International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Learning Methodology of Lung Cancer Detection and Diagnosis Using CT Images: A Systematic Approach

Sadhana B.[1], Dr. Pramod Kumar Naik [2]

Submitted: 28/01/2024 **Revised**: 06/03/2024 **Accepted**: 14/03/2024

Abstract: In the present era lung nodule is the very dangerous and deadly cancer disease requires initial diagnosis which improves patient survival probability. Many techniques have played the major and important role in the medical field in detection of lung cancer by analyzing lung medical image. Here we are carried out a systematic article surveys of different article published from last five years. The four different databases like (Science Direct, Scopus, web of science, and IEEE), used during last five years and chosen around 15 articles to carry out systematic survey on lung cancer detection. The major goal of the survey work is to consolidate and concise recent advancement in lung cancer detection medical field, diagnosis considering various detection algorithms and methods. This article concise and summarizes the deep knowledge by addressing the results and findings of recent research which enhances and provides sufficient knowledge in the relevant field. We included challenges, applications, and recommendation for further enhancements after analyzing different research articles in detail. The lung image screening, diagnosis of cancer using CT imaging, in which detailed scanned images of lung is included. Further enhancements by using of CAS ("computer-assisted systems") DL concepts were promoted to interpret lung nodule detection of CT images. The goal of this article is to cover the over-view related to DL techniques, detection of lung cancer using DL techniques and its related applications and benefits. This paper mainly focuses on two different methods of DL while screening and diagnosis of lung cancer, like classification and segmentation methods. The shortcomings and benefits, advancement of DL models are also be elaborated, the analysis picturises and demonstrates there is a potential significance of use of DL techniques to enable and provide accurate, precise, effective- computer assist for lung cancer diagnosis and screening by considering CT images. The part of this review article a set of potential recommendations for future scope and work which improves available applications of DL methods to enhances the better and efficient accurate results in identification of lung cancer.

Keywords: *Lung Cancer Detection,Image Pre-Processing, Image Segmentation ,Convo-lution Neural Network.*

1. Introduction

As per American cancer society estimations for lung cancer during 2023 in US there are 238340 new lung cancer patients out of which 117550 are men and 120790 are women. And the mortality counts due to lung cancer in men 67,160 and in women 59,910. Despite of covid-19 pandemic, in considerations of other reasons of death rate, the cancerogenic death- rate has continued in decline from 2019-2020 by 1.5%, 33% overall reduction from 1991.Because of advancements in treatment techniques there is a rapid decline in the mortality (2% annually during 2016-2020) for kidney cancer, melanoma, leukaemia etc., and moderate declines in the case of lung cancer. In contrast there is rapid decline in -the mortality ratio because of recent advancements in the clinical treatments, [1].

In defiance of advances in understanding the risk factor developments in immunization control, treatments for lung cancer options, it remained the major cause for human death. Smoking habits and tobacco consumptions remains the major factor for cancel developments also nontobacco risks include infections in lung, lifestyle food habits, chronic lung infections etc. Avoidance of tobacco consumptions may reduce the risk of lung cancer along with healthy diet plans, physical activities etc. [2].

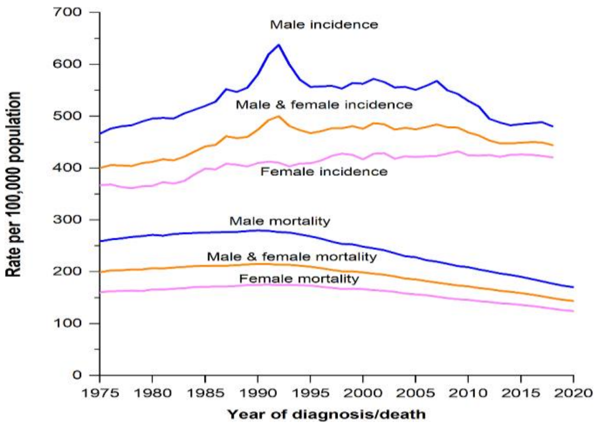


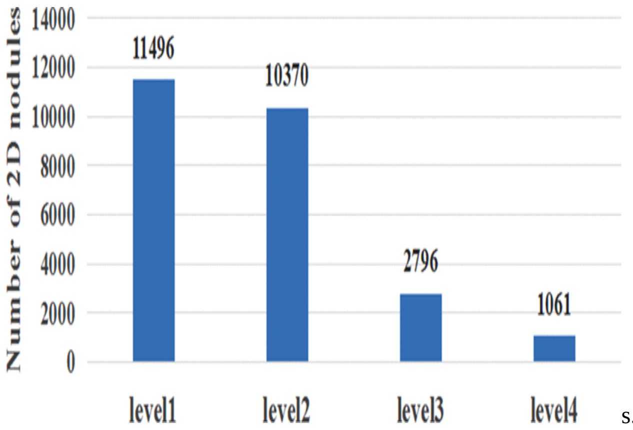
Fig .1. cancer incidence trends (1975–2019) mortality (1975–2020wrt sex, US

The author has proposed an efficient and novel computed aided automatic nodule detection which can minimize false positive results. To avoid juxta –pleural –nodules and automatic wall

1 Research Scholar DSU Bangalore, Assistant professor Department of ISE, CEC Mangalore, INDIA,
ORCID ID : 0000-0003-1426-6507
2 Associate Professor, Department of CSE, DSU BANGALORE INDIA
ORCID ID : 0000-0001-5767-256

mending technique was introduced. The pre-processing method adapted in this study was vascular elimination, which can enhance and highlights nodule portions by weakening the vessels. Four types of CNN structures based on the 4 nodules levels were applied. Two distinct group's forms pairs of images like group 1 and group 2 were the inputs to the CNN model. In comparisons with the Convention CNN model, the adapted method can simplify the data, and provides better results by considering merits of two group correlations.

By considering diagnosis information and data which will be provided to the radiologist and clinicians, helps them to predict structure of lung nodule



s.

An automated leaning based – nodule detections help the radiologist to early detect and identify relevance information's which are un covered from the big data.[3].

2.1 Deep Learning Methods Overview

Currently collecting lung images and interpretations of images for diagnosing is a challenging task. The high-resolution image capture techniques are available like MRI, XRAYS, CT scans etc. After lung image pre-processing step, the model extracts relevant features and information's from the images and build a training model there by unknown image can be tested for cancer detection. The traditional cancer detection technologies were un-able to detect the lung cancer and to produce reliable results. Deep-learning methods were successfully introduced and developed which will effectively diagnose the disorders.

The techniques under deep-learning are the subset of ML, that will enable estimation of outcomes based on provided dataset using the training model. The DL techniques uses NN with many layers like input, hidden, output layer. Based on the learning techniques and methods the outcome of the model can be validated. The deep learning models may be either i) reinforced learning ii) unsupervised learning models iii) semi-supervised learning models iv) supervised- learning models.

