IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A COMPREHENSIVE SURVEY ON EARLY DETECTION OF CANCEROUS TUMOR IN THE LUNGS

[1]Shafiulla Shariff, [2]Dr. K.G. Avinash

[1]Assistant Professor, [2]Associate Professor
[1]Department of Computer Science & Engineering,
[1]Jain Institute of Technology Davangere, Visveswaraya Technological University,Belagavi, India

Abstract: Among all disease Pulmonary cancer has repeatedly appeared as most crucial diseases in the history of mankind. It ranks among the frequently occurring types of cancer and a leading cause of death. The disease frequently shows no symptoms in its initial stages, which makes it very difficult to detect early. This cancer is characterized by the uncontrolled growth of cells within the lung tissue. While lung cancer cannot be entirely prevented, the likelihood of developing it can be reduced. As a result, Early detection of lung cancer is crucial. for patient survival. In this review, we have tried to look at a variety of artificial intelligence techniques for identifying lung cancer., as well as propose a new Artificial Intelligence strategy for identifying subtle patterns and delivering reliable information to pathologists.

***Index Terms* -** Artificial intelligence AI, Classifier, Data Set.

I. INTRODUCTION

Cancer is defined as an uncontrolled cell proliferation within the body. Unchecked cell growth in lung tissue leads to the development of lung cancer. After HIV, Lung cancer ranks as the second most common cause of death globally, resulting in nearly a million fatalities annually.. According to an estimate, around 33% of the world's population is affected by lung cancer., and an estimated nine million new cases are occurring every year. Lung cancer is caused primarily by cigarette smoking. According to the data [3], smoking is responsible for 80% of lung cancer deaths.. Tuberculosis may elevate the risk of developing lung cancer by generating substantial and long-term inflammation in the lungs, which can result in whole tissue damage, fibrosis, scarring, and genetic changes. Lung cancer predominantly affects men, comprising approximately, accounting for roughly one-third of all malignancies in men. Lung cancer is still the most common cause of cancer death according to GLOBOCAN 2020 [3] with an anticipated 1.8 million deaths (18%). Table 1 below shows recent cases and Deaths for major Cancers.

Table 1: Number of fresh cases and Deaths of 3 types of Cancers

Cancer	New Cases (%)	New Deaths (%)
Breast Cancer	11.70	6.9
Lung Cancer	11.40	18
Prostate Cancer	7.30	3.8

Figure 1 below shows the top 3 kinds of cancer estimated cases and deaths worldwide.

