



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 11, November 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

A Systematic Review on Lung Cancer Prediction Methods

Devashish Sharma, Shanu Gour

M.Tech Scholar, Department of Computer Science and Engineering, Bharti Vishwavidyalaya, Bhilai,
Chhattisgarh, India

Assistant Professor, Department of Computer Science and Engineering, Bharti Vishwavidyalaya, Bhilai,
Chhattisgarh, India

ABSTRACT: Lung cancer is a form of volatile cancers in addition to tough to discover. It normally motives lack of lifestyles for each gender guys & ladies consequently, so it's far greater want for care to right away & effectively test nodules. Accordingly, several techniques had been completed to come upon lung cancer inside the early ranges. In this paper a comparative evaluation of numerous strategies based on gadget getting to know for detection lung cancer were supplied. There are techniques advances in overdue years to stumble on lung cancers, most of them the usage of CT scan takes a look at photographs in addition to some of them using x-ray pictures. In addition, multiple classifier techniques are connected with diverse distribution techniques to use image recognition to turn out to be privy to lung cancer nodules. From this examine it is been located that CT scan pictures are greater suitable to have the ideal effects. Therefore, widely speaking CT scan images are utilized for classification of cancer. In Addition, the effects that acquired from the strategies primarily based deep learning techniques achieved better accuracy than the methods that have been carried out the usage of classical machine learning techniques.

KEYWORDS: Lung Cancer Classification, Machine Learning, Deep Learning, CT scan.

I. INTRODUCTION

Lung cancers is a form of cancers that advances inside the lungs and is one of the usual and dangerous types of cancers globally. The usage of detection algorithm to perceive humans who've an excessive contingency of obtaining lung cancer has become more popular nowadays. Prediction methods may also moreover assist become aware of those who want screening or extra information, early analysis and intervention are critical for improving affected individual consequences [1].

Lung cancer is a massive public health problem worldwide and traces a number of the leading 5 primary motives of deaths due to maximum cancers. The World Health Organization report that lung cancers calculated for 11.6 percent of all most cancers instances and 18.4 percent of most cancers-associated death. The 365 days 2020 have become projected to look 2.2 million new instances of lung cancer globally, resultant in 1.8 million deaths. The prevalence of lung cancers varies broadly during special regions, with Europe and North America having the very best charges on the equal time as Africa and Asia have the bottom. This discrepancy is essentially attributed to variations in smoking pervasiveness, it is an essential immediate for lung cancer.

As a result, the acceptance of smoking ending programmes and extra awareness towards the dangers of smoking, the prevalence of lung cancer has little by little decreased in lots of industrialised international locations. However, lung cancers remain a big health task, especially in growing countries in which smoking rates maintain to boom [3]. The international burden of lung cancer highlights the want for persevered efforts to grow powerful prevention and early detection techniques, further to modern remedy strategies to enhance affected individual effects.

Predictive analysis for lung cancer employs diverse kinds of information, which consist of medical, genetic, as well as imaging information, so one can shape fashions which could predict efficiently the opportunity of growing lung cancer. In advance of characteristics appear, those fashions may be capable of choose out individuals at an immoderate hazard for lung cancer, ensuing in early intervention and higher survival charges [4]. In this context, the progression of correct and dependable prediction method for lung cancer carries vivid covenant for enhancing affected person attention and

lowering the weight of this illness. In spite of enhancing attainability of medical information and encourages in machine learning, the capability for emerging correct prediction method for lung cancer is growing rapidly.

Lung cancer predictive model utilizes of machine learning methods has enhanced considerable reputation nowadays. Many machine-learning algorithms are utilized to grow to be privy to patterns and association inside big datasets, taking into consideration the growth of correct predictive methods. Several machine learning algorithms, together with logistic regression, decision tree, and support vector machines, were used for lung cancer prediction [5]. These algorithms make use of numerous scientific and imaging data, inclusive of age, smoking history, tumor size, and tumor location, to increase prediction methods that can effectively find out people at huge chance of growing lung cancer.

Deep learning methods, a form of machine learning, are also being executed to the prediction of lung cancers. These models utilize artificial neural networks to research and perceive patterns in big datasets. Convolutional neural networks (CNNs), is a form of Deep Learning method regularly utilized for image classification obligations, were utilized to investigate medical pictures, which incorporates CT scans, to come across lung cancer [6].

