|| Volume 5 || Issue 6 || June 2020 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

CNN BASED METHOD FOR LUNG DISEASE DETECTION

Shrikant N Borude[1], Saurabh R Kolhe[2], Harshal R Patil[3], Prof Laxmikant Malphedwar [4]

Computer Engineering, Dr D Y Patil School of Engineering Academy, Pune, India [1,2,3,4]

Shrikantborude8@gmail.com,kolhesaurabh23@gmail.com,harshalpatil560@gmail.com,

---------- ***----------

Abstract: - Lung Cancer is a Disease of uncontrolled cell development in tissues of the lung. Disclosure of Lung Cancer in its underlying stage is the key of its fix. All things considered, a measure for in front of calendar arrange lung malady assurance basically consolidates those utilizing X-shaft waist motion pictures, CT, MRI, etc. In various pieces of the world expansive screening by CT or MRI isn't yet down to earth, with the goal that midriff radiology remains in beginning and most fundamental framework. Initially, we will use a couple of frameworks are critical to the task of therapeutic picture mining, Lung Field Segmentation, Data Processing, Feature Extraction, Classification using neural framework and CNNs. The schedules used as a piece of this desk work states to gather mechanized X-shaft waist motion pictures into two classes: conventional and bizarre. Differing learning assessments were performed on two particular data sets, settled on by strategy for feature decision and CNN arranged with assorted parameters; the results are taken a gander at and detailed.

Keywords -Frequent itemset, closedþ high utility itemset, lossless and concise representation, utility mining, and data mining

---------- ***----------

I INTRODUCTION

Lung diseases are one of the most dangerous diseases a human can ever have. It is very hard to identify it in its beginning periods as its side effects seem just in the propelled stages.

Computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography(PET), mammography, ultrasound and X-ray, have been used for the early detection, diagnosis, and treatment of diseases which caused large volumes of patient data being generated and are becoming widely available. We have identified this type of problems which is solved by this proposed system. 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.

II PROBLEM STATEMENT

You have some kind of health problem and visit to nearest hospital, consultant suggests for CT scan of Lung. So you visit for CT scan and ideally you need to wait for 2 days for getting report. Again then you need to consult doctor regarding reports and then they will suggest you further medical treatments. So this is long process you need to go under.

Objectives:

The objective of proposed system is to build Predictive Diagnostic System" of infectious lung. Proposed system works on the concept of image processing techniques such as feature subtraction, extraction, selection along with CNN classifier algorithms (machine learning) so as to precisely predict & detect various diseases of infectious lung. As an integrated approach this system will not only predict and detect various lung diseases, but it will also generate test reports that can be used for a preventative treatment.

Motivation:

Deal with automatic system instead of manual. Eliminate 5 lung diseases using advanced

U machine technique.

Lung Disease is a Disease of uncontrolled cell growth in tissues of the lung. Discovery of Lung Disease 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.

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

= {u1,u2,……...un}

Q = Query

Q = {q1,q2…….……qn}

P = Process

P = {CNN}

CNN = Convolution Neural Network

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

III LITERATURE REVIEW

1. Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm

Author:- 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].

2. **Hybrid Approach For Feature Extraction of Lung Cancer Detection**

Author:- 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.

3. **CNN-based Method for Lung Cancer Detection in Whole Slide Histopathology Images**

Author:- Matko ˇ Sari´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 classificiation 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. **Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images**

Author:- K.Punithavathy1, M.M.Ramya2, Sumathi Poobal3

Our methodology successfully developed automatic lung cancer detection for PET/CT images using texture analysis and FCM. 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%.

IV PROPOSED SYSTEM

Objective of the proposed system is to introduce a unique "Predictive Diagnostic System" The original image is transformed to gray scale image. From that point onward, expulsion of the commotions and difference improvement is accomplished for getting the upgraded pictures. 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.