Supervised Deep Learning Models:

During the initial training phase for SDL type of learning model, labelled and specified data is needed. The DL model practices the training phase by considering every input combination which covers the target class labels. The list of training models was developed in order to forecast the labels of unknown samples, most frequently used DL models may be CNN, LSTM, RNN, and GRU.

Unsupervised Deep Learning Models:

The enhanced DL models doesn't need any tagged practice of the data. The model will analyse data's inheritance features considering few pertinent features and characteristics. This type of models were best suited for feature reduction and clustering methods. The commonly used learning models here are Auto-Encoders (AE) and Restricted Boltzmann Machines (RBM)

Semi-Supervised Deep Learning Models:

These types of learning models can use both labelled and unlabelled data, RNN, LSTM, GRU, GAN are the popular learning models.

Reinforced Deep Learning Models:

This learning model focuses more on chooses best courses of actions to reduce the rewards, and learn and perform behavioural study by considering the interactions with outside world. The DL is a branch of which uses multiple layered complex network structure to extract and learn the features from enormous amount of dataset. These network models have the capabilities to detect the linkage and patterns within the complex data for conventional ML, since they can learn automatically the complex features from different category of medical images like 2D CT images, 3D CT images, MR images etc.

The CT scanned images need more detailed lung images as they are most oftenly used in lung cancer identification. The detailed verification of shape, size, texture, brightness and intensity of the images, Deep learning methods can be used to identify and recognize the defected lung nodules. Deep learning methods uses complete images of the lung capacity, 3D CT image offers more accuracy in results compared with the 2D CT images.3D CT images exhibits high accuracy to detect the abnormalities. To reduce the exposure of radiations while screening lung cancer detection, low –dose CT methodologies are more often, which may be coupled with filtering, augmentation of image techniques. But MR images has more capabilities to reveal the tissue information's, blood flow.

Based on the observed features DL methods can analyse lung images to locate lung nodule including additional abnormalities. To increase the image quality, reduction in noise filtering methods like Gaussian, Median filters are implemented in the pre-processing stages. Region proposal and sliding window techniques were adapted in the case of deep learning approaches.

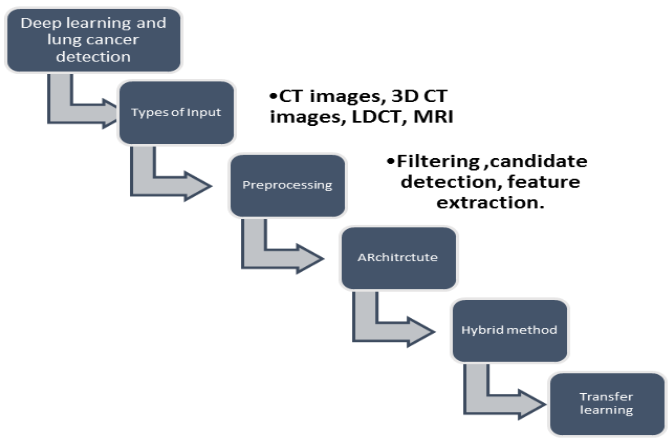


Fig.2. Cancer detection process

Unfortunately, disease detection with patient of stage 4 is associated with less survival rate and high symptoms. Early detection of cancer disease may reduce the mortal ration. The initial standard care for stage I and stage II including patients with stage IIIA, NSCLC is a surgical extirpation. The surgical removal of cancer cells follows some systematic procedure and therapies. The advantage of adjuvant cancer therapy varied from stage to stage includes stage 1B (Tumor size is greater than equal to 4cm), with a 3% reduction of risk of death during first 5 years. The usage of chemotherapy for stage III patients increased to around 13% comparatively.

This author stated that the CAD model designed is helpful to detect cancer in the early stage using CT lung images. This author has adapted four steps to detect cancer cells in the CT images taking in to account of medical professional's challenges to diagnose the lung cancer nodules from CT images. The first step filters the noisy contents by applying filtering concepts either by using Weiner Filter, Median, Gabor filter, Min and Max filter etc. In the preceding step thresholding like multiple and optical thresholding, active contour method, shape-based method etc were used. Segmentation was introduced to segment suspected nodules. The third step is the feature extraction like geometric features, shape –size features, gray scale features and statistical features need to be extracted from the CT lung images to detect malignant or non-malignant. To extract the features various classifiers like vector machine, generic algorithm, ANN. Rule based classifier, Linear Discriminate Analysis methods were adapted. As concluded by the research author still there is more of improvements required with respect to sensitivity, accuracy and specificity in the existing deionisation system. [3]

T Aggrawal et al. (2015) [4], in the proposed research paper the classification of cancer nodules, the author has introduced a model which uses gray scale characteristics and valuable thresholding values to perform segmentation process of lung nodules by threshold values, Gray scale characteristics to perform segmentation process. The proposed system achieved accuracy, sensitivity, Specificity of 84%, 97.14% and 53.33%.

The research author has proposed a model in which CT images were pre-processed to remove the noise. In the next step fuzzy k-mean algorithm was used to perform segmentation process, later on the result was improved by k mean approach. In the third step feature extraction was carried out from the CT lung images like correlation, entropy SSIM, PSNR, homogeneity using statistics GLCM feature extraction method. Finally, categorization is carried out using supervised NN like BPNN for the detection of lung cancer. This author has achieved 90.7% of accuracy but this accuracy can be enhanced using improvised classification technique like support vector machine. [5]

This author has proposed CNN model as the classification approach in the CAD system and achieved 84.6%, 82.5% of accuracy and sensitivity and specificity of 86.7% [6].

The author has proposed data mining methods, lung cancer patient database includes images of human upper half body X-rays that classifies like normal or malignant, benign. CAD system uses pattern recognition, feature extraction also classification process. The proposed model used X-ray images and accuracy gained is less comparing with CT images. This study elaborates that to achieve better accuracy implementation work can be enhanced and extended to apply the CT images in superiors' diagnosis of detection of lung cancer in the human body.[7]

The author has developed an enhanced model for segmenting the lung CT images by combining kernel graph cut algorithm and proposed a mathematical model. In the next stage the proposed algorithm was compared with K –mean algorithm, cluster variance algorithm and concluded with the comparative studies between two approaches [8].

In the proposed research article author has developed a CAD based system which detects cancer from lung CT images. This research has adapted four steps like pre-processing, segmentation, classification, and the validation. In the first step filtering process was carried out to enhance the input CT lung image followed by segmentation of blood vessels and pulmonary nodules by applying double level thresholding, morphological operations. In the next step features can be extracted from the segmented lung images which are HOG feature VH features, statistical features of first, second order GLCM etc. Fourth step is to get the better accuracy using MFFNN, radial basis NN, SVM.The validation process was carried out and the results obtained was CAR is -99.06%, S is -100% and SP is -99.2%. As per this article still there is a scope in improvements in early detection and classification of benign and non-benign tumor. [9]

The proposed article focused on how lung nodules can be diagnosed using skeletonization and Gini coefficient techniques which produces lung images as class malignant or class benign from the Data set. The Gini coefficient technique helps to identify the nodule distribution and the skeletonization helps to analyse the shape of the lung nodules. By considering these two parameters research author has proposed a prototype which performs discriminant analysis malignant or else benign. The results validation was depending on classification methodology and ROC curve. [10].