The utilization of machine learning and deep learning methods for prediction of lung cancer has validated promising consequences, with a few research reporting huge accuracy-rate. These models also can boom the recognition precision of lung cancers and allow focused on screening and remedies. In spite of that there are challenges also attached with the growth to implement the methods, such as the want for huge variety of datasets and the capacity for bias within the statistics and methodology. More research is wanted to determine out the way to deal with the ones issues and result that machine learning and deep learning models can be used in medical steerage to expect lung cancers. In this paper, lung cancer prediction methods are reviewed as well as analysed.

II. LITERATURE SURVEY

Naseer et al. [7], in this paper, author proposed lung cancer classification makes use of modified U-Net based totally lobe segmentation and nodule detection model including 3 levels. The first section segments lobe the use of CT slice and anticipated masks the usage of modified U-Net architecture and the second one segment extracts candidate nodule the use of predicted masks and label using modified U-Net architecture. Finally, the third section is based totally on modified AlexNet, and a help vector system is carried out to categorise candidate nodules into cancer and non-most cancers. The changed AlexNet-SVM type version achieves 97.98% of accuracy, 98.84% of sensitivity, 97.47% of specificity, ninety seven.Fifty three% of precision, and ninety seven.70% of F1 for the class of lung cancer.

Sreeprada et al. [8], in this paper, author proposed a hybrid OCNN-SVM method for classifying lung CT images is provided. The goal changed into to increase accuracy because that's what any computer-aided detection device strives for. The dataset of chest X-ray pictures changed into used to use the method. The classification accuracy percent was 98.70%. It demonstrates the effectiveness of the recommended hybrid OCNN-SVM and its applicability. Additionally, the keep in mind, precision, and F1 scores of the counseled OCNN-SVM approach are 98.34%, ninety eight.49%, and 98.Seventy six, respectively.

Shailesh et al. [9], in this paper, author shows region of interest (ROI) is calculated using region-based segmentation algorithm. Circle fit algorithm is used to extract the desired nodule. Radius, Mean Intensity, Area, Euler Number and ECD features are extracted in feature extracting step. Finally, Back propagation algorithm is used to train Artificial Neural Network (ANN) in categorization stage.

Q. Firdaus et al. [10], In this paper, author layout to advances a lung cancer classification system evolved on CT-Scan photos. This recognition technique has 4 main steps, which include preprocessing of CT-Scan pictures to beautify image features, section to locate and distribute the obtained cancer objective from the encircling, traits extraction based totally on location, divergence, entropy, power, as well as homogeneity. The categorization of lung cancer into benign or malignant cancer. From the architecture trial, the efficiency level based totally on the structure selection in obtaining the prognosis of lung cancer is benign or malignant became 83.33%.

Bhandary et al. [11], These paintings provide you with two different deep learning practices to assess the thought about trouble: (I) The underlying DL strategy, named a modified AlexNet (MAN), is achieved to order chest X-

Beam pix into ordinary and pneumonia elegance. In the MAN, the grouping is accomplished using with Support Vector Machine (SVM), and its presentation is concept approximately in opposition to Softmax. Further, its exhibition is authorized with other pre-prepared deep getting to know strategies, like AlexNet, VGG16, VGG19 and ResNet50. (ii) The following deep gaining knowledge of process includes out a aggregate of hand-crafted as well as discovered functions within the MAN to future advances the characterization precision all through cellular breakdown inside the lungs assessment. The exhibition of this DL shape is attempted through the benchmark cell breakdown inside the lungs CT pics of LIDC-IDRI and general order exactness of >97.27% is finished.

Shakeel et al. [12], In this paper, the accrued photographs are inspected by way of applying the staggered splendor protective method which without a doubt seems at every pixel, takes out the commotion and furthermore increment the nature of the lung image. From the clamor removed lung CT picture, impacted locale is sectioned via regarding similarly advanced deep neural network that portions district as a ways as making use of layers of enterprise and special highlights are separated. Then the powerful highlights are chosen with the help of cross breed twisting streamlining smart summed up harsh set method, and people factors are arranged utilising ensemble algorithm. The talked about approach builds the cell breakdown within the lungs most cancers prediction charge which is analyzed utilizing MATLAB-primarily based effects, for instance, logarithmic, accuracy, overview and F-rating.