- ✓ Hanta Virus
- ✓ Pneumonia
- ✓ Bronchitis
- ✓ Lung Cancer

Risk identified such as technical

1. Accuracy
2. System Failure

Operational

1. Human Error
2. File Handling

Business Risk

1. Virus Attack
2. System Hacking

V SYSTEM ARCHITECTURE

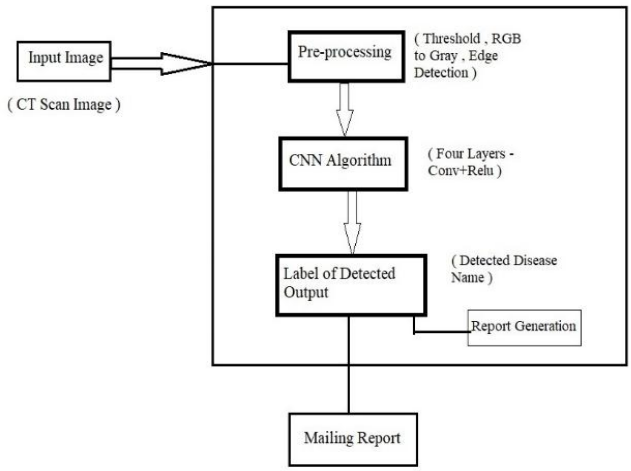
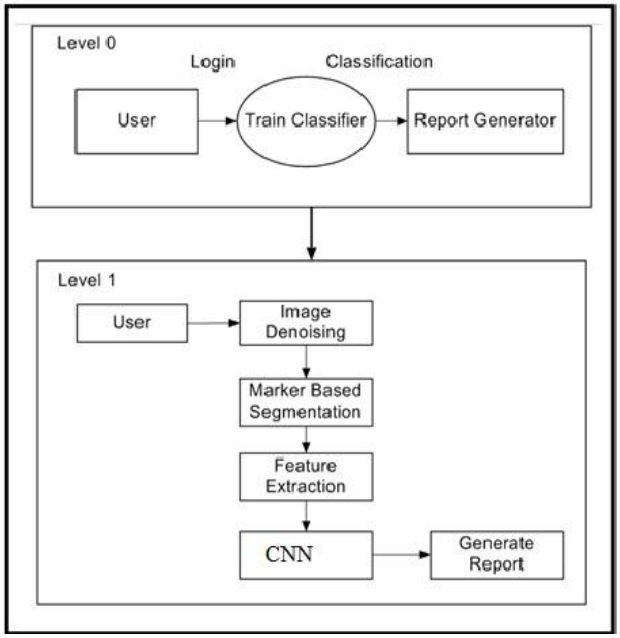


Figure 1: System Architecture



Research gap identified

1. CNN algorithm can be used
2. It works with four layers and works effectively on images.
3. It provides accuracy
4. And can be reused on the same data

In previous we could only find or detect the lung cancer.

EXISTING SYSTEM

Medical information mining is one of the major troubles in this contemporary world. Medical problems are frequently in each and every human being. Cancer is one of the most dangerous sicknesses a human can ever had. Lung cancer is certainly one of them. Lung cancer is disorder that occurs due to the uncontrolled cell boom in tissues of the lung. It is very tough to discover it in its early ranges as its signs appear most effective inside the superior tiers.

PROPOSED SYSTEM

We will try strategies that are fundamental for the assignment of restorative picture mining, Lung Field Segmentation, Data Processing, Feature Extraction, Classification utilizing neural system and CNNs.. Different learning experiments were performed on two different data sets, created by means of feature selection and CNNs 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.

VI RESULTS AND DISCUSSIONS

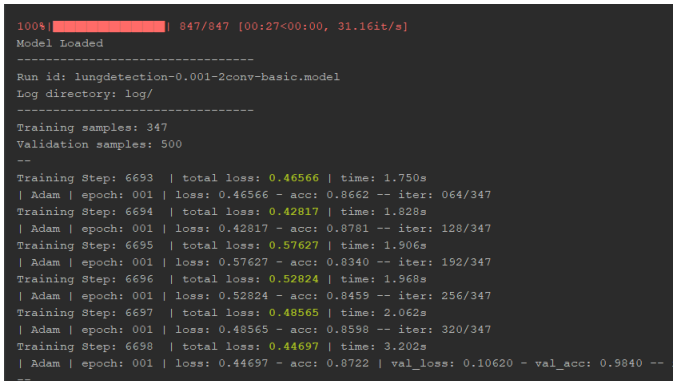


Figure 1 Training Model Loaded Successfully

Table 1 Modelling Table

	precision	recall	f1-score	support
0	1.00	0.83	0.91	81
1	0.93	1.00	0.96	68
2	0.96	1.00	0.98	292
3	1.00	0.97	0.98	59
accuracy			0.97	500
macro avg	0.97	0.95	0.96	500
weighted avg	0.97	0.97	0.97	500

Process finished with exit code 0

Formulas:

$$Accuracy = \frac{True\ positives + True\ Negatives}{All\ Samples}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

VII CONCLUSION

The objective of proposed system is to build "Predictive Diagnostic System" of infectious lung.

Proposed system works on the concept of image processing techniques such as feature subtraction, extraction, selection along with CNN classifier algorithms (machine learning) so as to precisely predict & detect various diseases of infectious lung.

As an integrated approach this system will not only predict and detect various lung diseases, but it will also generate test reports that can be used for a preventative treatment.

REFERENCES

[1] STUDY OF DETECTION OF VARIOUS TYPES OF CANCERS BY USING DEEP LEARNING: A SURVEY VINOD B BHARAT DR. NAVNEET MALIK 2019/8/31 International Journal of Advanced Trends in Computer Science and Engineering

[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.

Guide: Prof Laxmikant Malphedwar

Name: Shrikant N Borude
Email : shrikantborude8@gmail.com

Name: Saurabh R Kolhe
Email : kolhesaurabh23@gmail.com

Name: Harshal R Patil
Email : harshalpatil560@gmail.com

****