

A Review on Automatic Lung Lesion Detection from Various Imaging Techniques

Ayush Bhargava

Electronics & Communication
Engineering
Maulana Azad National Institute of
Technology
Bhopal, Madhya Pradesh, India
ayushbhargavaab@gmail.com

Madhu Shandilya

Electronics & Communication
Engineering
Maulana Azad National Institute of
Technology
Bhopal, Madhya Pradesh, India
madhu_shandilya@yahoo.in

Vijayshri Chaurasia

Electronics & Communication
Engineering
Maulana Azad National Institute of
Technology
Bhopal, Madhya Pradesh, India
vijayshree21@gmail.com

Abstract— Lung cancer is horrible disease that may takes human life. It is the most diagnosed cancer in the world that may considered as life threatening disease. If it can be diagnosed earlier then treatment is a solution that may saves human life. There are various imaging techniques through which this disease can be diagnosed and treated accordingly. But Computed Tomographical image commonly known as CT scan image is the better option for diagnosing disease with better level of accuracy as compare to the other imaging techniques such as X-Ray, Ultrasound and many more. If a disease can be diagnosed automatically then it saves medical professional time as well as human life. A routine checkup can become easier and it can be processed in less time. The intension of this paper is to review various previously implemented systems related to the automatic diagnosis of lung cancer. There are so many researches have been done that are based on CNN, DNN, Edge Detection and various image processing or machine learning techniques. The objective of this paper is to define the limitations and drawbacks of various systems that are lacking somewhere.

Keywords—Lung Cancer Detection, Lesion Classification, Imaging Techniques, CNN, DNN, X-Ray, CT Scan.

I. INTRODUCTION

Lung cancer is also one of the reasons for loosing human lives. It is hard to detect on ground level that it emerges and shows symptoms in last stage. Be that as it may, death rate can be diminished by early detection and treatment of the infection. There are various imaging tactics but CT imaging is better for lung cancer detection since it can detect as malignant or benign. Traditionally it could be diagnose only through medical professionals through CT Scan images of lungs. It was a time consuming process and patient has to wait for a day to obtaining his report. But now it has been enhanced and it can be diagnosed automatically through various techniques with distinct pre-processing techniques [1]. As of now, to help radiologists and doctors detect the cancer precisely then automatic diagnosis has become

supplement and promising apparatus [2]. There has been numerous framework created and research is still going on detection of lung cancer. In certain case, a few frameworks don't have satisfactory precision of detection and a few frameworks actually must be improved to accomplish most elevated exactness tending to 97 to 98 %. Image processing methods and AI procedures have been executed to detect lung cancer. It has been proper investigated; which technique is more suitable for developing automatic lung cancer detection in the next section. The general 5-year endurance rate for lung cancer patients increments from 14 to 49% on the off chance that the sickness is detected on schedule. Despite the fact that CT scan can be more productive than X-ray. An issue appeared to converge because of time requirement in detecting the presence of lung cancer in regards to on the few diagnosing strategy.

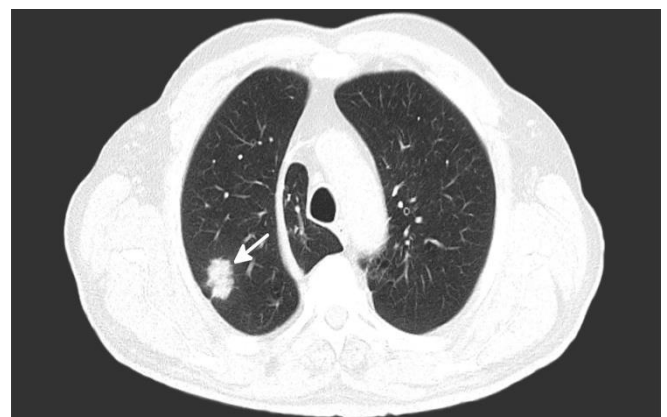


Fig. 1. Lung Cancer Lesion Detection in CT Scan Image [3]

Lung cancer is a sickness of unusual cells duplicating and developing into a growth. Cells can expand from the lungs through the circulatory system or lymph liquid that encompasses the lung tissue. For the most part, cancer cells regularly spread toward the focal point of the chest because

of the normal progression of lymph. Metastasis happens when cancer cells spread to different organs. The course of early detection of cancer assumes a significant part to keep cancer cells from duplicating and spreading. Past explores have been conducted for investigating lung cancer, for example, utilizing bunching strategy in microarray information.