Shakeel et al. [13], This paper manages improvement of the character of lung photograph in addition to finding of cellular breakdown in the lungs by using lessening misclassification. Different ghastly highlights are gotten from the impacted district. These are analyzed with the aid of applying profound advancing without delay organized neural community for expecting mobile breakdown inside the lungs. In the cease, the framework is analyzed by using the proficiency of the framework utilizing MATLAB based reproduction effects. The framework ensures that ninety eight.42% of precision with least grouping mistake 0.038.

Reddy et al. [14], The proposed shape includes diverse means, for example, photograph getting, pre coping with, department, encompass extraction, binarization, thresholding and neural framework distinguishing proof. In initial segment, Binarization methodology is used to modified over -fold photo and later then balance it with area motivating pressure with recognizing lung most cancers development. In 2nd level, department is completed to parcel the lung CT picture and a robust part extraction approach has been familiar with disposing of a few primary additives of separated photographs. Isolated highlights are used to installation the brain framework subsequently the shape. The execution of the proposed system indicates nice outcomes and proposed strategy gives ninety six.Sixty seven % precision.

Bhatia et al. [15], Author shows a manner to cope with recognize cellular breakdown in the lungs from CT assessments making use of profound leftover studying. We painting a pipeline of preprocessing methods to feature lung locales helpless against malignant increase and pay attention highlights utilising UNet and ResNet fashions. The listing of talents is sorted into one-of-a-kind classifiers, viz. XGBoost and Irregular Woodland, and the singular expectations are ensembled to foresee the opportunity of a CT check being dangerous. The exactness executed is eighty four% on LIDC-IRDI beating past endeavors.

Makaju et al. [16], Author utilized was that mobile breakdown within the lung discovery techniques had been arranged and recorded primarily based on their vicinity precision. The strategies had been investigated on each step and in fashionable obstacle, dangers have been delivered up. Discovered some has low exactness and some has better precision however no longer towards one hundred percentage. Consequently, our exploration focuses to amplify the precision toward a hundred percentage.

Faisal et al. [17], In this exploration paper endeavours to evaluate the discriminative pressure of a few indicators in the evaluate to construct the productiveness of mobile breakdown within the lung's identification thru their aspect effects. Various classifiers which include Support Vector Machine (SVM), Neural Network, C4.5 Decision tree, Multi-facet Perceptron and Naïve Bayes (NB) are assessed on a benchmark dataset were given from UCI vault. The presentation is also contrasted and extremely good clothes, for instance, Arbitrary Backwoods and Greater part Casting a poll. In light of execution checks, it's far seen that Slope helped Tree beat any remaining man or woman as well as collecting classifiers and finished 90% precision.

Singh et al. [18], In this paper, Author applied seven wonderful classifiers called multinomial naive Bayes classifier, stochastic gradient descent classifier, random forest classifier, okay-nearest pals' classifier, assist vector device classifier, choice tree classifier, and multi-layer perceptron (MLP) classifier. Author have utilized dataset of 15750 medical photos comprising of each 6910 innocent and 8840 dangerous cell breakdowns within the lungs associated pics to prepare and test these classifiers. In the got outcomes, it's miles found that exactness of MLP classifier is higher with really worth of 88.55% in correlation with distinct classifiers.

Alam et al. [19], This paper proposed framework can likewise foresee the likelihood of mobile breakdown inside the lungs. In every segment of characterization picture development and department have been completed independently. Picture scaling, range space change and differentiation upgrade have been utilized for picture improvement. Threshold and marker-controlled watershed based totally division has been applied for division. For characterization reason, SVM paired classifier become applied. Author's proposed approach show extra substantial level of precision in cellular breakdown in the lung's reputation and forecast.

Abdillah et al. [20], In this paper, Author carries out and dissect the picture managing method for recognition of cell breakdown within the lungs. In this examination, writer proposed a region method for cellular breakdown in the lungs in mild of picture division. Picture division is one of halfway degree in photo dealing with. Marker manipulate watershed and locale developing technique are utilized to section of CT have a look at photo. Recognition ranges are trailed by way of image improve making use of Gabor channel, photograph department, and highlights extraction. From the final results by means of the test, writer tracked down the adequacy of our method. The results demonstrate that the nice method for primary factors computation is watershed with masking approach which has high exactness and accuracy.