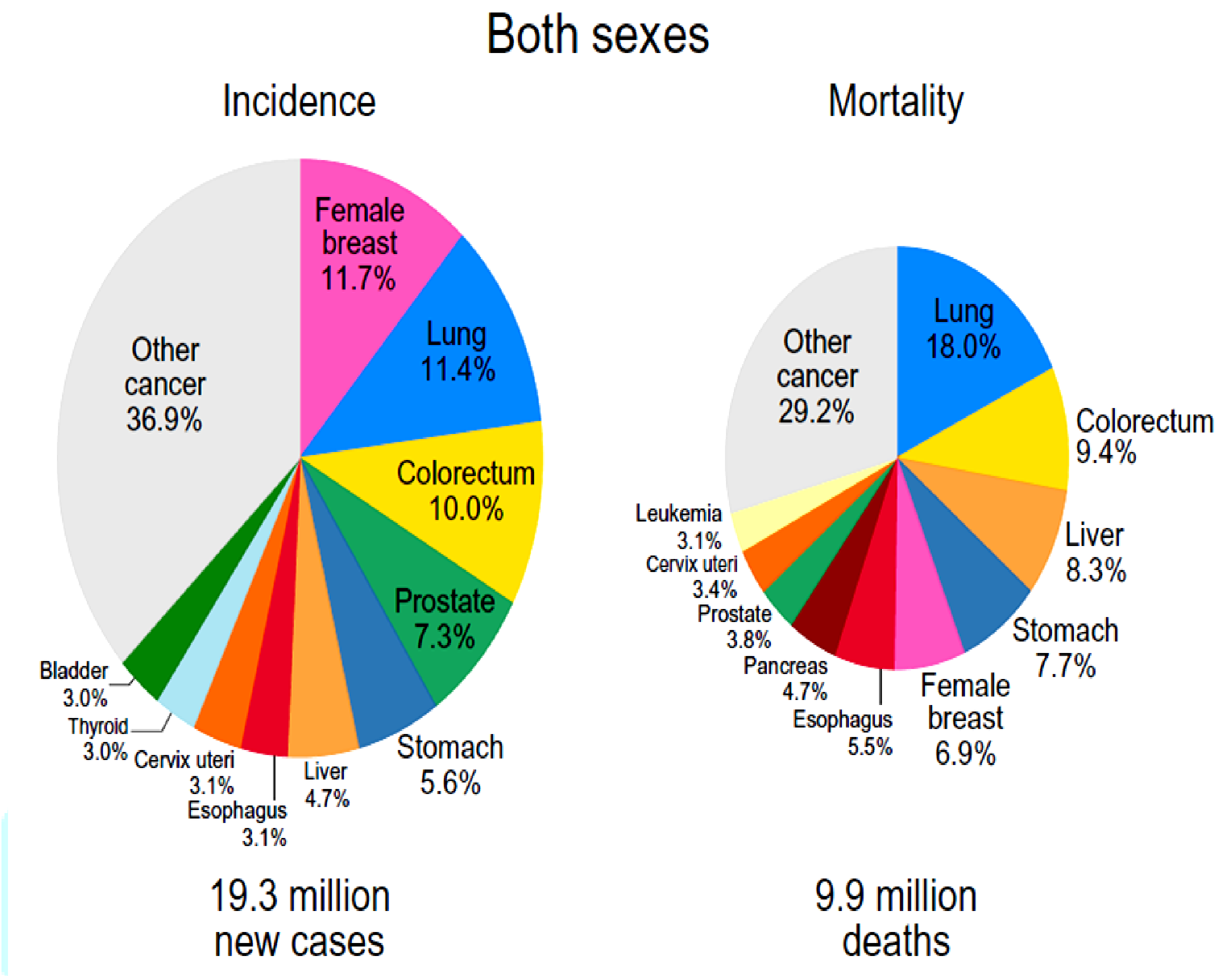

Figure 1: The Top 3 Most Common Cancers: Cases and Deaths Distribution in 2020.

The prognosis for lung cancer patients depends significantly on the stage of the disease at the time of diagnosis. Therefore, detection of pulmonary cancer in its initial phases is critical. Identifying lung cancer in its initial stages can significantly help reduce the death rate. In healthcare engineering, medical imaging has become a crucial area of study with the aim of helping medical professionals identify diseases early on.

II. Related Work

In 2022, Akitoshi Shimazaki et al. created and assessed a deep learning (DL) model aimed at detecting pulmonary cancer in chest radiographs through a segmentation method [2]. They trained and validated the DL model on a training dataset using a five-fold cross-validation method. The training dataset included 629 radiographs containing 652 nodules/masses. Subsequently, they evaluated the model's sensitivity and mean false positive indicators per image (mFPI) using an independent test dataset comprising of 151 radiographs with 159 nodules/masses. The DL-based model exhibited a mean false positive indicator (mFPI) of 0.13 and a sensitivity of 0.73 in the test data. Interestingly, it was observed that lung tumors located in overlapping blind spot regions like the pulmonary apices, pulmonary hila, chest wall, heart, and subdiaphragmatic space had lower sensitivity (ranging from 0.50 to 0.64). Specifically, sensitivity was reduced for lung tumors in these overlap regions compared to non-overlapping regions. For the 159 malignant lesions, the average Dice coefficient was 0.52. Despite challenges in overlapping areas, the DL-based model demonstrated its capability to detect lung tumors in chest radiographs, achieving a low mFPI.