II. RELATED WORKS

A. Related Works

P.B.Sangamithraa et al. [4] proposed a system which is based on EK-Mean Clustering. This framework first portions the ROI and afterward examinations have been achieved for impairment detection to look at the disease. Indeed, even with a few lung growth divisions that have been introduced and upgrading cancer division techniques which are as yet intriguing in light of the fact that lung growth CT images has certain characteristics, like variety in cancer appearance and dubious cancer limits. To resolve this issue, growth division technique for CT Images this dismantles non-upgrading lung cancers from lung tissues that has been done by clustering strategy. At first, the division of the CT images has been done by utilizing K-Means clustering strategy. To the bunched result, EK-Mean grouping is applied. Further the highlights like entropy, Contrast, Correlation, Homogeneity and the region are extracted from the tumorous piece of Fuzzy Ek-Means portioned Image. For highlight extraction, measurement technique called Gray Level Coherence Matrix (GLCM).

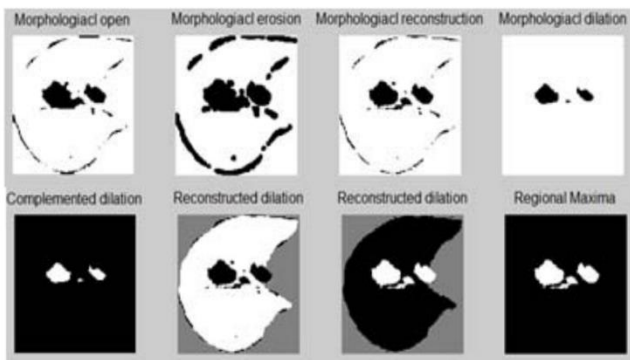


Fig. 2. EK-Means Algorithm [4]

Selin Uzelaltinbulat et al. [5] proposed a system which is based on median filter and morphological operation. The work includes distinctive image processing apparatuses which effectively accomplished the necessary objectives when it has been progressively applied to the system. The division framework includes various stages to at last arrive at its objective which is to fragment the lung growth. Image pre-processing occurs first where some improvement procedures are utilized to upgrade and decrease clamor in images. The following stage is the place where the various parts in the images are separated to section the cancer area. In this stage edge was selected consequently which guarantees the right selection of all images since the growth have distinctive dim levels in each image. One more strategy was likewise utilized here to eliminate the noise from the thresholded image i.e. erosion. At last, the lung

growth is precisely portioned by subtracting the thresholded and the other image.

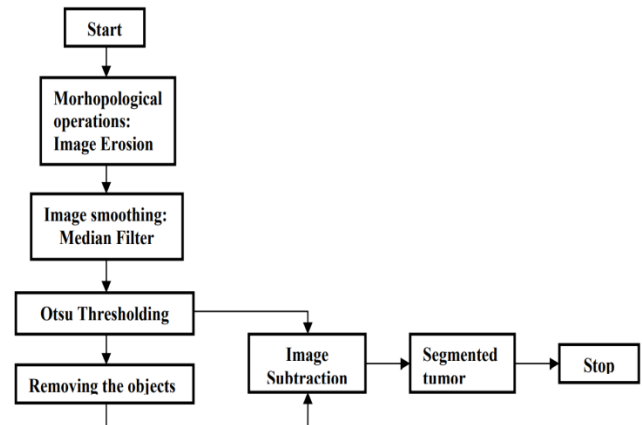


Fig. 3. Block diagram of proposed algorithm [5]

G.Niranjana et al. [5] reviewed various systems in the field of automatic lung cancer detection. By reviewing various researches about the lung lesion division techniques, it is concluded that applying edge on different CT images that segment the lung area from non-ROI. Staggered Thresholding is the least difficult and proficient strategies for lung division. For lesion division it gives preferred outcomes over other division techniques. Characterization of the lesion depends on shape and surface elements. Considering the above conversation, the component based methodology was viewed as better for classification.

