Detection of Lung Cancer through Image Data Mining Techniques - A Survey

[1]C. Jamunadevi, [2]M. Mangayarkarasi, [3]S. Deepa

[1,3]Assistant Professor, Kongu Engineering College, Perundurai-638 052, Tamiladu, India

[2]Assistant Professor (Selection Grade), Kongu Engineering College, Perundurai, Tamiladu, India

ABSTRACT

According to World Health Organization stated that, 7.6 million deaths globally each year are caused by cancer. Cancer represents 13% of all global deaths. National Cancer Institute stated that, by the end of 2015 there will be 221,200 new lung cancer diagnoses and 158,040 lung-cancer related deaths in the USA. These reports indicate that Lung cancer is the number one cancer killer. Due to this reason research of lung cancer lead to increase the survival rate of people. These literature paper illustrate the lung cancer identification in various characteristics.

I.   INTRODUCTION

Lung cancer is the uncontrolled growth of abnormal cells that start off in one or both lungs. The abnormal cells do not develop into healthy lung tissue; they divide rapidly and form tumors. Primary lung cancer originates in the lungs, while secondary lung cancer starts somewhere else in the body, metastasizes, and reaches the lungs. Manifestation of Lung cancer in the body of the patient reveals through early symptoms in most of the cases. [1]. Treatment and prognosis depend on the histological type of cancer, the stage (degree of spread), and the patient's performance status. Possible treatments include surgery, chemotherapy, and radiotherapy Survival depends on stage, overall health, and other factors, but overall only 14% of people diagnosed with lung cancer survive five years after the diagnosis.

The most common symptoms of lung cancer are:

- A cough that does not go away or gets worse.
- Coughing up blood or rust-colored sputum (spit or phlegm)
- Chest pain that is often worse with deep breathing, coughing, or laughing.
- Hoarseness.
- Weight loss and loss of appetite.
- Shortness of breath.
- Feeling tired or weak.

Lung cancer mostly affects people between the ages of 55 and 65 and often takes many years to develop [1]. In this work a study is carried out on various data mining techniques to detect lung cancer.

II.   RELATED WORKS

Prasanna et.al [2] illustrated the detailed survey about the lung cancer idenfiication. Kawsar Ahmed et.al [3] collected initially 400 cancer and non-cancer patients' data from

different diagnostic centers, pre-processed and clustered using a K-means clustering algorithm for identifying relevant and non-relevant data. Next significant frequent patterns are discovered using AprioriTid and a decision tree algorithm. Finally using the significant pattern prediction tools, lung cancer prediction systems were developed. This lung cancer risk prediction system should prove helpful in detection of a person's predisposition for lung cancer.

Conclusions: Most of people of Bangladesh do not even know they have lung cancer and the majority of cases are diagnosed at late stages when cure is impossible. Therefore early prediction of lung cancer should play a pivotal role in the diagnosis process and for an effective preventive strategy.

Rajneet Kaur [4] proposed data mining classification techniques such as neural network & SVMs for detection and classification of Lung Cancer in X-ray chest films. Due to high number of false positives extracted, a set of 160 features were calculated and a feature extraction technique was applied to select the best feature. They classify the digital X-ray films in two categories: normal and abnormal. The normal or negative ones are those characterizing a healthy patient. Abnormal or positive ones include types of lung cancer. Some procedures like Data Preprocessing, Feature Extraction etc. are also used. They use classification methods in order to classify problems aim to identify the characteristics that indicate the group to which each case belong.

The major cause of cancer related deaths is due to lung cancer. Lung cancer is caused by various abnormalities and one such abnormality is the lung nodule. When these lung nodules are detected at an early stage the survival rate is improved. CT image is having a large no of slices of images which makes the manual diagnosis a tedious process. It also takes a large time and energy of the radiologists. Hence an automatic approach for the detection of lung nodule is projected by Thangamani. M [5]. Zakaria Suliman Zubi et. al [6] used some data mining techniques such as neural networks for detection and classification Lung Cancer in X-ray chest films. 60 X-ray chest films multimedia database was taken as a training dataset. These set of images are considered 70 percent as a training value of the systems and 15 percent for testing it. Fifteen splits of the data collection were considered to compute all the results in order to obtain more accurate result of the system potential. A classification of the digital X-ray chest films was obtained into two categories: normal and abnormal. The normal ones are those characterizing a healthy patient. The abnormal ones could include types of lung cancer. Finally, classification, neural network methods were used to classify problems aiming at identifying the characteristics that indicate the group to which each case belongs. The system through all its processing steps will detect whether the supplied lung image had cancer or not. The results obtained from the classification process of the lung cancer image are 100% classification accuracy for normal images, 95% classification accuracy for begign images and 85% for malignant images.

Juliet R Rajan et. al [7] have planned a learning method based on unsupervised learning which can be used in building a predictive model for early detection of lung cancer. They also showed that Artificial Neural Networks can be used to predict the disease even with the occurrence of new symptoms. Also, the disease can be further analyzed by extracting the resultant weight vector after the training process.

CONCLUSION AND FURTHER DIRECTION

In this survey, the performance of different image data mining techniques is assessed. Many researchers have applied the data mining algorithms for predicting cancers especially Lung cancer. If studies on the detection of lung cancer continuous then it is likely that the use of data mining techniques for image data will become much more useful in diagnosing lung cancer.

REFERENCES

[1]. V.Krishnaiah , Dr.G.Narsimha, Dr.N.Subhash Chandra, Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques, Vol.3. No.3, pp. 131-134, 2013.

[2] V. Prasanna, M. Thangamani, "Survey on Lung Cancer Detection Using Data Mining Approach", International journal of Advanced Trends in Computer Applications, Vol1, No. 4, , pp. 5-8, April. 2015.

[3] Kawsar Ahmed, Abdullah-Al-Emran, Tasnuba Jesmin, Roushney Fatima Mukti, Md Zamilur Rahman, Farzana Ahmed, "Early Detection of Lung Cancer Risk Using Data Mining" *Asian Pacific Journal of Cancer Prevention, Vol 14, pp.595-598, 2013*

[4] Ada, Rajneet Kaur, "A Study of Detection of Lung Cancer Using Data Mining Classification Techniques" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3, pp.131-134, March 2013.

[5] Dr. M. Thangamani, Implementation of automatic detection of lung cancer using Adoptive Neuro Fuzzy System, International Journal of Advances in Science and Technology, Special issue, pp.28-32, December 2014

[6] Zakaria Suliman Zubi, Rema Asheibani Saad, "Improves Treatment Programs of Lung Cancer Using Data Mining Techniques" Journal of Software Engineering and Applications, Feb 2014, 7, 69-77

[7] Juliet R Rajan, Jefrin J Prakash, Early Diagnosis of Lung Cancer using a Mining Tool.