In 2022, Sharmila Nageswaran et al. conducted an experimental study using a dataset comprising 83 CT scans from 70 patients[1]. To enhance image quality, the geometric mean filter was utilized during the preprocessing phase. This was followed by image segmentation using the K-means method to identify specific regions within the images. Machine learning classification algorithms, such as ANN, KNN, and RF, were subsequently employed. The findings indicated that the The ANN model performed better than other machine learning methods in predicting lung cancer with higher accuracy.

In 2020, Meza Silvana et al. introduced an algorithm for detecting pulmonary cancer using CT scan images. Their system utilizes basic image processing techniques, including Lung Border Extraction, Erosion, Dilation, Median Filtering, and Outlining. This method notably decreases false positives cancer detections and achieves an acc of 89%, though it still presents a notable false positive rate.

In 2019 Sarfaraz Hussein et al proposed both machine learning strategies supervised learning and the unsupervised learning to enhance the tumor characterization, first method is by using 3D Convolution Neural Network and the Transfer Learning and in second method with inspiration of learning from label proportion (LLP) approaches in computer vision [5].
In 2018 Muhammad et al proposed a method in order to improve lung cancer detection efficiency via their signs and symptoms, The authors aim to evaluate the discriminative power of various predictors in the study [6]. Various classifiers are implemented on a set of data which were obtained from UCI repository. Classifiers used are C4.5 Decision tree, Support Neural Network, Vector Machine (SVM), Naïve Bayes and Multi-Layer Perceptron. The performance was also compared with Random Forest and Majority Voting methods.Based on performance evaluations, it is observed that Gradient-boosted Tree performed well compare to other ensemble classifiers.

In 2018, IEEE Janee Alam et al. proposed a method for diagnosing and predicting lung cancer using a multi-class Support Vector Machine (SVM) classifier.. Multi-stage classification was employed in this cancer detection method. This method can also be used to forecast the risk of lung cancer. For every classification step, segmentation and image quality enhancement have been carried out independently. Scaling, color space modification, and contrast enhancement have all been done to enhance the image quality. Utilizing the marker-controlled watershed approach and segmentation threshold, an SVM binary classifier was employed for classification.

In 2016 Arnaud A. A. et al gave a description of the proposed design, that includes multiple layers of 2-D ConvNets, to get the final classification, the outputs are grouped and combined with the help of fusion technique [8].

In 2016 Junyuan Xie et al proposed a model using unsupervised learning, it explains about clustering [11]. Clustering is a key component of many data-driven application fields and has been extensively studied in terms of grouping techniques and distance functions. There hasn't been much focus on learning representations for grouping. On the other hand, misclassifying any image does not provide an estimated outcome.

Table below shows the comparison of the different methods using accuracy and sensitivity difference for lung nodule characterization [[2]3][4][5].

Table 2: Different classifiers used for detection of lung cancer

Sl no.	Classifiers	Accuracy %	Sensitivity %
1.	ANN backpropagation	89	80
2.	Clustering and RF	76.74	58.59
3.	Clustering and SVM	76.04	57.08
4.	SVM GIST	81.56	71.31
5.	GIST features and LASSO	76.8	67.5
6.	GIST features and RR	76.4	67.44
7.	3D CNN and LASSO	86	50.30
8.	3D CNN and RR	88	49.7
9.	3D CNN with Multi task learning	91	45.9
10.	Auto MLP	78.33	68

11.	Naïve Bayes	85.00	50.1
12.	Decision tree	78.33	68.3
13.	Random forest	79.17	65.4
14.	Gradient boosted tree	90	50

Figure 2 presents a histogram illustrating the accuracy and sensitivity of different machine learning methods employed for pulmonary cancer detection. The data were compiled from keyword searches conducted across PubMed, CiteSeer, Google Scholar, the Science Citation Index, and other online sources.

