Impact Factor (SJIF): 5.301

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444
Volume 6, Issue 12, December-2019

Lung Disease Detection By Using CNN Algorithm

Dr. A. A. Dandawate
Miss. P. R. Mule
JSPM college of engineering, Handewadi-Hadpsar

ABSTRACT -

Lung Cancer is a Disease of uncontrolled cell growth in tissues of the lung. Discovery of Lung Cancer in its initial stage is the key of its cure. All in all, a measure for ahead of schedule stage lung disease determination essentially incorporates those using X-beam midsection movies, CT, MRI and so forth. In numerous parts of the world far reaching screening by CT or MRI is not yet pragmatic, so that midsection radiology stays in starting and most basic system. Firstly, we will utilize a few systems are key to the errand of medicinal picture mining, Lung Field Segmentation, Data Processing, Feature Extraction, Classification utilizing neural system and SVMs. The routines utilized as a part of this paper work states to group computerized X-beam midsection movies into two classes: ordinary and unusual. Diverse learning examinations were performed on two distinctive information sets, made by method for highlight choice and SVMs prepared with diverse parameters; the outcomes are looked at and reported.

Keywords- Frequent item set, closed high utility item-set, lossless and concise representation, utility mining, data mining

Introduction

Lung Disease Is A Noteworthy reason for Mortality in the western world as exhibited by the striking factual numbers distributed consistently by the American Lung Disease Society. They demonstrate that the 5-year survival rate for patients with lung malignancy can be enhanced from a normal of 14% up to 49% if the ailment is analyzed and treated at its initial stage. Medicinal pictures as a vital piece of therapeutic determination and treatment were focusing on these pictures for good. These pictures incorporate success of concealed data that misused by doctors in settling on contemplated choices around a patient. Then again, removing this important shrouded data is a basic first stride to their utilization. This reason inspires to utilize information digging systems abilities for productive learning extraction & find concealed lung. Mining Medical pictures includes numerous procedures. Medicinal Data Mining is a promising zone of computational insight connected to a consequently break down patients records going for the disclosure of new information valuable for restorative choice making. Affected information is expected not just to increment exact determination and effective infection treatment, additionally to improve security by diminishing blunders. The systems in this paper arrange the advanced X-beam midsection movies in two classes: ordinary and strange. The irregular ones incorporate Type of lung tumor; we will utilize a typical arrangement technique specifically CNN & neural systems.

Problem Statement

The existing system is Time consuming process, and It is very difficult to detect it in its early stages as its symptoms appear only in the advanced stages. Implementing the system to automate the classification process for the early detection of Lung Cancer.

Objectives

The key objectives of current work are as given below:

To proposed Some techniques are essential to the task of medical image mining, Lung Field Segmentation, Data Processing, Feature Extraction, Classification using and SVMs. The methods used in this paper work states to classify digital X-ray chest films into two categories: normal and abnormal. Different learning experiments were performed on two different data sets, created by means of feature selection and (Support Vector Machine)SVMs trained with different parameters; the results are compared and reported.

Mathematical Model:

Mathematical model of the proposed system

INPUT:-

Let S is the Whole System Consists:

Let S is the Whole System Consist of

S= {I, P, O}

Where,

I = input.

I = {U, Q}

U = User

U = {u1,u2,….un}

Q = Query

Q = {q1,q2……qn}

P = Process

P = {MBWA, WA, GLCOMA, SVM}

MBWA =Marker-Based Watershed Algorithm.

WA = Watershed Algorithm.

GLCOMA = Grey level co-occurance matrix algorithm.

SVM = Support Vector Machine

OUTPUT: The predicted result will be the output of the system

SCOPE

We describe what features are in the scope and what are not in the scope of the system to be developed. Our project is like a component which can be used on different ways in future. Component can be implemented in hospital management system to improve and support doctor's work. It can be also be used in as Android App for more generalized purpose.

In Scope

- Project can be used in every hospital for diagnosis of patients.
- This project can be used as inbuilt component in Hospital Management System.

It is the application in which we are trying to implement Classification algorithm for that we develop small kind of hospital management system.

Out Scope

- System can be further modified into android application.
- The modules used in this system can be used in Cognitive Computing model.