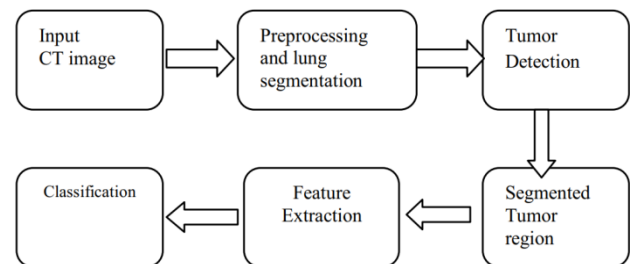


Fig. 4. Structure of Conventional System [6]

Qing Wu et al. [7] proposed a system which is based on Entropy Degrading method. In this review, the author proposed an EDM AI calculation with vectorized histogram elements to detect SCLC for early malignant cancer prediction. While it shows that EDM has sensibly great prediction, there is an enormous opportunity to get better before the calculation can be utilized in the clinical situation. The extreme objective of this review is to foster a clinical dynamic framework for radiologists to get readily prediction over vindictive lung cancer detection from SCLC with processed tomography (CT) imaging. Suren Makaju et al. [8] proposed a system which is based on Gaussian Filter, Median Filter and Watershed Segmentation. The proposed model has no satisfactory precision and doesn't classify the cancer affected cells. The proposed framework is utilized to detect the cancerous lesion from the lung CT scan image utilizing watershed division for detection and classifying the cancer cells as Malignant or benign. Proposed model detects the cancer with 92% precision which is higher than other model and classifier has obtained the 86.6% of precision.

Generally it can be seen the improvement in the proposed framework in contrast with current best model however, this proposed doesn't orders into various stages as stage I, II, III, IV of cancer. Consequently, as future degree improvement, this should be possible by executing arrangement in various stages. Likewise, further precision can be expanded by legitimate pre-processing and disposals of noise effectively.

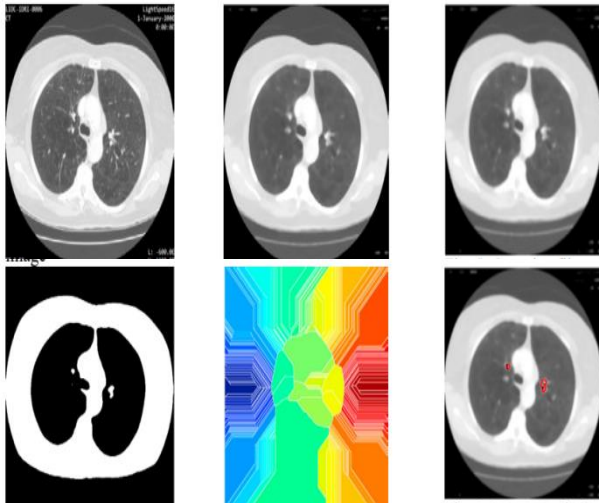


Fig. 5. System Diagnosis Phases [8]

SHANCHEN PANG et al. [9] et al. proposed a system which is based on DenseNet and Adaptive Boosting. In this paper, author built a model of lung cancer CT images with the help of DenseNet and AdaBoost calculations. First and foremost dataset is augmented by interpretation and transformation strategies, which can further develop the speculation capacity of the model and try not to predisposition the clustering results. Then, at that point, DenseNet is created to deal with the lung cancer datasets and characterize the collected information. At last adaboost calculation is utilized for classifying results to further develop better execution. Test results show that the model accomplishes bit better, and the precision of the test came to 89.85%. Later on clinical practice, this technique for lung cancer image classification can help the radiologists for lung cancer determination. It can further utilize with more lung cancer CT images to handle the order of lung cancer. Here authored transform data for enhancing the performance of the system. But DenseNet has densely connected network that may complicate the network and directly affect the computational efficiency. Aman Agarwal et al. [10] proposed a system which is based on Alexnet CNN. In this paper, author intended to find the malignant part of cancer from input lung CT scan image. Here the network has been trained with various shape and sizes of cancerous pattern from CT scanned lung images. It has been implemented based on the process of classification and take a decision whether it is malignant or benign. It is also extracting the cancer area from the image by applying multiple thresholding technique and certain morphological approaches and the final classification has been done using Alexnet. Here the system has been achieved 96 % of accuracy by classifying cancerous and non cancerous images. Alexnet is not a very deep model or network, it may

struggle to extract the features of the image that is why the performance of the system or network is poor.

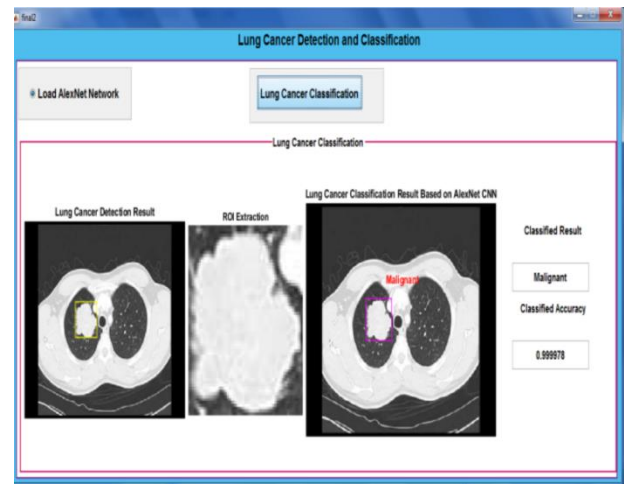


Fig. 6. System GUI [10]