In the present medical era lung cancer has become the major health challenge in human body. Early-stage detection and diagnosing of lung cancer is quite difficult with less rate of survival, risk factor is high if it is not early detected and diagnosed. Different levels of cancer treatments are available in the medical and hospital science like chemotherapy, radiography, surgery depending upon the stages of the lung cancer. The survival rate is 14% for 5 years since cancer develops with in the respiratory system epithelium and bronchial trees of the lung and which spread across the human body. [11]

Early-stage detection of lung cancer required a powerful technology to assist experts provide a desirable treatment. The techniques like ML techniques, artificial intelligence techniques and image processing can process the lung cancer images, if necessary, pre-processing and training of medical CT and X Ray images by adapting NN, ML techniques.[12]

The researcher has developed a prototype to improvise lung cancer detection using Computed-Tomography lung images, are mainly best suited for detection of pulmonary the nodules in the Lung cancer. [13, 14].

In the present era lung cancer is the major cause which reflects modern lifestyle and contamination and colossally expansion of lungs issue. The lung illness treatments corresponding to the image processing techniques using CAD frameworks implementations. In this prototype initially pre-processing was carried out by highlighting the features from CT lung images. This researcher has adapted advanced image processing, machine learning technologies in the implementation step, later on comparison analysis of diverse classification mechanisms was carried out which will help to improve % accuracy while detecting lung cancer and suggest the use of advanced classification and segmentation technologies.

To achieve the set objectives the researcher has used an effective segmentation techniques like thresholding, ANN, SVM classification techniques. But the obtained results can be further enhanced by improving the accuracy while detecting lung cancer to avoid false positive results because cancer in lung may be treat-able only it is detected in the initial stages. As per the cancer study lung cancer doesn't exhibits early stage and initial symptoms and noise CT images doesn't give exact results in CAD system. The proposed system eliminates the risk and challenges by using filtering techniques like Auto encoder system, segmentation and classification techniques like OTSU algorithm, decision tree, CNN. The researcher has clearly discussed the LDD and how progressing can fulfil the medical filed using R-CAD (Robust) system to overcome the challenges of future work. [15].

Medical cancer detection tools are more essential in initial stage diagnosation and to monitor lung cancer while treatment. For lung cancer detection different medical image modalities like MR images, X-Ray images, computed and position emission tomography etc. have been extensively used. The medical techniques adapted have some major limitations like automatic classification of lung cancer nodules is the major challenge. Deep learning technique is the quick growing concept in medical imagining area by considering emerging applications like textual and image-based modalities. The author has proposed a deep learning image-based system for advance detection of the lung nodules. The recent achievements in the deep learning approaches for lung cancer detection segmentation and classification has proven a wide variety of techniques to achieve enhanced accuracy. Most of the system uses CT image datasets to train the networks and clinical and experimental results demonstrates deep learning technique may be superiorly used by the radiologists which enhances the effectivity of the segmentation, classification and detection. Deep learning approach is a potential technique which solves many of the medical challenges in spite of limitations like clinical verifications, privacy protection, legal accountability etc. [16].

Lung cancer detection is the challenging task for clinical and medical professionals to identify the true cause for cancer with in the human body and treatment is the risky and challenging task because exact treatment have not been invented so far. If cancer is detected in the initial stage that can be partially treated. In the medical image processing system involves reduction of noise, feature extraction, detection of damaged segments and comparison with the of medical history which used to locate exact region where lung cancer is located. This research has shown an accurate detection and prediction of lung cancer with the help of image processing techniques and Machine learning techniques. The image filtration steps initiated by the researcher using K-mean method and in the next step images were segmented. For classification ML Techniques like KNN, RF and ANN were used. Among these ML techniques ANN model produces better accuracy for prediction of Lung Nodules.

Approximately one –million lung cancer patients claiming the lives because of lung cancer disease since it is a deadliest type of cancer disease. According to current affairs in medical challenges it is difficult that lung nodule detection using CT scanned images. Even CAD system is also a crucial for the detection, hence image processing techniques is basic activity that is to be carried out in a wide range of medical scenario. This method proposed and demonstrated accurate prediction and classification technique which enables enhanced results with the help of image processing and machine learning techniques. Lung images were collected and pre-processed using geometric mean filter to increase the quality of the images. K mean approach was used to segment the enhanced images, in the next step ML based categorization algorithms were used for classification purpose [17].

Lung cancer staging and its investigation in the initial stage is the basic and major predictor of survival since it determines the treatment options. Using CT images there are much computer-aided diagnostics were developed and this majorly concentrates on CNN-based technology to detect various stages of lung cancer. The proposed work focused on the development of a model by which an improved rate of accuracy classifications of lung nodules and their analysis can be achieved.

Datasets

The usage of available datasets plays an important role in automatic lung nodule detection as well as classification. To attain accurate and correct performance outcomes using any computational techniques the availability of related datasets is an important parameter. Presently public availability of dataset usage has high impact fir detection, classification and identification of lung cancer. The Cancer detections includes identification and distinguish between presences of nodules in the lung image, classification includes separation of cancer nodules. CAD system allows timely identification of lung cancer and early diagnosis for the same by using CT images using advanced techniques like AI approach. This detection system using CAD based decision engine, which examines the CT images as an input image, considering different methodologies for image segmentation and classification of medical images. The various databases available may be like ANODE09 database, ELCAP database, LUNA16 dataset, LIDC-IDRI etc. Among available databases, LIDC-IDRI is most commonly standard dataset for evaluation of lung nodules [18].

To detect the shape of the nodules the polygon approximation methods can be deployed. The author has adapted 3 steps which include region separation using multi-thresholding method, filtering and smoothing the regions and boundaries morphological method are considered. In the next step polygon approximation methods are used for nodule extraction. Further in the following step hybrid feature vector is created in considerations with the histogram of intensity of the image, oriented -gradients, geometrical features of enhanced lung nodule candidates. The selected features were classified using SVM classifications, the image intensity, geometric image features vector is fused in to classifier like SVM for lung nodule identification. The proposed system has achieved accuracy of 98.8%, sensitivity of 97.7%, specificity 96.2% by using LIDC dataset. In the modern lung cancer technology 3D based segmentation methods can be adapted to detect lung nodules accurately [19].

