International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 7 ǁ July 2022 ǁ PP. 61-67

Efficient Classification and Detection of Lung Nodules Using Deep Learning

[1] Renukesh G, [2] Shreyas KM, [3] Suhas G, [4] Nagateja Y,
[5] Dr J T Thirurkrihshna

[1][2][3][4] BE Students, Department of Information Science and Engineering
Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India
[5] Associate. Professor, Department of Information Science and Engineering
Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India

ABSTRACT
Lung cancers is one of the deadliest illnesses with inside the world. However, early diagnosis and treatment can save lives. Although CT scans are the best imaging modality in the medical field, it is difficult for doctors to interpret and identify cancer from CT scan images. Therefore, computer-aided diagnosis can be helpful for doctors in accurately identifying cancerous cells. It usually causes death for both men and women, so it is more important for care to immediately and correctly examine any nodules. Several methods have been developed to detect lung cancer at an early stage. Many computer- aided techniques have been used to detect lung cancer. The method used was to sort the techniques based on their accuracy in detecting lung cancer. From this study it has been found that CT scans provide more accurate results. Therefore, mostly CT scan images are used for detection of cancer. Also, marker-controlled watershed segmentation provides more accurate results than other segmentationtechniques.

Date of Submission: 22-06-2022 Date of acceptance: 04-07-2022

I. INTRODUCTION

Cancer is a group of diseases involving the abnormal growth of cells with the ability to invade or spread to other parts of the body. These tumors are contrasted with benign, non-spreading tumors. Possible signs and symptoms include a lump, unusual bleeding, persistent cough, unexplained weight loss, and stool changes. Smoking is responsible for about 22% of cancer deaths. About 5 to 10% of cancer cases are caused by genetic abnormalities. Cancer can be detected by certain signs and symptoms or by screening tests. It is then often studied further with medical imaging and confirmed by biopsy. About 5 to 10% of cancer cases are caused by genetic abnormalities. Cancer can be detected by certain signs and symptoms or by screening tests. It is then often studied further with medical imaging and confirmed by biopsy. Lung cancer, also known as lung carcinoma (because about 98-99% of all lung cancers are carcinomas), is a malignant lung tumor characterized by uncontrolled growth. control of cells in lung tissue. Most cancers that start in the lungs, known as primary lung cancers, are carcinomas. The two main types are small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC). The most common symptoms are cough (including coughing up blood), weight loss, shortness of breath, and chest pain.

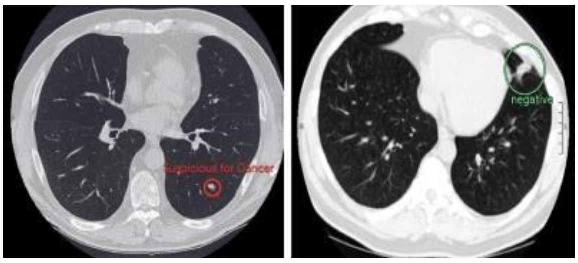


Fig 1: Detection of lung cancer

The vast majority (85%) of lung cancer cases are caused by long-term smoking. About 10-15% of cases occur in people who have never smoked. Lung cancer can be seen on chest X-rays and computed tomography (CT) scans. The diagnosis is confirmed by biopsy, usually done by bronchoscopy or CTguidance.
According to global cancer statistics, the five most common types of cancer in 2018 are lung cancer (2.09

million), breast cancer (2.09 million), colon cancer (1.8 million), and prostate cancer (128). There were 10,000 cases) and skin cancer (1.28 million cases). There was no melanoma). (1.04 million cases). The five major causes of cancer death are lung (1.76 million deaths), colon cancer (862,000 deaths), stomach (783,000 deaths), liver (782,000 deaths), and breast cancer (627,000 deaths). Died). It is the most common type of cancer in the world and is the leading cause of cancer-related death. However, a unique feature of lung cancer is that it can prevent more than 75% of cases by eliminating tobacco carcinogens. In addition, low-dose CT scan screening has been shown to be effective in reducing lung cancer mortality, as well as screening programs for diseases such as colon and breast cancer. Therefore, smoking cessation and screening can reduce the incidence of lung cancer and improve mortality. Of the most common cancers, lung cancer is probably the most dramatic advancement in screening, diagnosis, and treatment in the last decade. Intrabronchial ultrasonography and navigation bronchoscopy has revolutionized the clinical staging of lungcancer (6).

