Analysis of various Lung Cancer Detection Methods

Parshad Suthar
Research Scholar
parshad.suthar@gmail.com

Abstract - The lung cancer detection is the combination of image processing and machine learning. The lung cancer detection technique has various phases which are pre-processing segmentation, feature extraction and classification. The various features are extracted like colour, textural features for the classification. The threshold based technique is the common type of technique which can be applied for the segmentation. In the last various algorithms of the classification can be applied which classify data into cancer and non cancer part. In this review paper, various techniques of lung cancer detection are reviewed in terms of certain parameters

Keywords - *Lung cancer, segmentation, feature extraction, classification*

I. INTRODUCTION

The highest projected incidence and mortality rate of each type of cancer and also the lowest 5-year net survival rates are included in the lung cancer. The cells are developing uncontrollably in the body during the occurrence of cancer. When cancer is commenced in the lungs, this disease is recognized as thelung cancer. Lung cancer leads to the death and it is considered as the second most detected cancer in people of US. This disease was increased for decades after that lung cancer rates are reduced nationwide becauseless people smoke cigarettes. The Cigarette smoking is one of the leading reasons of lung cancer. Other kinds of tobacco, breathing second hand smoke, the exposure of substances at home or work and occurrence of lung cancer in the family history are some of the other reasons leading to lung cancer. Initially,a collection of malignant cells is commenced in the lungs and then the lung cancer grows into thenearby tissues and destroys them [1]. Thecancerous tumours are developed by clusters of these malignant cells. It is essential to detect the lung cancer at the initial phase as most of the diagnoses take place in the later phases of the disease that include greater mortality rates. The lung cancer has various diverse categories and that categories are split further into two major groups namely SCLC and non-SCLC. There are a variety of medical imaging methods including X-rays, CT scans, MRI etc. to diagnose the lung cancer.

A number of systems are constructed and a research is conducted to detect the lung cancer. But there are some systems that do not provide theacceptable accuracy for detecting the lung cancerand some systems still requires enhancement for attaining the uppermost precision tending to 100%.Image processing techniques are popularly applied for the detection and classification of the lung cancer. Lung cancer detection based on image processing includes several steps such as imaging acquisition, pre-processing, segmentation, feature extraction. Pre-processing is an improvement of the image data thatsuppresses unwanted distortions or enhances some imagefeatures for further processing. The image Pre-processing phase begins with image enhancement.Imageenhancement aims to enhance the interpretability of the image information for analysts [2]. This process delivers better input for other automatic image processing methods.Image enhancement techniques can be divided into two broad categories of spatialdomain and frequency domain methods. The image enhancement is generally performed using three techniques, namely Gabor filter, Auto-enhancement and Fast Fouriertransform (FFT).The presentation of image depending on Gabor function creates an exceptional local and multiscale decomposition in terms of logons. These logons localize optimally, at once, both in space and frequency domains. A Gabor filter refers to a linear filter. A harmonic function multiplied by a Gaussian function defines the impulse response of this linear filter. In Gabor filter, the Fourier transform of impulse response denotes the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function.The second technique of auto enhancement is based on subjective observation andstatistical operations including mean and variance calculation to a large extent.The method of FFT (Fast Fourier Transform) performs on Fourier transform of a specified image [3]. The frequency domain represents a space wherein the value of every picture at image location F denotes the amount of intensity values in image “I”, that changes over a particular distance regarding F. The use of FFT (Fast Fourier Transform) method is quite common for the image filtering or image enhancement.The role of Image segmentation stage is very crucial for almost all following image analysis tasks. Especially, numerous available image description and recognition methods rely on the outcomes of segmentation stage to a large extent. Segmentation process does the division of the image into its integral parts or objects. The segmentation of clinical images in two-dimensional images per slice is very beneficial for the