Lung image dataset REVIEW:

The lung image dataset collection method includes 7 academic centers and 8 medical imaging collection companies combinable identified addresses to solve challenges like technical, clinical to create a strong database. The LIDC/IDRI Database contains 1018 CT images. In the preliminary stage each radiologist separately reviewed CT scanned images and identified nodules belonging to any one of these categories' nodule size ≥3mm, nodule size < 3mm, non-nodule size≥3mm [20],[21]. The LIDC/IDRI Dataset can be used to equip an important in research of medical imaging resource wrt CAD development, validations of the results, and diffusion of clinical applications.

Table 1: Lung image dataset review

Reference	Year of publication	Datasets of CT images	Results	%
27	2020	3D-UNet	LUNA 16	DSC-95.30
28	2020	DB-ResNet	LIDC-IDRI	DSC-82.74
29	2021	END TO END DEEP LEARNING	1916 LUNG TUMOURS IN 1504 PATIENTS	SENSITIVITY 93.29
30	2021	3D ATTENTION UNET	COVID 19 DATASET	ACCURACY 94.43
31	2021	IMPROVED U NET	LIDC-IDRI	PRECISION 84.91
32	2021	U-NET	LUNA 16,LIDC-IDRI	DSC 89.79,90.35
33	2021	DENSE R2U CNN	LUNA	SENSIIVITY 99.4 + OR – 0.2%
34	2021	CIRDL	First Affiliated Hospital of Guangzhou University	SENSIIVITY 87.63
35	2022	DENSENET201	Seoul St Mary's hospital dataset	SENSIIVITY 96.2

[22] An automated identification of nodules in lung CT images is the active and potential area of research from last two decades. Even though there is limited researches have been carried out which includes comparative analysis of different systems wrt to common database. The automatic nodule identification technique involves LUNA16 which analyses public dataset of CT images like LIDC-IDRI dataset.

The LUNA 16, researches can develop and upload predictions on CT scanned images in two different tracks, full or complete nodule detection technique in which full CAD system can be developed. And in the second approach Erroneously reduction in which a complete set of lung nodule classification be carried out. The proposed system adapted the leading solution in which CNN model was deployed, the combined solution exhibits excellent sensitivity of 95% at lesser than 1.0 false -positive/scan.

[23] ELCAP is the publicly available database which include set of CT images which can be used in different lung diagnosis systems. This database contains 50 LDCT lung images &379 unduplicated CT images of lung nodule

[24] In this proposed study authors have developed new dataset and conducted testing of computer aided pulmonary nodule detection and related few strategies in which intended to complement the existing datasets by adding additional focuses on radiological variabilities along with clinical and local reality. The nodule detection, segmentation, characterization techniques and its outcomes were tested and done comparative study with respect strategies of radiologist and computer aided systems. This method adapted Lung Nodule Database (LNDb)with includes 294 CT images (Centro Hospital) - Universitario de São Joãao.

[25] In this proposed system author has adapted maximum intensity project method while pre-processing and segmentation method included was segmentation system like SquExUNet and classification model like 3D-NodNet by considering publicly accessible database of lung images. Like (LIDC), LNDb Challenge- Dataset, Indian Lung CT Image Database (ILCID). The researchers were introduced cascaded 2D and 3D CNN model, the results were compared with the previous existing techniques of cancer detection and segmentation methods and they attained a Dice –coefficient metrics for segmentation of lung nodules and sensitivity of nodule detection strategies.

The authors developed an image database which is digital, contains 247 chest radiographs which includes with and without nodules. Around 154 chest radiographs with nodules and 93 without nodules were designated from 14 medical centres, later these images were digitised in to matrix size of 2048 × 2048 (0.175-mm pixels) and a gray scale size of 12-bit. The lung nodules were categorized in to 5 major groups like level 1, very subtle (detection may be very crucial since this level has low contrast, small size, or overlap with a normal structure); level 2, very -subtle (detection itself is crucial (very difficult); level 3, subtle (detection: difficult); level 4, relatively obvious (detection : relatively easy); and level 5, obvious (detection : easy). The final outcomes were analysed by considering 20 radiologist observations which is later implemented to ROC analysis (receiver operating characteristic) to detect pulmonary nodules. [26]

Table 2 Lung nodule segmentation approaches.

Reference	Dataset	Sample Number
20	LIDC	1018 CT lung images
21	LIDC-IDRI	1018 CT images of 1010 patients
22	LUNA 16	888 Lung CT Images :dataset is LIDC-IDRI
23	ELCAP	LDCT images 50 and 379 unduplicated CT images
24	LNDb	294 CT images (Centro hospitaller Universiatario de Sas Joaao
25	ILCID	CT Images (400 patients)
26	JSRT	154,93 nodules and non nodules with lables

The author has proposed a CT image segmentation methodology which is based on 3D-UNet, Res2Net and developed CNN model referred as 3D-Res2UNet.This new model has identical connection including all the relevant attributes extraction capacities. This enables the established network to enhance multiple scaled features with in a excellent granularity as long as elaborating the individual filed layers of the model. The proposed technique figures out the depth problems but the modelled network is not suitable for gradient dis-appearance and exploration problem which may enhance the segmentation - accuracy. The developed network ensures about feature map size, repair effectively the lost image features. The data set used was LUNA16, dice-coefficient indexed value = 95.30% , recall rate=99.1% which indicates this method has achieved high performance while segmenting Lung image nodules [27].

Here the authors have proposed a data driven model called as DB-ResNet and integrated two techniques to improve the model performance in which first method can continuously capture multiple viewed and scaled features of CT lung-nodules. During second method researchers have combined intensity features and CNN and adapted pooling methods like intensity of the pooling layer to take intensity parameters of voxel (center) of adapted system model block, in the later stage CNN model was used to get convoluted lung features of voxel of the block. In additions to this weighted sampling strategy was considered which is depends on the weighted scores of boundaries of the lung nodules which will improve the results accuracy. The model was evaluated using LIDC dataset which contains 986 lung nodules. The carried experimental results exhibits DB-ResNet is superior segmentation performances including Dice score=82.74% for LIDC database [28].

The outcomes of proposed study were validated with the available findings of four radiologists using the same dataset. The average dice score of software was higher than the average dice score of human experts by 0.49 percent. Hence, proposed research work is just as better as an experienced medical radiologist's.[29] The main purpose of accurate estimation of lung tumor in the case of biomedical image is an important aspect for real identification and monitoring therapeutic response. An automated detection is challenging and key role in spite of advancements in the Deep learning area. The authors have assumed that oncology department contours may provide a large database of 3D segmentation to achieve best models. And they developed and validated a model using DL to detect and segment lung tumours using CT images. By considering 1,916(lung tumors) of 1,504- patients with radiation treatments in which segmentation of image quality was verified by oncologist with the help of custom web-based application. While comparing the performance evaluation an external test set of CT images as well as segmentation of 59 patients who had single primary /metastatic type of tumor. The sensitivity of detection of tumor with the definition in which correct prediction at least one voxel with in ground truth tumor was 93.2%, Dice coefficient=0.67 with IQR range from 0.53 to 0.85. The researchers have achieved strong co-relation between physicians predicted tumor size and model predicted with r=0.69, P<0.0001.

