International Journal of Advanced and Innovative Research (2278-7844)/

Volume 7 Issue 6

A Survey over Lung Cancer Detection Using Digital Image Processing and Artificial Neural Networks

Chandni Kumari[1], Kamlesh Chandravanshi[2], Deepak Tomar[3]

[1]Research Scholar, CS, Technocrats Institute of Technology College, Bhopal, India
[2]Guide, CS, Technocrats Institute of Technology College, Bhopal, INDIA

[1]Email-ID: singhxchandni@gmail.com
[2]Email-ID: kamlesh.vjti@gmail.com

Abstract: ***Data mining is the process of extraction of data from input resource and finding a usable entity from it. Medical field produces multiple data which can actually use for medical analysis and predictions. MRI and other CT scan produces the structure of internal body components and symptoms can be recognize using these images. An expensive human skillset is always required to predict the immense disease such as lung cancer. Digital image processing is the approaches which help in analysis of an image using its feature extraction and classification process. Various past approaches are limited to certain criteria and feature observation. In this paper a survey of various approaches which takes a part for lung cancer analysis and other disease prediction is discussed. An image processing algorithm and their usage has been discussed by different author is presented. Thus our proposed system is needed to present and enable of extraction the proper lung cancer detection. A proposed technique using the A proposed BATLO (Bat classification with Linear function Optimization) approach is used for efficient accuracy; precision, low error rate with hybrid classification approach is going to defined by our work which can predict the decision efficiently.***

Keywords: - lung cancer, image processing, artificial network, histogram, feature extraction.

Introduction

In computer science, image processing is any kind of signal processing for which the input is an image, such as photographs or frames of video; the output of image processing can be an image. It can also be a set of characteristics or a set of image related parameters [1]. An image-processing technique mostly involves treating the image as a two-dimensional signal and it constitutes applying various signal-processing techniques to the image.

Image processing includes different forms of information processing where the input is an image. Image processing techniques are mostly derived from the signal processing techniques application in the domain of images, such as pictures or videos [2].

According to recent surveys, cancer-related medical expenses and labor loss cost annually 10,000 billion dollars all over the world [1]–[3]. Lung cancer is a

number one killer among all cancer-leading deaths, due to late stage detection and environmental conditions, such as air pollution, working conditions, life-long smoking habits. For instances, 225,000 new cases were detected in the United States in 2016, and 4.3 million new cases in China in 2015 [2], [6]. Like other types of cancers, early detection is viewed to be the best strategy to save lives. Computerized Tomography (CT) is 3D imaging modality which has been widely used for lung cancer screening and diagnostics. 3D images are reconstructed from thousands of 2D X-ray transmission projections. Advanced 3D reconstructions [9] were developed for better image quality and diagnostic accuracy.

Most current machine learning based Computer Aided Diagnostic (CAD) researches are focusing on non-small cell lung cancer (NSCLC) [10]. CAD systems help to reduce the workload of radiologists significantly [10], [13]. So far there are very few work on small cell lung cancer (SCLC) detection, which is an extremely difficult task for human observer because the image with SCLC looks almost identical to one without. There are machine learning algorithms that may be candidates for SCLC detection task, such as convolution neural network based deep learning method, which starts with building neurons and layers, where a dynamic parameter set is used to calculate forward propagation. During the training process, parameters in each layer are updated by back propagation from cost function (i.e. a distance metric between forward propagation of input data and labels) [15]. However deep learning algorithms usually requires an extremely large training dataset, which is not always available.

Literature Review

In this section survey of previously given approach which are used for the lung cancer detection and prediction. A literature of the previously given approaches is presented.

In [4] they have worked with lung cancer detection from the image dataset where the CT scanned level of images were taken and used for the experiment purpose which is taken from NIH/NCI Lung Database Consortium and automation of the lung cancer detection is performed using the given technique. A different level of extraction is used from the dataset which includes region of interest selection at first, then a slicing technique used for binary conversation of the CT scanned image. Further a segmentation technique is applied over the selected region of interest and then a classification using the rule mining approach is performed. Finally a accuracy of 80% is achieved with proper classification technique over lung cancer CT scanned image dataset.

Author [5] they have implemented an computerized system algorithm for the computational purpose. The CT scanned images were taken in this paper for experiment. Lung images segmentation is performed initially with the images, enhancement of edges and weak areas were performed with the images, further a classification based on feature extraction is performed. The threshold segmentation technique is

used by the author to outperform their given work at best level. A feature vector mechanism is used for anomalous region extraction. A fuzzy classifier is then applied on the extracted data which is used with the given input images and features extracted from the CT scanned image dataset.