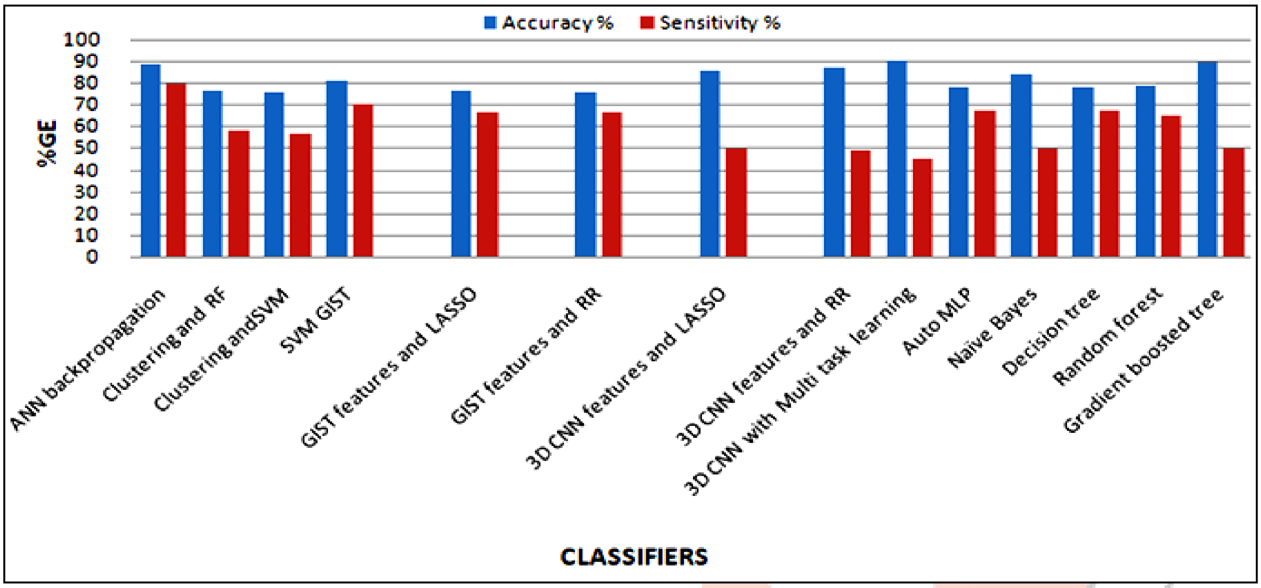

Figure 2: Histogram showing the accuracy and sensitivity of various M L- techniques

III. PROPOSED DESIGN PROCESS

The pulmonary cancer is predicted to become first deadliest disease by coming years. As mentioned, detecting a tumor in its initial stages is highly challenging. If it tumor is detected in its beginning stages, the cancer can be cured and mortality rate can be decreased. Radiographic tests for lung cancer are accessible, yet they are not effective in detecting cancer during its initial phases. In this regard we are planning to design a framework using AI with invasive/non invasive methods using biomarkers and images to predict/detect lung cancer in initial stages. Figure 3 below illustrates the design process for detecting cancer in its early stages.