Non-small cell lung cancer

The three major subtypes of NSCLC are adeno carcinoma, squamous cell carcinoma, and large cell carcinoma. Rare subtypes include lung intestinal adenocarcinoma. Almost 40% of lung cancers are glandular cancers, usually originating from peripheral lung tissue. Most cases of adeno carcinoma are associated with smoking, but the most common lung cancer in people who smoke less than 100 cigarettes in their lifetime ("non-smokers") and former smokers with a moderate smoking history. It is also a typical form. Bronchoalveolar carcinoma, a subtype of adenocarcinoma, is more common in female non-smokers and may have a higher long-term survivalrate. Squamous cell carcinoma causes about 30% of lung cancers. They usually occur near large airways. At the center of the tumor is often caries and associated cell death. About 10-15% of lung cancers are large cell cancers. These are so named because the cancer cells are large, with excess cytoplasm, large nuclei, and prominentnucleoli (5).

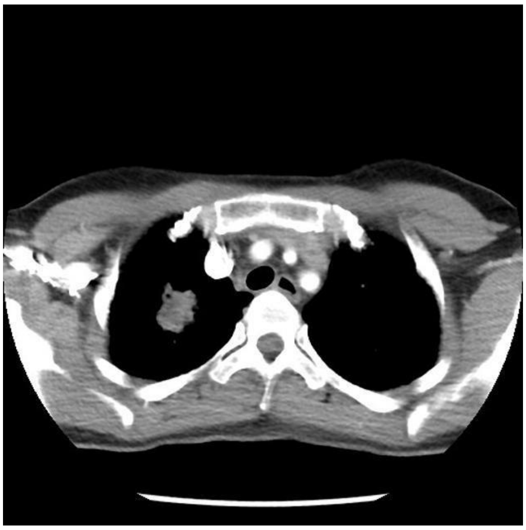


Fig 2: Non-Small Cell Lung Cancer

Small Cell Lung Cancer

In SCLC, cells contain dense neurosecretory granules (vesicles containing neuro endocrine hormones) that indicate an association of this tumor with endocrine syndrome or paraneo plastic syndrome. Most often it occurs in the larger airways (primary and secondary bronchi). Approximately 60-70% suffer from a widespread illness (which cannot be treated in a single radiation therapy field).

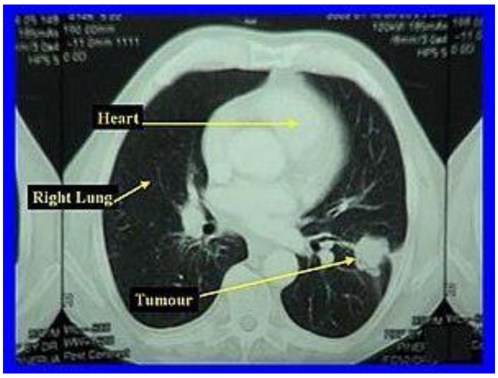


Fig 3: Small Cell Lung Cancer through CT Scan

CT Scan

CT scans, also known as computer tomography scans (formerly known as axial computer scans or CAT scans), are radiation to obtain detailed internal images of the body non-invasively for diagnostic purposes. It is a medical imaging technology used in medicine (X-ray). The person performing the CT scan is called a radiologist or radiologist. The CT scanner uses a rotating x- ray tube and a series of detectors located in the gantry to measure the attenuation of x-rays by different tissues in the body. Multiple X-ray measurements from different angles are then processed on a computer using a reconstruction algorithm to generate a tomographic (cross- section) image (virtual "slice") of the body.

Since its development in the 1970s, CT has proven to be a versatile imaging technique. CT is most prominently used in diagnostic medicine, but it can also be used to create non-living images. The 1979 Nobel Prize in Physiology or Medicine was jointly awarded to South African-American physicist Allan M. Cormack and British electrician Godfrey N. Hounsfield "For the Development of Computer-Assisted Tomography".

II. LITERATURE SURVEY

Moffy Vas et al, says that methodology used in this project Automated lung cancer detection system. Application of Median filter to remove impulse noise in images success. Also, morphological manipulation Contributed to satisfactory results in the process segmentation. Artificial neural networks have proven themselves a classifier with acceptable accuracy methodology. As a result of being adopted in this project, the accuracy was 92%. Hospital database. This system the accuracy and speed of the lung cancer detection system. This, it also helps detect cancer in the early stages (1).