Zirong Li et al. [11] proposed a system which is based on ResNet. In this paper, author portrayed a strategy to detect density from chest image by using ResNet. By analyzing and testing the other networks and the architecture of the networks, it has been analyzed that RCNN architecture is bit faster and can produce the result more efficiently and the component extraction a piece of the network utilizing RESNET that shows better execution. It has been classified whether the image belongs to the healthy lung or cancerous. It has been tested that it is better than the Alexnet and the number of convolutional layers are bit higher than the Alexnet. But there is a problem with this paper is that, it is using chest X-ray instead of CT scan images of lung. X-ray image of lung cancer is not able to pertain good level of accuracy. X-ray has not potential to highlight the cancerous area of lung cancer. It is better to use CT scan images that can produce high level of accuracy as compare to the X-ray images.

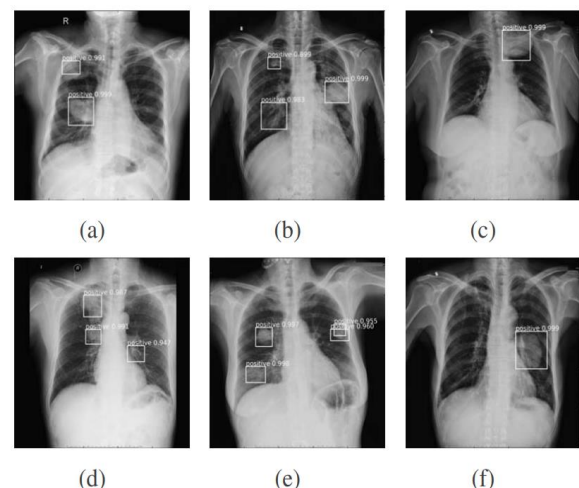


Fig. 7. Lesion Detection from Chest X-ray Images [11]

Hui Shi et al. [12] proposed a system which is based on VGG16. In this paper, 2 dimensional CNN has been used to extract the lung cancer lesions and for declaring the health of the lung. Here the system has been trained with VGG16

where 16 layers have been used with distinct shape and sizes of the lesions. They are using anchors and anchors are responsible to predict the region of the lesion and predict the diagnosis score of the system. System has claimed to detect the lesion of various shapes and sizes with high precision. But system uses poor datasets for training and testing modules i.e. Luna16. It has average sensitive lesions and small lesions are not able to detect from this datasets. Here the system achieved 82.62% of accuracy which is not very effective as compare to the other models. Moffy Vas et al. [13] proposed a system which is based on Feed Forward ANN. The approach which has been adapted in this paper expects to foster a framework for lung cancer detection. Here author uses median filter for removing the clamor in the images ended up being a triumph. System also uses morphological approaches for additionally contributed towards satisfactory outcomes during the classification and segmentation. ANN ended up being a decent classifier with expected precision. The strategy embraced in this project brought about a precision of 92% for the emergency clinic database. This framework targets expanding the exactness and speed of the lung cancer detection framework. It likewise helps in detecting the cancer at prior stages. Here system converted all the lung images into binary image that decreases the size of the image and then it applies erosion technique for eliminating the noise from the image which is encounter in the category of morphological operation. It eliminates the background and foreground noise of the image but it may also erode the sensitive data because blood vessels, veins are very much similar to the lesion.

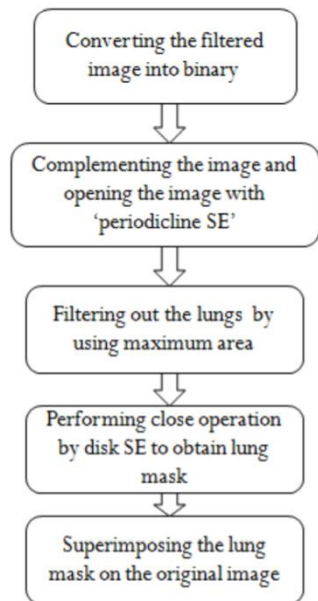


Fig. 8. Segmentation Stage [13]