This article helps to develop a systematic model for cancer diagnosis, treatment using a new segmentation approach which meets the necessity of lung CT- images processed during COVID -19 pandemic. Initially the selected extracted area in which implement patch mechanism to fulfil the 3D network applicability, eliminated the unwanted background. In the second step 3-D model was developed to attain a targeted area. The improvements in the network convergence and loss function is added to boost gradient while optimizing and training direction. At last augmentation method, conditional random field was used to resample and segment a given data. This method was assessed using some comparative study experiments and proposed system achieves better performance so it has high clinical applications [30].

This paper implements DNN established a lung cancer classification using optimizations like hyper parameter. This optimization of DNN is costly, surrogate –assisted evolution algorithm introduced to search automatically an optimal hyper parameter configuration on DNN.This proposed model is distinct from the available surrogate models which adapts covariance function which is stationary ie kernel to compute difference among the hyper parameter points. The non- stationary surrogate models adapt a function whose smoothness changes with respect to spatial- location of the given input. The ML-CNN model was built for lung nodule classification, optimization of available hyper parameter configuration is carried out using non-stationary kernel. This system outperforms manual tuning, optimization(hyperparameter), search (random grid), Gaussian processes, The Parzen Estimator Approach which is Tree-structured, Hyper parameter Optimization using RBF, HORD (Dynamic coordinate search) [31].

Performance Evaluation Metrics:

The performance evaluation measures of segmentation, classification of DL models like ACC, SN, AUC, TP, SP are basic factors need to be reviewed in the research area. Sensitivity means recalling the probability of recognizing true pixels of segmented images. SP entitles an ability of identification of negative pixels. Acc can be defined as the ratio between truly identified pixels. The total pixels with in the image may be (TP + FP + TN + FN). In the proposed system author has presented an automated lung nodule detection technique by initializing automated region(seeded) growing (SRG) to perform the segmentation technique of lung images. Without considering homogeneity condition SRG can segment and CT image automatically, it will not focus on similarities between the neibouring pixels. This system approaches a code technique like 3D chain for identification of juxta-pleural nodules. It extracts segmented lung image voxelized regions and boundary to identify regions and boundaries and detect infected areas. Use of 3D chain coding is more effective in order to exhibit the better accuracy in results, it does not need training of data set, classification methodologies to initialize beginning and endpoint gaps between a nodule. [22].

3. Methodology Of Proposed System

The automatic detection of lung nodule using ML techniques was adapted in which the technique uses two different phases like training and testing phase. During train-ing phase the lung images are augmented means preprocessed and next fed in to deep learning model. The output of deep learning model produces features acts as an input to the ML algorithm. Under testing phase after image screening classification of lung nodules using objective or subjective analysis methodologies were adapted.

Training Phase:1.Data Preparation:In the training phase, a diverse dataset of multi-Model medical images, such as X-rays and CT scans, is prepared. This dataset contains labeled examples indicating whether each image contains cancerous regions (positive class) or not (negative class).

2.Data Augmentation:

Data augmentation techniques are applied to create variations of the training im-ages. This includes transformations like rotation, scaling, and flipping to increase the diversity of the training dataset.

3.Model Initialization:

The multi-Model CNN model is initialized with random weights and architecture designed to handle different image Modalities. It usually includes convolutional layers for feature extraction, followed by fully connected layers for classification.

4.Loss Function and Optimization:

During training, the model computes a binary cross-entropy loss, which measures the dissimilarity between the predicted output and the true labels. The optimiza-tion algorithm, typically Adam or SGD (Stochastic Gradient Descent), adjusts the model's parameters to minimize this loss.

5.Forward and Backward Pass: The training dataset is processed in batches. For each batch, a forward pass com-putes the model's predictions, while a backward pass updates the model's weights using gradients obtained from the loss function.

6.Epochs:

The training process occurs over multiple epochs, where each epoch represents on complete pass through the entire training dataset. This repetition allows the model to learn iteratively.

7.Validation:

A validation dataset, separate from the training data, is used to monitor the mod-el's performance during training. This is important for early detection of overfit-ting.

Testing Phase:

1.Data Preparation:

In the testing phase, a separate dataset (the testing dataset) is used, which the model has not seen during training. This dataset is also composed of multi-Model medical images with known ground truth labels.

2.Forward Pass:

The testing images are passed through the trained model for inference. The model predicts whether each image contains cancerous regions based on the learned fea-tures.

3.Performance Metrics:

Various performance metrics are calculated based on the model's predictions and the ground truth labels in the testing dataset. These metrics include:

- Accuracy: The percentage of correct predictions.
- Precision: The ratio of true positives to the sum of true positives and false pos-itives.
- Recall (Sensitivity): The ratio of true positives to the sum of true positives and false negatives.
- F1-Score: The harmonic mean of precision and recall.
- ROC-AUC: The area under the receiver operating characteristic curve, which measures the trade-off between true positive rate and false positive rate.

4.Interpretability:

Methods for interpretability, such as generating heatmaps or saliency maps, may be applied to visualize which regions of the image influenced the model's predic-tions. This can help in understanding why the model made specific decisions.

5.**Fine-Tuning (Optional):**

If the model's performance is found to be unsatisfactory during testing, further fi-ne-tuning or optimization of hyperparameters may be considered to improve ac-curacy.The testing phase evaluates the model's real-world performance in accurately identifying lung cancer regions in multi-Model medical images and provides valua-ble insights into its reliability and effectiveness in clinical practice.