Author [7] describes Lung carcinoma is one of the most lethal of cancers worldwide. Positron emission tomography (PET) data has greater sensitivity and specificity in the staging of lung cancer than computed tomography (CT) or magnetic resonance imaging (MRI). KNN and SVM approach for the better classification and effective analysis of lung cancer detection is performed by the paper work.

Author [11] proposed a noninvasive detection method of lung cancer combined with a sort of virtual l SAW gas sensors and imaging recognition method. Patients breathe goes through an electronic nose with solid phase micro extraction (SMPE) and capillary column for pre concentration and separation of volatile organic compounds (VOCs) respectively.

Author [12] has developed a neural- digital computer aided diagnosis system (CAD) system based on a parameterized two level convolution neural network (CNN) Architecture and on a special multipliable output encoding procedure. The developed architecture was trained, tested and evaluated specially on the problem of diagnosis of lung cancer nodule found on digitized chest radio- graphs. The system performs automatic suspect localization feature extraction and diagnosis of a particular pattern class-aimed at a high degree of “true-positive fraction” detection and “low positive fraction” detection.

Author [13] presented the prevalence of respectable lung cancer detected by LDCCT at baseline screening was low at 0.23%, but there was a high rate of significant incidental pathology. Low dose chest computed tomographic scanning (LDCCT) can detect early stage asymptomatic lung cancer in a high risk urban population. Four hundred forty patients underwent surgery for primary lung cancer, and 45 normal cases were selected. Eight radiologists participated in observer tests.

Authors	**Algorithm/Technique**	**Remark/ Further extension**
Poulami Raha [4]	CT Scanned Images (They have taken the CT scanned images and an feature extraction analysis using the curve information was presented)	Previous approach has worked on the basic KNN classification. Proposed an automatic computer aided diagnosing system for detection of lung cancer by analyzing these

		lung CT images.
Shreya K Chari [5]	Detected the lung nodules with the help of CT scan images.(They have worked with the image segmentation, the prediction performed according to node detail in images)	A single image feature extraction and detection over the single phase is performed in previous section. The threshold segmentation technique was applied, for removing background and extracts the nodules from an image.
Kazuya Kawai [7]	Positron emission tomography (PET). (A nearest neighbor detection and finding relation between the extracted component were taken into part for prediction. SVM classifier approach is used for the prediction of lung cancer detection.	An approach using the basic NB classifier is used by the previous work in this paper. The proposed work given which is Using k-nearest neighbor and support vector machines (SVM) classifiers.
Asif I. Zia, Nasrin Afsarimanesh [11]	Noninvasive detection method of lung cancer.(A size and measurement based approach is used for the lung cancer detection and prediction.)	Invasive way of detection is presented in previously given solution. Patients breathe goes through an electronic nose with solid phase micro extraction (SMPE) and capillary column.
J. Fu [13]	Lung cancer detected by LDCCT (A feature extraction using the radiologist images were used to perform the prediction).	Previously given approach worked on the classical dataset. Whereas the

		proposed work is performed on Eight radiologists participated in observer tests.

Table 1: Comparison analysis and working drawbacks and future extend for the available recent algorithms.

In the comparison table 1 above, some existing recent algorithms are discussed. The work shows the previously given approaches, their remark and proposed solution given by them on previous problems.

Proposed Work

The proposed technique starts with preprocessing, segmentation, classification which we have used in our paper.

A. Pre-Processing:

Preprocessing is the initial step of a CAD system for noise removal and image enhancement. An ultrasound machine produces raw images having low contrast and speckle noise. Speckle noise is multiplicative and random in nature. The presence of speckle noise, blurring edges and lower contrast resolution make an accurate segmentation of nearly impossible which in turn leads to misdiagnosis of cancer. A lot of preprocessing techniques are available like linear filtering, Wiener filtering, wavelet de-speckling, wavelet filtering etc. however, speckle reduction anisotropic diffusion (SRAD) algorithm is the most commonly used preprocessing technique for ultrasound images.

B. Segmentation:

Segmentation is the most important step after preprocessing of the BUS images. An efficient segmentation extracts the required region of interest (ROI), which is the breast mass in BUS images from its background. A lot of segmentation algorithms are available like histogram thresholding, Otsu's method, level set method, active contour model (ACM), Markov random field (MRF) etc. In most of the methods, segmentation is done after manual selection of mask over the region of interest. Here segmentation is done without any mask selection, which makes this CAD system fully automatic. Marker based watershed algorithm is used to segment mass from the background or surroundings.

