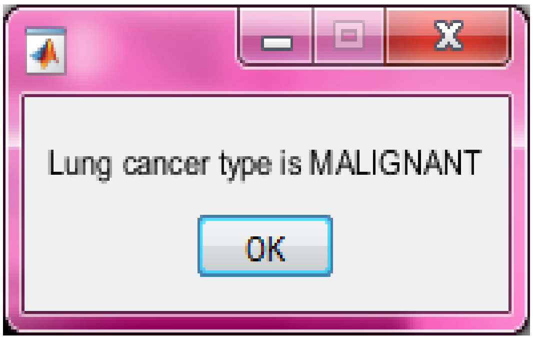

Fig.4.7 Detected Type of cancer from lung CT image

Cancer is detected as Maligant cancer by proposed model.

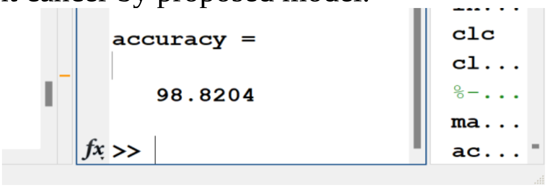

Fig.4.8 Percentage of classification accuracy

Performance analysis of proposed model gives accuracy of 98.8204 which is very high accuracy. And hence proposed model works superior to existing state of art techniques.

Conclusion

In this project, we embarked on the endeavor of leveraging deep learning techniques within MATLAB to address the critical challenge of lung cancer detection and staging. Lung cancer, a major public health concern, demands early and accurate diagnosis for improved patient outcomes. For automatic detection of lung cancer we used a Computer Aided Diagnosing (CAD) system. Using these we can extract the features and generating diagnosis rules. We take multiple images to perform experiment from the hospital. In result we see that the proposed system CAD can detect the false positive nodules correctly. Proposed autoencoder based deep learning model gives higher accuracy of 98% which is very higher accuracy for prediction of cancer from lung CT images.

References

[1] Jiang, H., Qian, W., Gao, M., Li, Y. " An automatic detection system of lung nodule based on multigroup patch-based deep learning network" IEEE Journal of Biomedical and Health Informatics,22(4):1227-1237.

[2] Xin-Yu Jin, Yu-Chen Zhang, Qi-Liang Jin "Pulmonary Nodule Detection Based on CT Images Using Convolution Neural Network." 9Th International Symposium on Computational Intelligence and Design. 2016.

[3] Armato, S.G., M.L. Giger and H. MacMahon, 2001.Automated detection of lung nodules in CT scans:Preliminary results. Med. Phys., 28: 1552-1561.http://www.ncbi.nlm.nih.gov/pubmed/11548926

[4] Fiebich, M., D. Wormanns and W. Heindel, 2001.Improvement of method for computer-assisted detection of pulmonary nodules in CT of the chest.Proc. SPIE Med. Image. Conf., 4322: 702-709.DOI: 10.1117/12.431147

[5] Ginneken, B.V., B.M. Romeny and M.A. Viergever,2001. Computer-aided diagnosis in chestradiography: A survey. IEEE, Tran. Med. Image.,20: 1228-1241. ISSN: 0278-0062

[6] R. Pandian, V. Vedanarayanan, D.N.S. Ravi Kumar, R. Rajakumar, Detection and classification of lung cancer using CNN and Google net, Measurement: Sensors Volume 24, 2022, 100588, ISSN 2665-9174,https://doi.org/10.1016/j.measen.2022.100588.

[7] Suzuki K., et al., “False-positive Reduction in Computer-aided Diagnostic Scheme for Detecting Nodules in Chest Radiographs by Means of Massive Training Artificial Neural Network”, Academic Radiology, 12, No 2, February 2005, pp. 191-201.

[8] Non-Small Cell Lung Cancer, Available at: http://www.katemacintyrefoundation.org/ pdf/non-small-cell.pdf, Adapted from National Cancer Institute (NCI) and Patients Living with Cancer (PLWC), 2007, (accessed July 2011).

[9] Tarawneh M., Nimri O., Arqoub K., Zaghal M., Cancer Incidence in Jordan 2008, Available at: http://www.moh.gov.jo/MOH/Files/Publication/Jordan%20Cancer%2 0Registry_2008%20Report_1.pdf, 2008, (accessed July 2011).

[10] Lung Cancer Database, Available at: https://eddie.via.cornell.edu/cgi- bin/datac/signon.cgi, (accessed July 2011).

[11] Gonzalez R.C., Woods R.E., Digital Image Processing, Upper Saddle River, NJ Prentice Hall, 2008.

[12] Cristobal G., Navarro. R., Space and frequency varient image enhancment based in Gabor representation, Pattern Recognition Letters, Elsevier, 1994, 15, p. 273-277.

[13] Krishan A., Evaluation of Gabor filter parameters for image enhancement and segmentation, in Electronic Instrumentation and Control Engineering, Master. Punjab: Thapar University, 2009, p. 126.