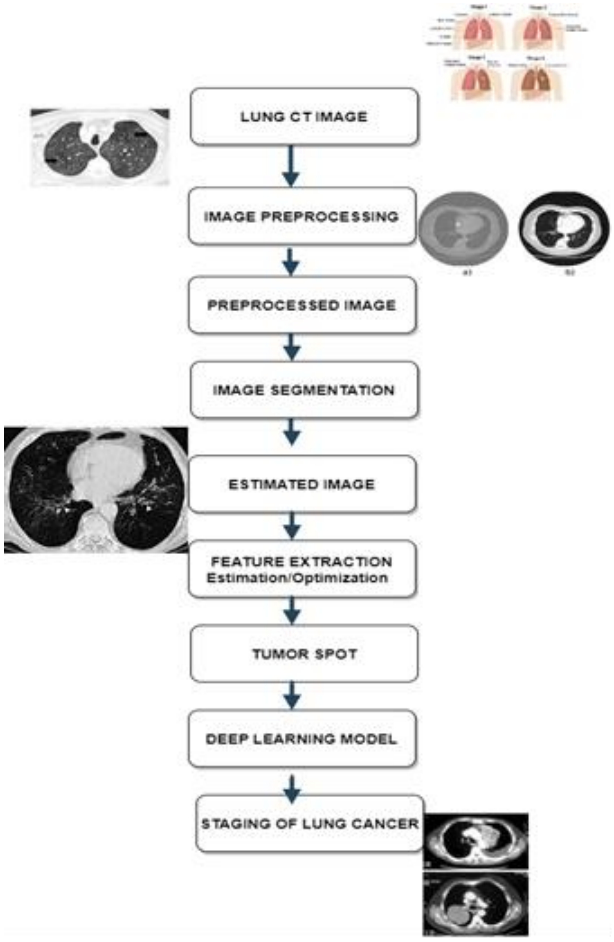


Fig. 3. Proposed Methodology for Lung Cancer Detection Model

4. Results And Discussions

The proposed method uses Deep Learning model for preprocessing segmentation and lung nodule detection. The Figure. 4 shows lung image of a malignant case using machine learning algorithm. The performance of the model achieved with accuracy of 99.48%, the accuracy and model loss is show in Figure 5 and Figure 6 respectively. The presence cancer using CT images like normal or malignant cases are listed in Figure 7.

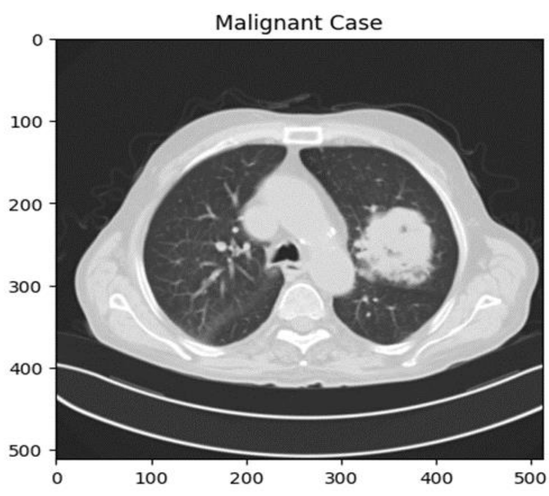


Fig.4. Malignant case in CT image

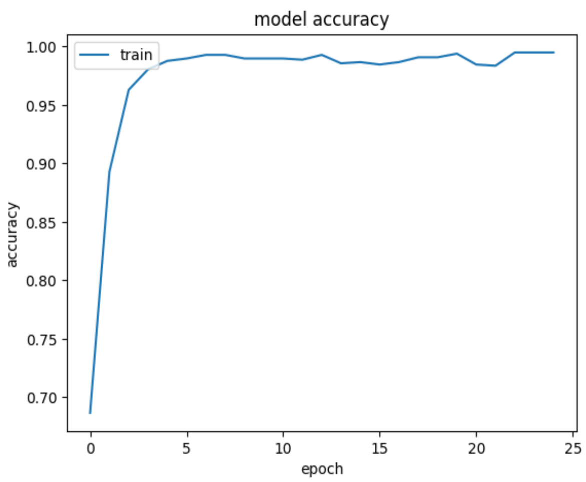


Fig.5. Model accuracy of the proposed system

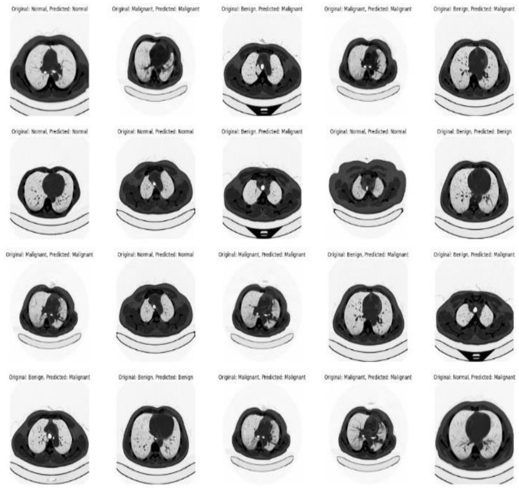


Fig .8. lung cancer detection of CT images for different cases

5. Conclusion

The proposed system discusses an automatic detection of lung cancer using Deep learning techniques using CT images. The machine learning algorithms and Deep learning techniques were used for segmentation and validation purpose. The adapted method accurately detects lung nodules with accuracy of 96% .In the future scope an android based Lung cancer detection application can be introduced which may help medical experts to detect lung cancer disease during initial stages

6. Authors' Note

The author(s) declare(s) that there is no conflict of interest regarding the publication of this article. Authors confirmed that the data and the paper are free of plagiarism

References

[1]. D. P. Naidich "Lung Cancer Detection and Characterization: Challenges and Solutions" SpringerLink DOI: 10.1007/978-3-642-18758-2_17.

[2]. Xiaodan Chen, Shouting Feng, Daru Pan" An improved approach of lung image segmentation based on watershed algorithm" - ICIMCS '15: Proceedings of the 7th International Conference on Internet Multimedia Computing and Service August 2015 Article No.: 39Pages 1–5https://doi.org/10.1145/2808492.2808531.

[3]. Manikandan T "Challenges in lung cancer detection using

computer-aided diagnosis (CAD) systems – a key for survival of patients. Archives of General Internal Medicine Volume 1 Issue 2.

[4]. Taruna Aggarwal; Asna Furqan; Kunal Kalra" Feature extraction and LDA based classification of lung nodules in chest CT scan images" Published in: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) DOI: 10.1109/ICACCI.2015.7275773.

[5]. P. B. Sangamithraa, S. Govindaraju "Lung tumour detection and classification using EK-Mean clustering" Published 23 March 2016 Computer Science, Medicine 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) DOI:10.1109/WISPNET.2016.7566533.

[6]. Wei Li, Peng Cao, Dazhe Zhao and Junbo Wang "Pulmonary Nodule Classification with Deep Convolutional Neural Networks on Computed Tomography Images" Comput Math Methods Med. 2016; 2016: 6215085. Published online 2016 Dec 14. doi: 10.1155/2016/6215085.

[7].Zakaria Suliman Zubi, Rema Asheibani Saad "Improves Treatment Programs of Lung Cancer Using Data Mining Techniques" Journal of Software Engineering and Applications > Vol.7 No.2, February 2014 DOI: 10.4236/jsea.2014.72008.

[8]. Xinyan Li,Shouting Feng,Daru Pan" Enhanced lung segmentation in chest CT images based on kernel graph cuts" ICIMCS'16: Proceedings of the International Conference on Internet Multimedia Computing and ServiceAugust 2016 Pages 228–233https://doi.org/10.1145/3007669.3007690.