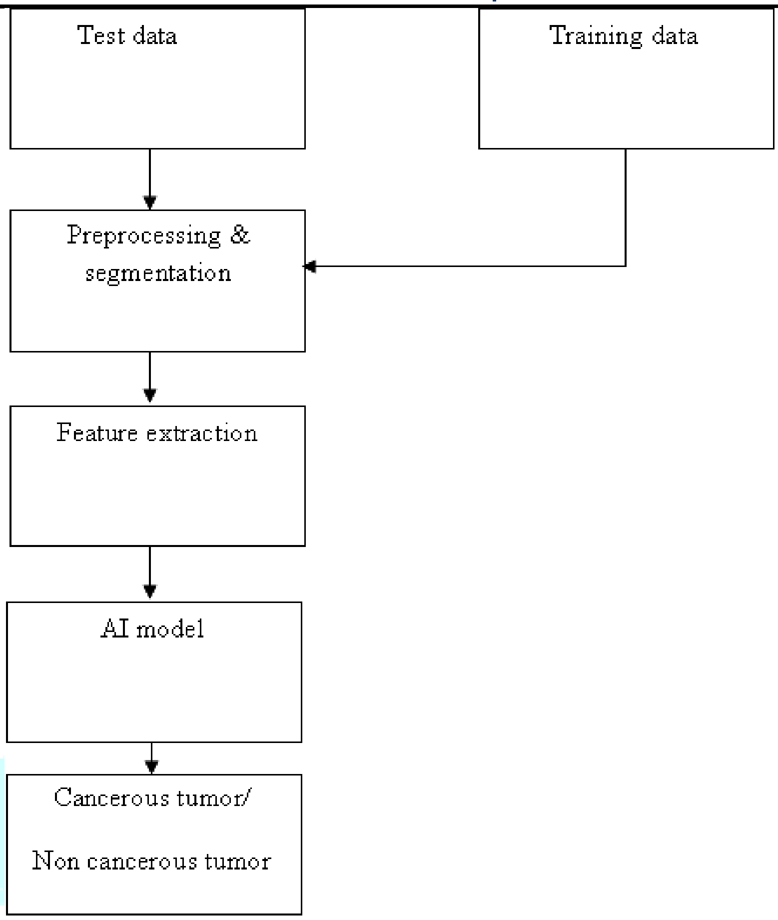

Figure 3: The proposed approach for detecting lung cancer.

Preprocessing: Most images are affected by artifacts from contrast variations and noise. Pre-processing techniques are used to reduce these artifacts while maintaining image details.

Image segmentation: This involves partitioning an image into several segments. This approach is used to identify objects and delineate boundaries within the image. It is a fundamental step in numerous subsequent image analysis techniques.

Feature extraction: This is a vital phase where algorithms and techniques are to identify and isolate specific desired parts or shapes within an image. Many algorithms are utilized in image processing to assess the normality or abnormality of an image depending on the segmentation outcome. Key parameters that help in classifying cancerous regions include area, perimeter, eccentricity, and average intensity.

The AI models implemented for classification using the given training data to detect the cancerous tumor. And the corresponding models are evaluated using test data to get the good accuracy of the models.

IV. CONCLUSION

We know that Now-a-days, Cancer is the leading cause of death in many populations.. Lung cancer fatality cases are steadily increasing among them. According to this review, we concluded that there was no notable improvement in sensitivity, the no. of false positives, or the automation capability for nodule detection during the studied period. However, the CAD system enhances automation levels when integrated with the Picture Archiving and Communication System (PACS). The survey reveals that many methods have been utilized to detect early-stage lung cancer, but employing deep learning models can enhance accuracy in predicting and detecting lung cancer at Stages I and II.

REFERENCES

[1] Sharmila Nageswaran , G. Arunkumar , Anil Kumar Bisht , Shivlal Mewada , J. N. V. R. Swarup Kumar , Malik Jawarneh , and Evans Asenso, Lung Cancer Classification and Prediction Using Machine Learning and Image Processing,2022, 2022 Hindawi,Biomed research International.

[2] Akitoshi Shimazaki1, Daiju Ueda1,2*, Antoine Choppin3, Akira Yamamoto1, Takashi Honjo1, Yuki Shimahara3 & Yukio Miki1 Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method, 2022 scientific reports.

[3] Hyuna Sung Jacques Ferlay Rebecca L. Siegel Mathieu Laversanne Isabelle Soerjomataram Ahmedin Jemal, Freddie Bray .Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries 2021

[4] Meza Silvana Ricky Akbar Hesti Gravina Firdaus. Optimization of characteristics using Artificial Neural Network for Classification of Type of Lung Cancer 2020 IEEE.