Matko Saric et al, says that fully automated deep learning, Base method for full slide detection of lung cancer Histopathological image. The VGG16 and ResNet50CNN architectures are compared and the first architecture shows a higher AUC and patch classification accuracy. The results presented show this Convolutional neural network can move the lungs Cancer diagnosis from the entire slide image, but more effort it is necessary to improve the classification accuracy. In future work the next step is to increase the size of your training set Image enlargement and speckle normalization. Besides, we will try Train from scratch instead of using pre-trained weights With ImageNet (2).

Wadood Abdul et al, says that the key input during this paper is that the application and analysis, without computing the morphology and texture, the architecture of CNN, a deep learning solution, in classifying the lung nodules as benign or malignant. The LIDC-IDRI database was tested and therefore the best results were obtained with the 97.2% accuracy, 95.6% sensitivity, and 96.1% specificity, which outperforms the results obtained with other learning techniques. The comparison shows that the proposed ALCDC system performsbetter than the present state-of-the-art systems. The proposed ALCDC are helpful in diagnosis research and health care systems (3).

Tiantian Fang et al, says that GoogLeNet-one of the foremost advanced CNNs in visual imagery analysis- was fine-tuned and applied to the sector of determining carcinoma malignancy. This significantly increased the proposed system's convergence rate and its accuracy, sensitivity, and specificity. In other words, less training time produced better output results. Additionally, MIPs were applied to concatenate multi-view information of three- dimensional CT scans into RGB images that are compatible with the fine-tuned GoogLeNet. MIPs enabled the proposed system to integrally learn features of malignant and benign lung nodules during the training, and procure high accuracy when tested on the validation sets (4).

III. METHODOLOGY

Lung cancer detection and classification with DGMM-RBCNN technique

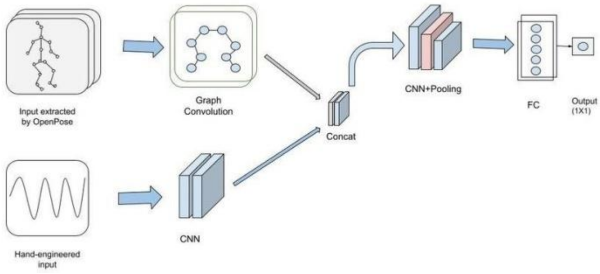
The next phase is included in the predicted ideal diagnosis of lung cancer using CT images of the lung as input. (1) Perform denoising preprocessing (2) Segment the data using Wiener and Gaussian filters. Segmentation to separate the lung region from the surroundings using area growth, (3) collect lung nodule features to reduce the number of dimensions, and (4) classify lung nodules based on successfully extracted properties. Develop the algorithm DGMM-RCNN model. (5) In conclusion, the bestfeatures were identified and lung nodules were identified (9).

LIDC-IDRI

The Lung Image Database Consortium (LIDC-IDRI) includes collections of thoracic computed tomography (CT) scans with an notations of lung cancer lesions. These scans can be used to predict and screen for lung cancer. The National Cancer Institute (NCI) is a national research institute that oversees the development of cancer-related treatments and research. NCI was founded through the active participation of many different organizations and individuals. This evaluates a consortium based on a consensus-based decision-making process (8).

Deep learning for lung Cancer detection and classification

The lung disease classification has a general structure which includes identifying the type of the lung disease, the cause of the lung disease, and the symptoms of the lung disease. Lung tissue pictures are taken using an offline technique. The input lung tissue is used in the online process. The adaptive bilateral filter is used to reduce the noise in the image, and the image contrast is increased. The histogram equalization process improves the quality of an image by removing any unwanted noise. After the image has been pre-processed, the removal of the lung region is the next step. The artificial beecolonywasusedtoharvestthelungregion.Asegmentationstrategyisused.Theholesinthelungs were filled with material called in the output of the ABC segmented image and mathematical morphological techniques were used. It then uses texture properties to locate malignant lung nodules. Lungs The next step is to classify thedata (10).