[9]. Hanan M. Amer,Fatma E. Z. Abou-Chadi,Sherif S. Kishk,Marwa I. Obayya" A Computer-Aided Early Detection System of Pulmonary Nodules in CT Scan Images" ICSIE '18: Proceedings of the 7th International Conference on Software and Information EngineeringMay 2018 Pages 81–86https://doi.org/10.1145/3220267.3220291.

[10]. Aristófanes C. Silva,Paulo Cezar P. Carvalho,Marcelo Gattass" Diagnosis of lung nodule using Gini coefficient and skeletonization in computerized tomography images" SAC '04: Proceedings of the 2004 ACM symposium on Applied computingMarch 2004 Pages 243–248https://doi.org/10.1145/967900.967954.

[11] H Witschi "A short history of lung cancer" PMID: 11606795 DOI: 10.1093/toxsci/64.1.4.

[12]. Orkun Furat1,Mingyan Wang, Matthias Neumann, Lukas Petrich, Matthias Weber, Carl E. Krill I and Volker Schmidt" Machine Learning Techniques for the Segmentation of Tomographic Image Data of Functional Materials" Research Article | Open Access Volume 2016 | Article ID 6215085 | https://doi.org/10.1155/2016/6215085.

[13]. E R Heitzman "The role of computed tomography in the diagnosis and management of lung cancer. An overview" MedlinePlus Health Information, The American College of Chest Physicians. Published by Elsevier Inc., 1986. 89(4): p. 5.

[14]. M A Howe, B H Gross" CT evaluation of the equivocal pulmonary nodule" Comput Radiol. 1987 Mar-Apr;11(2):61-7. doi: 10.1016/0730-4862(87)90012-6.

[15]. Vikul J. Pawar, Kailash D. Kharat, Suraj R. Pardeshi,Prashant D. Pathak "Lung Cancer Detection System Using Image Processing and Machine Learning Techniques" International Journal of Advanced Trends in Computer Science and Engineering nced Trends in Computer Science and Engineering.

[16]. Lulu Wang "Deep Learning Techniques to Diagnose Lung Cancer" Cancers 2022, 14, 5569. https://doi.org/10.3390/cancers14225569.

[17]. Sharmila Nageswaran , G Arunkumar , Anil Kumar Bisht , Shivlal Mewada , J N V R Swarup Kumar , Malik Jawarneh , Evans Asenso "Lung Cancer Classification and Prediction Using Machine Learning and Image Processing" Biomed Res Int. 2022; 2022: 1755460. Published online 2022 Aug 22. doi: 10.1155/2022/1755460.

[18]. Setio, A.A.A.; Ciompi, F.; Litjens, G.; Gerke, P.; Jacobs, C.; Van Riel, S.J.; Wille, M.M.W.; Naqibullah, M.; Sánchez, C.I.; Van Ginneken, B. Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks. IEEE Trans. Med. Imaging 2016, 35, 1160–1169. [Google Scholar] [CrossRef] [PubMed].

[19]. Naqi, S.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Lung nodule detection using polygon approximation and hybrid features from CT images. Curr. Med. Imaging 2018, 14, 108–117. [Google Scholar] [CrossRef]

[20] Paing, M.P.; Hamamoto, K.; Tungjitkusolmun, S.; Visitsattapongse, S.; Pintavirooj, C. Automatic detection of pulmonary nodules using three-dimensional chain coding and optimized random forest. Appl. Sci. 2020, 10, 2346. [Google Scholar] [CrossRef][Green Version].

[21]. Lung Image Database Consortium (LIDC). Available online: https://imaging.nci.nih.gov/ncia/login.jsf (accessed on 5 October 2022).

[22]. Armato Samuel, G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.A.; et al. Data from LIDC-IDRI. 2015. Available online: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI (accessed on 5 October 2022).

[23]. Setio, A.A.A.; Traverso, A.; de Bel, T.; Berens, M.S.N.; van den Bogaard, C.; Cerello, P.; Chen, H.; Dou, Q.; Fantacci, M.E.; Geurts, B.; et al. Validation, Comparison, and Combination of Algorithms for Automatic Detection of Pulmonary Nodules in Computed Tomography Images: The LUNA16 Challenge. Med. Image Anal. 2017, 42, 1–13. [CrossRef]

[24]. ELCAP Public Lung Image Database. 2014. Available online: http://www.via.cornell.edu/lungdb.html (accessed on 5 October 2022).

[25]. Pedrosa, J.; Aresta, G.; Ferreira, C.; Rodrigues, M.; Leito, P.; Carvalho, A.S.; Rebelo, J.; Negrao, E.; Ramos, I.; Cunha, A.; et al. LNDb: A Lung Nodule Database on Computed Tomography. arXiv 2019, arXiv:1911.08434.

[26]. Prasad, D.; Ujjwal, B.; Sanjay, T. LNCDS: A 2D-3D cascaded CNN approach for lung nodule classification, detection andsegmentation. Biomed. Signal Process. Control. 2021, 67, 102527.

[27]. Shiraishi, J.; Katsuragawa, S.; Ikezoe, A.; Matsumoto, T.; Kobayashi, T.; Komatsu, K.; Matsiu, M.; Fujita, H.; Kodera, Y.; Doi, K.Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristicanalysis of radiologists' detection of pulmonary nodules. Am. J. Roentgen. 2000, 174, 71–74. [CrossRef]

[PubMed]

[28] .Xiao, Z.; Liu, B.; Geng, L.; Zhang, F.; Liu, Y. Segmentation of lung nodules using improved 3D-Unet neural network. Symmetry 2020, 12, 1787. [CrossRef].

[29].Cao, H.C.; Liu, H.; Song, E.; Hung, C.C.; Ma, G.Z.; Xu, X.Y.; Jin, R.C.; Jianguo Lu, J.G. Dual-branch residual network for lung nodule segmentation. Appl. Soft Comput. 2020, 86, 105934. [CrossRef].

[30] Kashyap, M.; Panjwani, N.; Hasan, M.; Huang, C.; Bush, K.; Dong, P.; Zaky, S.; Chin, A.; Vitzthum, L.; Loo, B.; et al. Deep learning based identification and segmentation of lung tumors on computed tomography images. Int. J. Radiat. Oncol. Biol. Phys. 2021, 111(3S), E92–E93. [CrossRef].

[31] Chen, C.; Zhou, K.; Zha, M.; Qu, X.; Xiao, R. An effective deep neural network for lung lesions segmentation from COVID-19 CT images. IEEE Trans. Ind. Inform. 2021, 17, 6528–6538. [CrossRef]

[32] Zhang, M.; Li, H.; Pan, S.; Lyu, J.; Su, S. Convolutional neural networks based lung nodule classification: A surrogate-assisted evolutionary algorithm for hyperparameter optimization. IEEE Trans. Evol. Comput. 2021, 25, 869–882. [CrossRef].