doctors. Segmentation of medical image includes visualization and volume estimation of the required artifacts, detection of irregularities (e.g. tumours, polyps, etc.), tissue quantification and classification, and so on.Threshold segmentation is the most common segmentation technique [4]. The images that have variety in brightness levels use the Threshold-based segmentation. Here, depending upon the brightness level of the pixels, the images can be segmented into different regions.The largest interclass variance method called Otsu's is the most popular threshold segmentation algorithmic approach. This approach maximizes the variance amongst classes for selecting a globally best threshold. The major benefit of this approach lies in its simplified computation and faster operation speed. Especially, the segmentation impact can be achieved if the target and the background have high brightness level [5].The basic idea ofregion-based segmentation algorithms is to group neighbouring pixels and the pixels of similar values together and split groups of pixels of different values for operating in iterative manner.In the Marker-driven watershed segmentation method, the seeds are extracted in which objects or background presence is described at the particular locations of image. The regional minima set after the marker locations are placed in the topological surface. The watershed algorithm is carried out for the difficult operation of image processing that is separation of touching objects in an image.Image features extraction is one of the significant phases in which several desired parts or features are detected and separated of an image. There are two techniques available for predicting the lung cancer that has present. These techniques are binarization and masking that are based on the facts associated with the anatomy of lung and the information related to CT imaging of lung is also described in it [6].Binarization approach is based on the fact that in normal images of lung, the black pixels are higher than the white pixels. The average is acquired by counting the black pixels for images that are normal or abnormal. This average is considered as a threshold. The image is normal when the number of black images had found more than the threshold. On the other hand, a smaller number of thresholds is specified as the abnormal image.Masking approach is associated with the fact that the percent of cancer presence is increased with when the appearance of masses is similar to white areas that have connected inside region of interest [7]. The normal case is described by the blue colour appearance and the cancer presence is shown when the RGB masses are appeared. The case is normal or abnormal, this decision is taken on the basis of the Binarization and Masking approaches. These two approaches integrated and decision is taken on the behalf of assumptions of these approaches. The image is considered abnormal if the black pixels are greater in number as compared to white pixels but is it considered abnormal the numbers of black pixels are less.

II. LITERATURE REVIEW

Mesut Toğaçar, et.al (2020) suggested ahybrid model on the basis of 6 diverse ML algorithms, 3 Convolutional Neural Networks modelsand mRMR feature selection technique for detecting the lungcancer [8].An open dataset in which CT images were contained had employed to conduct the experiments. The features were extracted and the classification was performed using CNNs in the experiments. The AlexNet model was integrated with the K-Neural Network algorithms and the most efficientaccuracy obtained for classification was 98.74 %. Afterward, the effective attributes were selected by the means of mRMR technique in thedeep feature set. At last, the obtained success rate was computed as 99.51 % when the dataset was re-classified with the selected featuresand K-NN model.

Nidhi S. Nadkarni, et.al (2019) recommended anautomated approach to detect the lung cancer in Computed Tomography scan images [9]. Various techniques were utilized to suggest analgorithm while detecting lung cancer. This involved median filtering to pre-process the image after that the mathematical morphological operations were executed to segment the lung ROI.Median filtering method was proved efficient to remove the impulse noise from the images without blurring the image. Thelung tumour region was segmented accurately usingMathematical morphological operations. The evaluation ofgeometrical attributes was done from the extracted ROI and implemented forcategorizing the CT scan images into normal and abnormal through SVM.

ÖzgeGünaydin, et.al (2019) discussed that the ML techniques were compared during the detection oflung cancer nodule [10]. An anomaly was detected by carrying out PCA, K-NN, NB, DT and ANNs. The comparison of all these ML techniques was carried out subsequent to thepre-processing and without pre-processing. The outcomes of experiments demonstrated that the best accuracy obtainedafter image processing using ANN was evaluated 82.43% and without image processing, the best accuracy of 93.24% was obtained from DT.In future, several noise removal techniques would carry out for obtaining superior accuracy.