LITERATURE REVIEW

1	Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images	2015	K.Punithavathy M.M.Ramya2 Sumathi Poobal	cancer detection for PET/CT images using texture analysis and FCM. Pre-processing techniques enhance the accuracy of the cancer detection. Pre-processing techniques enhance the accuracy of the cancer detection. Morphological operations enable accurate lung ROI extraction and reduce the search space. The results of our methodology illustrate that the texture analysis yielded number of significant texture features. These features fed as input to the FCM classifier help in accurate detection of the lung cancer. Results of the proposed methodology are promising with an overall accuracy of 92.67%.
2	Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm	2017	Qing Wu and Wenbing Zhao	In this study, we proposed an EDM machine learning algorithm with vectorized histogram features to detect SCLC

				for early malicious cancer prediction. While we show that EDM has reasonably good prediction accuracy, there is a large room for improvement before our algorithm can be used in the clinical setting.The ultimate goal of this study is to develop a clinical decision-making system for radiologists to better predict a malicious lung cancer from SCLC with computed tomography (CT) imaging. For the future work, we would train the proposed method with larger training set and deeper network, and combine it with convolution neural network, which has been used in CT imaging for different applications [17], [23]..
3	CNN-based Method for Lung Cancer Detection in Whole Slide Histopathology Images	2017	Matko Sari ˇ c, Mladen Russo, Maja Stella, Marjan Sikora	In this paper we proposed the fully automatic deep learning based method for detection of lung cancer in whole slide histopathology images. VGG16 and ResNet50 CNN architectures were compared and the first one shows higher AUC and patch classification accuracy. Presented results shows that convolutional neural networks have potential to perform lung cancer diagnose from whole slide

				images, but more effort is needed to increase classification accuracy. In future work next steps will be increasing the training set size, adding image augmentation and stain normalization. Also, we will try training from the scratch instead of using weights pretrained on ImageNet.
4	Segmentation and Analysis of CT Chest Images for Early Lung Cancer Detection	2016	Rachid Sammouda Computer Sciences Department	In this study, we have achieved our purpose in developing an automatic CAD system for early detection of lung cancer by analyzing lung human CT images using several phases.
5	Hybrid Approach For Feature Extraction of Lung Cancer Detection	**2018**	Ms. Twinkal Patel Asst. Professor Mr. Vimal Nayak	We presume that the proposed algorithm is better than the existing algorithms as far as the efficiency of decisionmaking process gets simpler for Also, extracted features are can be easily evaluated for classification. Regardless of giving a decent yield the proposed algorithm in future it can in any case be enhanced to get precise outcomes

PROPOSED SYSTEM

Objective of the proposed system is to introduce a unique "Predictive Diagnostic System" The original image is transformed to gray scale image. After that, removal of the noises and contrast enhancement is done for obtaining the enhanced images. After that, removal of the noises and contrast enhancement is done for obtaining

the enhanced images. Firstly Image acquisition is done on the image, system performs pre-processing on image. Find out affected regions and their characteristics in form of data. This data is classified using CNN. CNN classify it as normal or diseases lung and identify lung diseases.

SYSTEM ARCHITECTURE

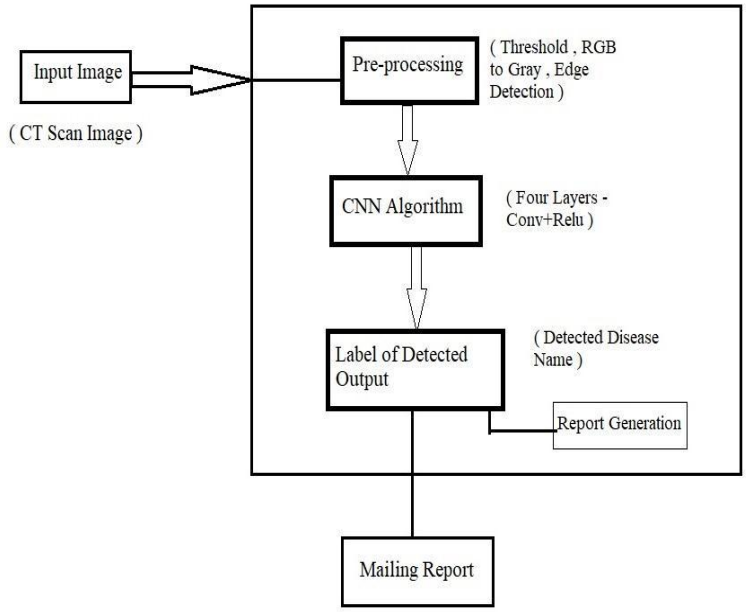

Figure: System Architecture

EXISTING SYSTEM

Medical data mining is one of the major issues in this modern world. Medical problems are often in each and every human being. Cancer is one of the most dangerous diseases a human can ever had. Lung cancer is one of them. Lung cancer is a disease that occurs due to the uncontrolled cell growth in tissues of the lung. It is very difficult to detect it in its early stages as its symptoms appear only in the advanced stages.