TABLE I. Comparisons of various techniques and method used in existing system

Ref.	Year	Datasets	Preprocessing	Methods	Results
[7]	2023	LIDC-IDRI & LUNA-16 database	Use an improved Gabor filter for pre-processing in our proposed method, which reduces parameters using the Gauss-Kuzmin distribution	AlexNet SVM classification model	97.98% of accuracy, 98.84% of sensitivity, 97.47% of specificity, 97.53% of precision, and 97.70% of F1
[8]	2023	chest X-ray images	Using Histogram Equalization	OCNN-SVM approach	98.70% of accuracy, 98.34% of recall, 98.49% of precision, and 98.76% of F1
[9]	2021	CT Scan images	Techniques used in pre-processing are Median Filtering. Histogram Equalization and morphological operation	ANN Classifier	95.6% of accuracy, 93.1% of sensitivity, 100% of specificity
[10]	2020	Lung Cancer CT Scan images	GLCM features on image	SVM classification algorithm	Technique determines the detection of benign or malignant lung cancer is 83.33%
[11]	2020	Database of Chest X-Ray and Lung cancer (LIDC-IDRI)	Morphological segmentation and watershed segmentation are used for automated nodule segmentation	MAN is used to classify chest X-Rays images and EFT is used to classify the lung CT images.	DL exactness is 96% for X-Ray images while the exactness is 97.27% for CT Images
[12]	2020	Database of cancer imaging archive (CIA) dataset	Multilevel brightness-preserving approach	Improved deep neural network and ensemble classifier.	The proposed system recognized the cancer with maximum exactness.

[13]	2019	Image was collected from Cancer imaging Archive (CIA) dataset	The noise is removed using weighted mean histogram equalization approach. In addition, improved profuse clustering method (IPCT) is applied for segmenting the affected region.	Deep learning instantaneously trained neural network (DITNN) is utilized.	exactness 98.42% & minimum classification error of 0.038.
[14]	2019	Database obtained from UCI repository	Picture securing, pre-handling, binarization, thresholding, division, feature extraction are applied.	The fuzzy neural system is used to test the neural system with machine learning approaches.	Exactness 96.67 %.
[15]	2019	Database of Lung Images Dataset Consortium image collection (LIDC-IDRI)	This step involves of segmentation is followed by normalization and zero centering.	A number of classifiers like XGBoost and Random-Forest are utilized.	Exactness 84%
[16]	2018	Dataset of Lung Image Dataset Consortium (LIDC)	Median filter and Gaussian filter are applied on the CT images.	Watershed segmentation for detection and SVM for classification of nodule as Malignant or benign.	Exactness 92%.
[17]	2018	Dataset obtained from UCI repository	Pre-processing for data cleaning is applied	A number of classifiers including: MLP, Neural Network, Decision Tree, Naïve Bayes, Gradient Boosted Tree, and SVM are assessed	Exactness 90%.
[18]	2018	Database of Lung Image Database Consortium (LIDC)	Converting to grayscale image, applying denoising methods such as median blur, Gaussian blur, and bilateral blur, then applying thresholding methods for converting the grayscale image into a binary image.	K-nearest neighbors classifier, support vector machine classifier, decision tree classifier, multinomial naïve Bayes classifier, stochastic gradient descent classifier, random forest classifier, and multi-layer perceptron (MLP) classifier are applied	Exactness 88.55%.
[19]	2018	The lung cancer database utilized for training is taken from UCI machine learning dataset	Image enhancement and segmentation has been done. Image scaling, color space transformation and contrast enhancement	multi-class SVM classifier	Precision of 97% for cancer recognition and 87% for cancer prediction
[20]	2016	Used CT-scan Images are from VIA and ELCAP dataset	Image enhancement using Gabor filter is applied	Region Growing, Marker Controlled Watershed, and Marker Controlled Watershed with Masking are applied.	Watershed with masking method has highest exactness and robustness.

III. CONCLUSION

This paper surveys, lung cancers is analysed at a starting section, it might be useful in light of the reality that the prescription will then, at that point, be started out to keep infection from having a negative final outcome. Thusly, this paper sums up an itemized study on distinct AI approaches to deal with institution lung malignancies using either CT filter out pics or X-ray pictures. Additionally, several classifiers have been worried via the analysts in the writing, for example, Gradient Boosted Tree, Decision Tree, k-nearest neighbour, random forest classifier, MLP, SVM, Naïve Bayes, Neural Network, stochastic gradient descent, and ensemble classifier.

Subsequently, and in view of the large take a look at that have been done on these details, it very well can be presumed that the strategies which used Conventional neural network methods were given higher results as a long way as precision than other conventional AI techniques.