R. Sathishkumar, et.al (2019) described that the Support Vector Machine and K-NN had implemented to detect the lung cancer [11]. First of all, the pre-processing ofmedical image or a CT scan image was carried out. The CLAHE Equalization method was executed in order to increase the contrast of the image. Subsequently,random walk segmentation technique was applied to segment the image.

During thesegmentation, there were 3 procedures executed. Firstly, the Region of Interest of image was segmented and the border was corrected later on. Lastly,the continuous pixel change was segmented.The classification was one of the main portions in whichpre-trained model had utilized to determine which part was cancerous and non-cancerous. The above-mentioned techniques had utilized theconventional way to process the image and data analytics. For future work, themodern XGboost algorithm would be implemented to enhance the accuracy using less data.

JaneeAlam, et.al (2018) presented an effectual LCD andprediction algorithm for which SVM algorithm was implemented [12]. The cancer was detected usingMulti-stage classification.This system was also capable for predicting the probability of lung cancer. The imageenhancement and image segmentation were performed independently under each phase ofclassification. The image was scaled, colour space was transformed and the contrast was improved in the image enhancement. The image was segmented by applyingsegmentation based on threshold and marker-controlled watershed. The Support Vector Machine binary algorithm was carried out to perform the classification. It was indicated that the presented method provided the superior precision to detect and predict the lung cancer.

Lingling Li, et.al (2018) intended a novel system to detect the lung cancer on the basis of CT technology [13]. The conversion ofRGB images was done into grayscale images, the noises were uprooted using the median filter and the Wiener filter, theCT images which were free from noise converted into binary images with the help ofOtsu thresholding technique andbody region exacted in binary images through REGIONPROPS function under the pre-processing phase. The statistic technique named GLCM was implemented for extracting a number of attributed in the next phase. The lung cancer was recognized from CT images by executing theextracted attributes along with SVM and BPNN. The adequate outcomes were achieved from the intended system and the obtained accuracy on Support Vector Machine was computed 96.32% and the accuracy on BPNN was 83.07%.

Qing Wu, et.al (2017) suggested a new algorithm based on NN that was recognized as EDM for detecting SCLC from the CT images [14]. Thisresearch assisted in detecting the lung cancers at initial phase. TheNational Cancer Institute offered thetraining data and testing data containing high-resolution lung CT scans. There were twelvecomputed tomography scans of lung were taken from the library, six scans were takenfrom healthy lungs and the remaining six were from patients having SCLC. Five scans were captured at randomfrom every group for training the suggested model and the rest of the scans were carried out in testing. The obtained precision using suggested algorithm was computed as77.8%.

Author	Year	Description	Outcomes
Mesut Toğaçar, Burhan Ergen, Zafer Cömert	2020	Suggested ahybrid model on the basis of 6 diverse ML algorithms, 3 Convolutional Neural Networks modelsand mRMR feature selection technique for detecting the lungcancer.	The obtained success rate was computed as 99.51 % when the dataset was re-classified with the selected featuresand K-NN model.
Nidhi S. Nadkarni, Sangam Borkar	2019	Recommended anautomated approach to detect the lung cancer in Computed Tomography scan images.	The evaluation ofgeometrical attributes was done from the extracted ROI and implemented forcategorizing the CT scan images into normal and abnormal through SVM.
ÖzgeGünaydin,MelikeGünay, ÖznurŞengel	2019	Discussed that the ML techniques were compared during the detection oflung cancer nodule. An anomaly was detected by carrying out PCA, K-NN, NB, DT and ANNs.	The outcomes of experiments demonstrated that the best accuracy obtainedafter image processing using ANN was evaluated 82.43% and without image processing, the best accuracy of 93.24% was obtained from DT.