PROPOSED SYSTEM

we will use some techniques are essential to the task of medical image mining, Lung Field Segmentation, Data Processing, Feature Extraction, Classification using neural network and SVMs. The methods used in this paper work states to classify digital X-ray chest films into two categories: normal and abnormal. Different learning experiments were performed on two different data sets, created by means of feature selection and SVMs trained with different parameters; the results are compared and reported.

ADVANTAGES OF SYSTEM

- Utilization of time management.
- Fast process.
- Problem of maintaining privacy can be solved.
- High security.
- Data is highly secure.

CONCLUSION

In this paper, various periods of picture handling were applied on Lung Nodules. These different image processing techniques, the fuzzy filter will provide the efficient daignosis. Division done by marker based watershed calculation, gives different district of picture.GLCM is used to extract the different features of image and which takes less time for generating the result. This results are passed through CNN Classifier, which classifies the nodules as benign or malignant. CNN classifier provides 92.5% accuracy.

REFERENCES

[1] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in Proc. 20th Int. Conf. Very Large Data Bases, 1994, pp. 487–499.

[2] R. Chan, Q. Yang, and Y. Shen, "Mining high utility itemsets," in Proc. IEEE Int. Conf. Data Min., 2003, pp. 19–26.

[3] A. Erwin, R. P. Gopalan, and N. R. Achuthan, "Efficient mining of high utility itemsets from large datasets," in Proc. Int. Conf. Pacific- Asia Conf. Knowl. Discovery Data Mining, 2008, pp. 554–561.

[4] H.-F. Li, H.-Y. Huang, Y.-C. Chen, Y.-J. Liu, and S.-Y. Lee, "Fast and memory efficient mining of high utility itemsets in data streams," in Proc. IEEE Int. Conf. Data Mining, 2008, pp. 881–886

[5] C. F. Ahmed, S. K. Tanbeer, B.-S. Jeong, and Y.-K. Lee, "Efficient tree structures for high utility pattern mining in incremental databases," IEEE Trans. Knowl. Data Eng., vol. 21, no. 12, pp. 1708– 1721, Dec. 2009.

[6] J.-F. Boulicaut, A. Bykowski, and C. Rigotti, "Free-sets: A condensed representation of Boolean data for the approximation of frequency queries," Data Mining Knowl. Discovery, vol. 7, no. 1, pp. 5–22, 2003.

[7] T. Calders and B. Goethals, "Mining all non-derivable frequent itemsets," in Proc. Int. Conf. Eur. Conf. Principles Data Mining Knowl. Discovery, 2002, pp. 74–85.

[8] K. Gouda and M. J. Zaki, "Efficiently mining maximal frequent itemsets," in Proc. IEEE Int. Conf. Data Mining, 2001, pp. 163–170.

[9] T. Hamrouni, "Key roles of closed sets and minimal generators in concise representations of frequent patterns," Intell.Data Anal., vol. 16, no. 4, pp. 581–631, 2012.

[10] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation," in Proc. ACM SIGMOD Int. Conf. Manage.Data, 2000, pp. 1–12.

[11] Y. Liu, W. Liao, and A. Choudhary, "A fast high utility itemsetsmining algorithm," in Proc. Utility-Based Data Mining Workshop, 2005, pp. 90–99.

[12] V. S. Tseng, C.-W.Wu, B.-E.Shie, and P. S. Yu, "UP-Growth: An efficient algorithm for high utility itemset mining," in Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2010, pp. 253–262.

[13] B. Vo, H. Nguyen, T. B. Ho, and B. Le, "Parallel method for mining high utility itemsets from vertically partitioned distributed databases," in Proc. Int. Conf. Knowl.-Based Intell. Inf. Eng. Syst., 2009, pp. 251–260.

[14] C.-W.Wu, B.-E. Shie, V. S. Tseng, and P. S. Yu, "Mining top-k high utility itemsets," in Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2012, pp. 78–86.

.