C. Classification:

Classification is the final step of the proposed CAD system. In this step, based on an extracted feature set, BUS images are classified into benign and malignant category. Classification can be divided into two step (1) Training and (2) Testing. Initially based on features, the classifier is trained and the

corresponding classes are developed. In testing phase, BUS images are classified as benign or malignant cases. Most commonly used classifiers are support vector machine (SVM), artificial neural networks (ANN), self-organizing map (SOM), K-Means clustering algorithm etc. In this CAD system, k-Nearest Neighbors algorithm (k-NN) is used for classification of BUS images. The k-NN classifier is a instance based classifier which classifies objects based on prediction method.

Conclusion

The Image classification is the field of process an image using the techniques which help in processing the image components and their pixels. Data mining help in classification of large amount of data into different segment according to input and requirements. Lung cancer is the disease which needs a prediction using the image retrieve through CT scan and other instrumental image outputs. Previously given approaches worked on classification such as SVM, K-Mean and other feature extraction such as color, edges, dots etc. Thus a proper classification is still lacking which need an enhance approach for lung cancer prediction with high accuracy. In this paper a survey of previous approach and the proposed solution to achieve high accuracy is presented. The proposed further work can be experimented to perform and show improvement through our approach.

REFERENCES

[1] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, “Global cancer statistics, 2012,” CA: a cancer journal for clinicians, vol. 65, no. 2, pp. 87–108, 2015.

[2] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2016,” CA: a cancer journal for clinicians, vol. 66, no. 1, pp. 7–30, 2016.

[3] I. Hwang, D. W. Shin, K. H. Kang, H. K. Yang, S. Y. Kim, and J.-H. Park, “Medical costs and healthcare utilization among cancer decedents in the last year of life in 2009,” Cancer research and treatment: official journal of Korean Cancer Association, vol. 48, no. 1, p. 365, 2016.

[4] Poulami Raha, Fully Automated Computer Aided Diagnosis System for Classification of Breast Mass from Ultrasound Images, IEEE WiSPNET 2017 conference.

[5] Shreya K Chari, Threshold Selection in Image Segmentation Using Parametric Entropy Measures, 2017 Fourth International Conference on Image Information Processing (ICIIP).

[6] W. Chen, R. Zheng, P. D. Baade, S. Zhang, H. Zeng, F. Bray, A. Jemal, X. Q. Yu, and J. He, “Cancer statistics in china, 2015,” CA: a cancer journal for clinicians, vol. 66, no. 2, pp. 115–132, 2016.

[7] Kazuya Kawai, A Robust PET Image Reconstruction Using Constrained Non-negative Matrix Factorization, 12 - 15 December 2017, Malaysia

[8] S. Xu, A. Uneri, A. J. Khanna, J. Siewerdsen, and J. Stayman, "Polyenergetic known-component ct reconstruction with unknown material compositions and unknown x-ray spectra," Physics in Medicine and Biology, vol. 62, no. 8, p. 3352, 2017.

[9] S. Xu, J. Lu, O. Zhou, and Y. Chen, "Statistical iterative reconstruction to improve image quality for digital breast tomosynthesis," Medical physics, vol. 42, no. 9, pp. 5377–5390, 2015.

[10] K. Suzuki, "Pixel-based machine learning in computer-aided diagnosis of lung and colon cancer," in Machine Learning in Healthcare Informatics. Springer, 2014, pp. 81–112.

[11] Asif I. Zia, Nasrin Afsarimanesh, Improved Detection Limits for Phthalates by Selective Solid-phase Micro-extraction, 2015 Ninth International Conference on Sensing Technology.

[12] Bassem Bouaziz, convolutional neural networks, 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications.

[13] J. Fu, Accurate Image Reconstruction of Few-view Ptychrography X-ray Computed Tomography, August 2016.

[14] H. Dang, J. Stayman, A. Sisniega, J. Xu, W. Zbijewski, J. Yorkston, N. Aygun, V. Koliatsos, and J. Siewerdsen, "Cone-beam ct of traumatic brain injury using statistical reconstruction with a post-artifact-correction noise model," in Proceedings of SPIE–the International Society for Optical Engineering, vol. 9412. NIH Public Access, 2015.

[15] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in Proceedings of the 23rd ACM international conference on Multimedia. ACM, 2015, pp. 689–692.