R. Sathishkumar, K. Kalaiarasan, A. Prabhakaran, M. Aravind	2019	Described that the Support Vector Machine and K-NN had implemented to detect the lung cancer.	The classification was one of the main portions in whichpre-trained model had utilized to determine which part was cancerous and non-cancerous. The above-mentioned techniques had utilized theconventional way to process the image and data analytics.
JaneeAlam, Sabrina Alam, Alamgir Hossan	2018	Presented an effectual LCD andprediction algorithm for which SVM algorithm was implemented. The cancer was detected usingMulti-stage classification.	It was indicated that the presented method provided the superior precision to detect and predict the lung cancer.
Lingling Li, Yuan Wu, Yi Yang, Lian Li, Bin Wu	2018	Intended a novel system to detect the lung cancer on the basis of CT technology.	The adequate outcomes were achieved from the intended system and the obtained accuracy on Support Vector Machine was computed 96.32% and the accuracy on BPNN was 83.07%.
Qing Wu, Wenbing Zhao	2017	Suggested a new algorithm based on NN that was recognized as EDM for detecting SCLC from the CT images.	The obtained precision using suggested algorithm was computed as77.8%.

III. CONCLUSION

In this paper, it is concluded that lung cancer detection is the major challenge of machine learning due to dynamic type of input data. The various algorithms are reviewed for the lung cancer detection which are used on the different phases. The different phases of lung cancer detection include feature extraction, segmentation and classification. In future, ensemble learning method can be designed for the lung cancer detection

IV. REFERENCES

[1] Shradha fule, "Lung Cancer Detection Using Image Processing Techniques", 2017, International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 12

[2] Ayushi Shukla, Chinmay Parab, Pratik Patil, Prof. Savita Sangam, "Lung Cancer Detection using Image Processing Techniques", 2018, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04

[3] BariqiAbdillah, AlhadiBustamam, and DevviSarwinda, "Image processing-based detection of lung cancer on CT scan images", 2016, The Asian Mathematical Conference

[4] FebrMokhled S. AL-TARAWNEH, "Lung Cancer Detection Using Image Processing Techniques", 2012, Leonardo Electronic Journal of Practices and Technologies.

[5] Muhammad Usman, Muhammad Shoaib and Mohamad Rahal, "Lung Cancer Detection Using Digital Image Processing", 2013, PIERS Proceedings, Stockholm, Sweden

[6] Anita Chaudhary and SonitSukhraj Singh, "Multiresolution Analysis Technique for Lung Cancer Detection in Computed Tomographic Images", 2012, International Journal of Research in Engineering & Applied Sciences, IJREAS Volume 2, Issue 2

[7] Disha Sharma, Gagandeep Jindal, "Computer Aided Diagnosis System for Detection of Lung Cancer in CT Scan Images", 2011, International Journal of Computer and Electrical Engineering, Vol. 3, No. 5

[8] Mesut Toğaçar, Burhan Ergen, Zafer Cömert, "Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks", 2020, Biocybernetics and Biomedical Engineering

[9] Nidhi S. Nadkarni, Sangam Borkar, "Detection of Lung Cancer in CT Images using Image Processing", 2019, 3rd International Conference on Trends in Electronics and Informatics (ICOEI)

[10] ÖzgeGünaydin,MelikeGünay, ÖznurŞengel, "Comparison of Lung Cancer Detection Algorithms", 2019, Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)

[11] R. Sathishkumar, K. Kalaiarasan, A. Prabhakaran, M. Aravind, "Detection of Lung Cancer using SVM Classifier and KNN Algorithm",2019, IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)

[12] JaneeAlam, Sabrina Alam, Alamgir Hossan, "Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVMClassifier", 2018, International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)

[13] Lingling Li, Yuan Wu, Yi Yang, Lian Li, Bin Wu, "A New Strategy to Detect Lung Cancer on CT Images", 2018, IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)

[14] Qing Wu, Wenbing Zhao, "Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm", 2017, International Symposium on Computer Science and Intelligent Controls (ISCSIC)