REFERENCES

- [1] Oudkerk, Matthijs, ShiYuan Liu, Marjolein A. Heuvelmans, Joan E. Walter, and John K. Field. "Lung cancer LDCT screening and mortality reduction—evidence, pitfalls and future perspectives." *Nature reviews Clinical oncology* 18, no. 3 (2021): 135-151.
- [2] Ferlay, Jacques, Murielle Colombet, Isabelle Soerjomataram, Donald M. Parkin, Marion Piñeros, Ariana Znaor, and Freddie Bray. "Cancer statistics for the year 2020: An overview." *International journal of cancer* 149, no. 4 (2021): 778-789.
- [3] Bade, Brett C., and Charles S. Dela Cruz. "Lung cancer 2020: epidemiology, etiology, and prevention." *Clinics in chest medicine* 41, no. 1 (2020): 1-24.
- [4] Tu, Yuqin, Yan Wu, Yunfeng Lu, Xiaoyun Bi, and Te Chen. "Development of risk prediction models for lung cancer based on tumor markers and radiological signs." *Journal of clinical laboratory analysis* 35, no. 3 (2021): e23682.
- [5] Tuncal, Kubra, Boran Sekeroglu, and Cagri Ozkan. "Lung cancer incidence prediction using machine learning algorithms." *Journal of Advances in Information Technology* Vol 11, no. 2 (2020).
- [6] Subramanian, R. Raja, R. Nikhil Mourya, V. Prudhvi Teja Reddy, B. Narendra Reddy, and Srikar Amara. "Lung Cancer Prediction Using Deep Learning Framework." *International Journal of Control and Automation* 13, no. 3 (2020): 154-160.
- [7] Naseer, Iftikhar & Akram, Sheeraz & Masood, Tehreem & Rashid, Muhammad & Jaffar, Arfan. (2023). Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2023.3285821.
- [8] Vs, Sreepada & Vedavathi, Dr. (2023). Lung Cancer Detection from X-Ray Images using Hybrid Deep Learning Technique. *Procedia Computer Science*. 230. 467-474. 10.1016/j.procs.2023.12.102.
- [9] Shailesh, M., Bhise, S., & Khot, P.S. (2021). Early Stage Lung Cancer Diagnosis using ANN Classifier. 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 22-27.
- [10] Q. Firdaus, R. Sigit, T. Harsono and A. Anwar, "Lung Cancer Detection Based On CT-Scan Images With Detection Features Using Gray Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM) Methods," 2020 International Electronics Symposium (IES), IEEE, 2020, pp. 643-648, doi: 10.1109/IES50839.2020.9231663.
- [11] Bhandary, A., Prabhu, G. A., Rajinikanth, V., Thanaraj, K. P., Satapathy, S. C., Robbins, D. E., ... & Raja, N. S. M. (2020). Deep-learning framework to detect lung abnormality—A study with chest X-Ray and lung CT scan images. *Pattern Recognition Letters*, 129, 271-278.
- [12] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2020). Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. *Neural Computing and Applications*, 1-14.
- [13] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2019). Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. *Measurement*, 145, 702-712.
- [14] Reddy, U., Reddy, B., & Reddy, B. (2019). Recognition of Lung Cancer Using Machine Learning Mechanisms with Fuzzy Neural Networks. *Traitement du Signal*, 36(1), 87-91.
- [15] Bhatia, S., Sinha, Y., & Goel, L. (2019). Lung cancer detection: A deep learning approach. In *Soft Computing for Problem Solving* (pp. 699-705). Springer, Singapore.
- [16] Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detection using CT scan images. *Procedia Computer Science*, 125, 107-114.

- [17] Faisal, M. I., Bashir, S., Khan, Z. S., & Khan, F. H. (2018, December). An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. In 2018 3rd International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST) (pp. 1-4). IEEE.
- [18] Singh, G. A. P., & Gupta, P. K. (2019). Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Neural Computing and Applications*, 31(10), 6863-6877,
- [19] Alam, J., Alam, S., & Hossan, A. (2018, February). Multi-stage lung cancer detection and prediction using multi-class svm classifie. In 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2) (pp. 1-4). IEEE
- [20] Abdillah, B., Bustamam, A., & Sarwinda, D. (2017, October). Image processing-based detection of lung cancer on CT scan images. In *Journal of Physics: Conference Series* (Vol. 893, No. 1, p. 012063). IOP Publishing



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details