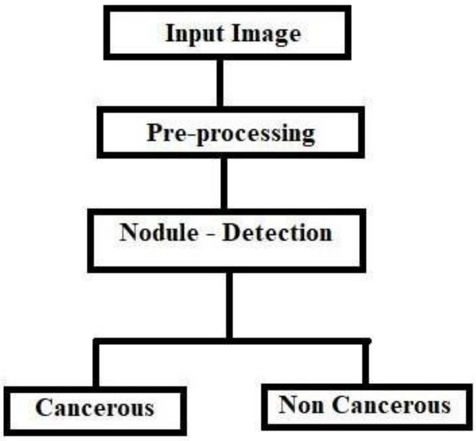
An automated lung cancer detection system based on machine learning algorithm

The lung cancer detection method has a number of stages, each of which must be completed in order for the method to be effective. At first, the work's flow is demonstrated. They are CT image pre- processing, segmentation, highlight extraction and classification. The CT image cannot be classified on the go, so it includes several important steps to achieve superior classification results. The pre- processing of CT images is intended to make them easier to handle for activities such as image differentiation. The main goal is to improve the clarity of the distinctions between locales. The Versatile Histogram Equalization (AHE) procedure is used to improve the image differentiation of CTscans.The purpose of picture division is to center the image on all parts so that the image can be more easily classified. This work demonstrates that Kernelized FCM (KFCM) is more efficient than standard FCM. The sectioned locales are passed to the next phase called highlight extraction, which extracts the curvelet also GLCM highlights. The features of these classifiers allow them to achieve much better results (11).

Novel Method For Lung Tumour Detection Using Wavelet Shrinkage-Based Double Classifier Analysis

This section provides information about the stages and process used to detect lung nodules on CT images. This LIDC dataset includes 48 scans from several subjects. There were several pictures in each of these scans. From the 48 scans that also provided nodule pattern information, 80 slices were chosen in total.These images were 512 512 pixels in size and were stored in commonly used formats, such as JPEG or JPEG. Each patient had 80 pieces of data extracted from this dataset. The first step is to obtain photos from the LIDC database and an 80-image in-house clinicaldataset (7).

PROPOSED ARCHITECTURE:



Pre-processing Stage:

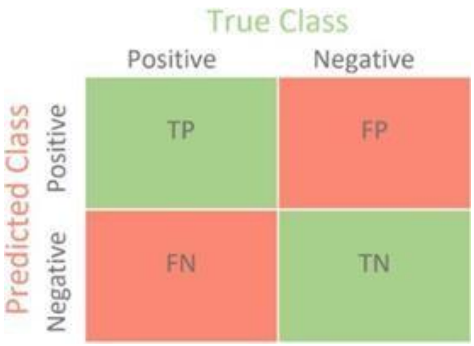
Each chest medical image is scanned and saved. The quality of the medical chest image is reduced by fluctuations, shifts, and disturbances during scanning. To produce accurate medical chest images, it is necessary to pre-process the data beforehand. This includes removing unnecessary noise from scanned images, without compromising the detail. This phase of the diagnostic and analysis procedure is crucial for understanding the problem. Filtering is an important step in the pre-processing process that helps to increase contrast and reduce noiseto improve imagequality.

Nodule Detection:

The lung nodule detection system includes segmenting the lungs and identifying potential nodules. Detection and reduction of false positives are important in the detection of malware. Some people have reviewed the method used for nodules. They say that it is effective. False positive detection and reduction is important to achieve overall sensitivity and reduce the number of false positives, but little comparative analysis is done to determine how effective the features used to reduce false positives are. We have outlined the different techniques for extracting features from a data set using a feature-based classifier. I tested the system's sensitivity and performance by reducing the number of False Positives and FPs detected. I also checked the accuracy of the system's reports by comparing them to those from traditional function-based methods. As far as we know, our reviews are comprehensive and up-to-date, 120 covering development in thisarea.

IV. RESULT:

Confusion matrix is a tabular way of visualizing the performance of your prediction model. Each entry in a confusion matrix denotes the number of predictions made by the model where it classified the classes correctly or incorrectly.



The metrics for the Confusion Matrix are:

True Positive (TP): It refers to the number of predictions where the classifier correctly predicts the positive class aspositive**.**

True Negative (TN): It refers to the number of predictions where the classifier correctly predicts the negative class as negative.

False Positive (FP): It refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.

False Negative (FN): It refers to the number of predictions where the classifier incorrectly predicts the positive class as negative. Common performance measures for the confusion matrix are:

Accuracy: It gives you the overall accuracy of the model, meaning the fraction of the total samplesthatwerecorrectlyclassifiedbytheclassifier.Tocalculateaccuracy, use the following formula:(TP+TN)/(TP+TN+FP+FN).

Precision: It tells you what fraction of predictions as a positive class were actually positive. To calculate precision, use the following formula: TP/(TP+FP).

Recall: It tells you what fraction of all positive samples were correctly predicted as positive by the classifier. It is also known as True Positive Rate (TPR), Sensitivity, Probability of Detection. To calculate Recall, use the following formula: TP/(TP+FN).