[33]. Banu, S.F.; Sarker, M.; Abdel-Nasser, M.; Puig, D.; Raswan, H.A. AWEU-Net: An attention-aware weight excitation u-net for lung nodule segmentation. arXiv 2021, arXiv:2110.05144. [CrossRef]

[34]. Dutta, K. Densely connected recurrent residual (DENSE R2UNET) convolutional neural network for segmentation of lung CT images. arXiv 2021, arXiv:2102.00663.

[35] Lin, X.; Jiao, H.; Pang, Z.; Chen, H.;Wu,W.;Wang, X.; Xiong, L.; Chen, B.; Huang, Y.; Li, S.; et al. Lung cancer and granuloma identification using a deep learning model to extract 3-dimensional radiomics features in CT imaging. Clin. Lung Cancer 2021, 22, e756–e766. [CrossRef]

[36]. Kim, H.M.; Ko, T.; Young, C.I.; Myong, J.P. Asbestosis diagnosis algorithm combining the lung segmentation method and deep learning model in computed tomography image. Int. J. Med. Inform. 2022, 158, 104667. [CrossRef]

[37] Deep Learning Algorithm for Classification and Prediction of Lung Cancer using CT Scan Images published in 2019 5th International Conference on Computing Communication Control and Automation (ICCUBEA)

[38] Random Forest based Classification Model for Lung Cancer Prediction on Computer Tomography Images by D. Jayaraj published in 2019 International Conference on Systems and Inventive Technology (ICSSIT).

[39] Predicting Outcomes of Non small Cell Lung Cancer Using CT Image Features by Samuel H. Hawkins; John N. Korecki; Yoganand Balagurunathan; Yuhua Gu; Virendra Kumar published in 2169-3536 2014 IEEE. Translations

[40] Prediction of Lung Cancer from Low-Resolution Nodules in CT-Scan Images by using Deep Features by Anand Gupta; Sagorika Das; Tarasha Khurana; Kamakshi Suri published in 2018 International Conference on Advances in Computing, and Informatics (ICACCI)

[41] Lung Nodule Classification in CT Images Using Convolutional Neural Network by Gilu K Abraham; Preethi Bhaskaran; V. S Jayanthi Published in: 2019 9th International Conference on Advances in Computing and Communication (ICACC) .

[42] Mutiullah, M. Bari, A. Ahmed, M. Sabir and S. Neveed, "Lung Cancer Detection Digital Image Processing Techniques: A Review", Mehran University Research Journal Engineering & Technology, vol. 38, no. 2, pp. 351-360, April 2019.

[43] S. Makaju, P. W. C. Prasad, A. Alsadoon, A. K. Singh and A. Elchouemi, "Lung cancer detection using CT scan images", Procedia Computer Science, vol. 125, pp. 114.

[44]S. Nageswaran, G. Arunkumar, A. K. Bisht, S. Mewada, J. N. Kumar,M. Jawarneh, and E. Asenso, “Lung Cancer Classification and prediction using machine learning and image processing,” BioMed Research International, vol. 2022, pp. 1–8, 2022.

[45]J. W. Yang, D. H. Song, H. J. An, and S. B. Seo, “Classification of subtypes including LCNEC in lung cancer biopsy slides using convolutional neural network from scratch,” Scientific Reports, vol. 12, no. 1, 2022.

[46]Q. Wu and W. Zhao, “Small-cell lung cancer detection using a supervised machine learning algorithm,” 2017 International Symposium on Computer Science and Intelligent Controls (ISCSIC), 2017.

[47]A. Gupta, Z. Zuha, I. Ahmad, and Z. Ansari, “A study on prediction of lung cancer using machine learning algorithms,” 2022.

[48]O. Obulesu, S. Kallam, G. Dhiman, R. Patan, R. Kadiyala, Y. Raparthi, and S. Kautish, “Adaptive diagnosis of lung cancer by deep learning classification using Wilcoxon gain and generator,” Journal of Healthcare Engineering, vol. 2021, pp. 1–13, 2021.

[49]P. Ramesh, R. Karuppasamy, and S. Veerappapillai, “Machine learning driven drug repurposing strategy for identification of potential RET inhibitors against non-small cell lung cancer,” Medical Oncology, vol. 40, no. 1, 2022.

[50]M. Saminathan, M. Ramachandran, A. Kumar, K. Rajkumar, A. Khanna, and P. Singh, “A study on specific learning algorithms pertaining to classify lung cancer disease,” Expert Systems, vol. 39, no. 3, 2021.

[51]A. Kumar Dutta, “Detecting lung cancer using Machine Learning Techniques,” Intelligent Automation & Soft Computing, vol. 31, no. 2, pp. 1007–1023, 2022.

[52]M. Mamun, A. Farjana, M. Al Mamun, and M. S. Ahammed, “Lung cancer prediction model using Ensemble Learning Techniques and a systematic review analysis,” 2022 IEEE World AI IoT Congress (AIIoT), 2022.

[53]E. Dritsas and M. Trigka, “Lung cancer risk prediction with machine learning models,” Big Data and Cognitive Computing, vol. 6, no. 4, p. 139, 2022.

[54]N. Banerjee* and S. Das, “Machine learning techniques for prediction of lung cancer,” International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 6, pp. 241–249, 2020.

[55]G. M. Alshmrani, Q. Ni, R. Jiang, H. Pervaiz, and N. M. Elshennawy, “A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images,” Alexandria Engineering Journal, vol. 64, pp. 923–935, 2023

[56]A. S. U., F. R. P. P., A. Abraham, and D. Stephen, “Deep learning- based BoVW–CRNN model for lung tumor detection in nano- segmented CT images,” Electronics, vol. 12, no. 1, p. 14,

2022.

[57]K. Ramana, M. R. Kumar, K. Sreenivasulu, T. R. Gadekallu, S. Bhatia, P. Agarwal, and S. M. Idrees, “Early prediction of lung cancers using deep saliency capsule and pre-trained deep learning frameworks,” Frontiers in Oncology, vol. 12, 2022.

[58]C. Anil Kumar, S. Harish, P. Ravi, M. SVN, B. P. Kumar, V. Mohanavel, N. M. Alyami, S. S. Priya, and A. K. Asfaw, “Lung cancer prediction from text datasets using machine learning,” BioMed Research International, vol. 2022, pp. 1–10, 2022.

[59]M. Humayun, R. Sujatha, S. N. Almuayqil, and N. Z. Jhanjhi, “A transfer learning approach with a convolutional neural network for the classification of lung carcinoma,” Healthcare, vol. 10, no. 6, p. 1058, 2022.

[60]R. P.R., R. A. S. Nair, and V. G., “A comparative study of lung cancer detection using machine learning algorithms,” 2019 IEEEInternational Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019.

[61]G. Zaman Khan et al., "An Efficient Deep Learning Model based Diagnosis System for Lung Cancer Disease," 2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2023, pp. 1-6, doi: 10.1109/iCoMET57998.2023.10099357.