[5] Sarfaraz Hussein, Pujan Kandel, Candice W. Bolan, Michael B. Wallace, and Ulas Bagci, Lung and Pancreatic Tumor Characterization in the Deep Learning Era: Novel Supervised and Unsupervised Learning Approaches 2019 IEEE.

[6] Muhammad Imran Faisal, Saba Bashir, Zain Sikandar Khan, Farhan Hassan Khan., An Evaluation of Machine Learning Classifiers and Ensembles for Early Stage Prediction of Lung Cancer , 2018 IEEE.

[7] Janee Alam1, Sabrina Alam2, Alamgir Hossan., Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classifier, 2018 IEEE.

[8] Arnaud A. A. Setio, Francesco Ciompi, Geert Litjens, Paul Gerke, Colin Jacobs, Sarah J. van Riel, Mathi"Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks"(2016).

[9] Junyuan Xie, Ross Girshick "Unsupervised Deep Embedding for Clustering Analysis"(2016).

[10] Tiantian Fang., A Novel Computer-Aided Lung Cancer Detection Method Based on Transfer Learning from GoogLeNet and Median Intensity Projections, 2018 IEEE.

[11] De-Ming Wong ; Chen-Yu Fang ; Li-Ying Chen ; Chen-I Chiu ; Ting-I Chou ; Cheng-Chun Wu Shih-Wen ChiuKea-Tiong Tang., Development of a breath detection method based E-nose system for lung cancer identification, 2018 IEEE

[12] Bohdan Chapaliuk, Yuriy P. Zaychenko., Deep Learning Approach in Computer-Aided Detection System for Lung Cancer, 2018 IEEE.

[13] Moffy Vas, AmitaDessai ., Lung cancer detection system using lung CT image processing, 2017 IEEE.

[14] Qing Wu and WenbingZhao ., Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm, 2017 IEEE.

[15] Lilik Anifah , Rina Harimurti , ZaimahPermatasari ,PuputWanartiRusimamto , Adam Ridiantho Muhamad., Cancer lungs detection on CT scan image using artificial neural network backpropagation based gray level coocurrence matrices feature, 2017 IEEE

[16] Ravindranath K, K Somashekar.,Early Detection of lung cancer by nodule extraction– A Survey, 2017 IEEE.

[17] Ryota Shimizu, ShusukeYanagawa, Yasutaka Monde, Hiroki Yamagishi, Mototsugu Hamada, Toru Shimizu, and Tadahiro Kuroda., Deep Learning Application Trial to Lung Cancer Diagnosis for Medical Sensor Systems 2016 IEEE

[18] Rachid Sammouda., Segmentation and Analysis of CT Chest Images for Early Lung Cancer Detection, 2016 IEEE.

[19] K.Punithavathy, M.M.Ramya, SumathiPooba., Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images, 2015 IEEE.

[20] Fatma Taher ,NaoufelWerghi ,Hussain Al-Ahmad., Rule based classification of sputum images for early lung cancer detection, 2015 IEEE

[21] Bhagyashri G. Patil, Prof. Sanjeev N. Jain., Cancer Cells Detection Using Digital Image Processing Methods 2014 IEEE.

[22] FatmaTaher, NaoufelWerghi, Hussain Al-Ahmad, RachidSammouda., Lung Cancer Detection Using Artificial Neural Network and Fuzzy Clustering Methods, American Journal of Biomedical Engineering 2011.

[23] Badrul Alam Mia, Mohammad Abu Yusuf , "Detection of lung cancer from CT image using image processing and neural network", International conference on Electrical Engineering and Information Communication Technology (ICEEICT) ,IEEE ,May, 2015.

[24] Nooshin Hadavi, Md Jan Nordin, Ali Shojaeipour , "Lung cancer diagnosis using CT-scan images based on cellular learning automata", International conference on Computer and Information Sciences(ICCOINS), IEEE , 2014.