Specificity: It tells you what fraction of all negative samples are correctly predicted as negative by the classifier. It is also known as True Negative Rate (TNR). To calculate specificity, use the following formula: TN/(TN+FP).

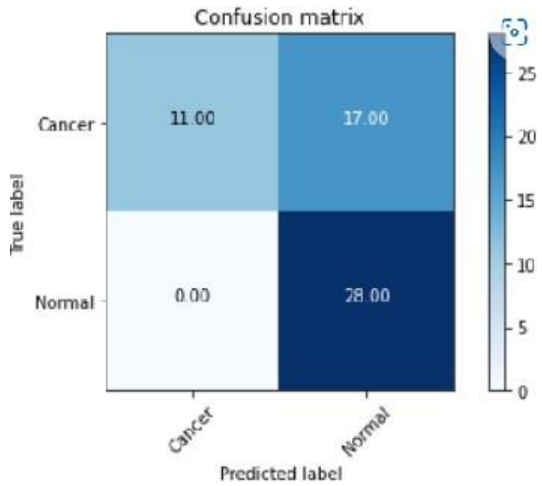


Fig 7: Confusion matrix designed system

By using the CNN model to detect the result of lung cancer in the current work. An accuracy of 70%, precision of 40% , recall of 1%,and specificity of 62% is obtained .

V. CONCLUSION:

GCN is a valuable tool for detecting lung cancer, as it is effective and efficient. GCN and CNN have been the most successful methods from the start, and have been widely implemented. By using the GCN and CNN model, we were able to detect lung cancer in the current work. A 70% accuracy, 40% precision, 10% recall, and 62% specificity is achieved with four different CNN structures. The CNN model's input consists of two types of candidate nodules: those with pairs of images, and those that do not. To get a cube, overlap the index from 0.

REFERENCES:

[1]. Vas, Moffy, and AmitaDessai. "Lung cancer detection system using lung CT image processing." *2017InternationalConferenceonComputing,Communication,ControlandAutomation(ICCUBEA)*. IEEE,2017.

[2]. Šarić, Matko, Mladen Russo, Maja Stella, and Marjan Sikora. "CNN-based method for lung cancer detectioninwholeslidehistopathologyimages."In*20194thInternationalConferenceonSmartand Sustainable Technologies (SpliTech)*, pp. 1-4. IEEE,2019.

[3]. Alhussein, M., Ali, Z., Imran, M. and Abdul, W., 2016. Automatic gender detection based on characteristics of vocal folds for mobile healthcare system. *Mobile Information Systems,2016*.

[4]. Tian,T.,Liu,H.,Li,L.,Yu,J.,Ge,S.,Song, X.andYan,M.,2017.basedbiosensorfornoninvasive detection of epidermal growth factor receptor mutations in non-small cell lung cancer patients. *Sensors and Actuators B: Chemical*, *251*,pp.440-445.

[5]. B. Kumar S and M. Vinoth Kumar, "Detection of Lung Nodules using Convolution Neural Network: AReview,"2020SecondInternationalConferenceonInventiveResearchinComputingApplications (ICIRCA), 2020, pp. 590-594, doi:10.1109/ICIRCA48905.2020.9183183.

[6]. B. Kumar S and M. Vinoth Kumar, "Detection of Lung Nodules using Convolution Neural Network: AReview,"2020SecondInternationalConferenceonInventiveResearchinComputingApplications (ICIRCA), Coimbatore, India, 2020, pp.590–594.

[7]. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–

424.https://radiologyassistant.nl/chest/solitary-pulmonary- nodule/benign-versus-malignant.

[8]. Pehrson, L.M., Nielsen, M.B. and AmmitzbølLauridsen, C., 2019. Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: a systematicreview. *Diagnostics*, *9*(1), p.29.

[9]. Jena, Sanjukta Rani, S. Thomas George, and D. NarainPonraj. "Lung cancer detection and classificationwithDGMM-RBCNNtechnique."*NeuralComputingandApplications*33,no.22(2021): 15601-15617.

[10]. Asuntha,A.,andAndySrinivasan."DeeplearningforlungCancerdetectionandclassification." *Multimedia Tools and Applications* 79.11 (2020): 7731-7762.

[11]. Lalitha,S."Anautomatedlungcancerdetectionsystembasedonmachinelearningalgorithm."*Journal of Intelligent & Fuzzy Systems* 40.4 (2021): 6355-6364.