Wadood Abdul et al. [14] proposed a system which is based on CNN model. System pertains the lesion from the structure and texture and train the network accordingly. Here system tries to enhance the back propagation of the network through which a system can take better decision while extracting the cancerous area from CT lung images. Author used LIDC-IDRI dataset for testing and obtained the result accordingly. System compared its performance with the state of art algorithm from various perspectives. Author complicated the network for getting effective result but it

degraded the computation efficiency as well as the computation time. Sanjukta Rani Jena et al. [15] proposed a system which is based on texture based analysis of cancerous cells. System uses various image pre-processing techniques for extracting the lesion and the location of the lesion. System also uses Local Binary Patterns for segmentation but it is a very conventional method for segmenting the image or region of interest and it loses the sensitive information from the image. Local binary patterns compare the image information with threshold value and erode and dilate the pixel accordingly. At the final stage of the diagnosis; Support Vector Machine has been used to classify the lesion but SVM does not work effectively with local binary patterns because of the conventional approach. For performing the testing phase, system attained certain normal as well as abnormal lung images and extract the histogram features of all the images. Total of 80 images have been used for classification. They used four parameters true positive, true negative, false positive and false negative. But they claim 100% of accuracy which is not feasible at all. Because no one can claim 100% of accuracy, it may degrade in a point when testing samples increases. System cannot be effective for all kind of images or samples. It may get distracted and return false alarm rate when distinct patterns have been received.

III. PROBLEM IDENTIFICATION

There are various researches have been done in the field of automatic lung cancer detection. Certain researchers use densely populated network that may complicate the network that degraded the computation efficiency of the system and execution time get increases and training phase also take longer time. Some of the researches are based on conventional CNN which is deeper but not effective to scan all kind of features from the lung CT images. A network should be light in weight by using light weight filters and trained strongly. Because poor trained system is not effective to attain good accuracy. Few researchers use LUNA16 dataset which is poor while extracting small lesions. There are so many datasets available that can pertain higher accuracy as compare to the LUNA16. Many researchers also use Median Filter for eliminating the noise from the input lung image but median filter is very conventional and simple filtration technique which is not effective for clinical data.

Table No. I Problem Findings

Method	Finding
ResNet	ResNet usually takes several weeks for the completion of training that makes it practically not feasible in real-world applications
VGG-16	VGG16 has so many weight and biased parameters due to that the model becomes very heavy in size
DenseNet	Excessive dense connections may complicated the network that affected computation efficiency

AlexNet	AlexNet is not very deep model due to that it struggles to scan for all the features that resulting poor in performing
CNN	Conventional CNN model is poor in training and building heavy network that directly affects the execution time
LBP	LBP is a very conventional approach that works on the basis of binary values that may erode the sensitive information

Table I shows the comparison among various methods which have been used to implement the automatic lung cancer detection. Table shows the problem which has been identified at each model of the system and where it suffers and what kind of perspective should be retain in mind to develop an ideal system that can diagnose lung cancer whether it is malignant or benign from lung CT scan images. System require a better approach that can train the system with less filters or light weight filter with proper training through which percision can be calculated high with less false alarm rate. In Table II, result comparison has been discussed and shows how much accuracy they pertained during testing phase of the system.

Table No. II Result Comparison

	Method	Modality	Precision
Zirong Li [11]	ResNet	X-Ray	0.5238
Hui Shi [12]	VGG-16	CT	0.826
Shanchen Pang [9]	DenseNet	CT	0.8985
P.B.Sangamithraa [4]	EK-Means	CT	0.9087
Moffy Vas [13]	FF-ANN	CT	0.92
Aman Agarwal [10]	AlexNet	CT	0.96

IV. CONCLUSION & FUTURE SCOPE

Automatic lung cancer detection is a process through which lesion can be detected from lung images and the area can also be extracted that states the size of the lesion which helps medical professionals to examine the disease level. This paper reviewed various researches and concluded at a point where problem has been detected with the current scenario. Poor network and heavy network both generate complications while training and testing the data. If it can be proceed through machine learning approach then network should be light weighted and less filters or light weighted filters should be used. System should not compromise with the layers and each layer should produce efficient result and system should pertain high level of accuracy at the end of

the layers. If system uses edge detection technique or any other image processing technique then it should be highly sensitive to the data because clinical data is very much sensitive and it should not be erode at all.

REFERENCES

- [1] A.M. Gindi, T.A. Al Attiatalla, M.M. Sami "A Comparative Study for Comparing Two Feature Extraction Methods and Two Classifiers in Classification of Earlystage Lung Cancer Diagnosis of chest x-ray images." *Journal of American Science*, 10 (6) (2014), pp. 13-22
- [2] K. Suzuki, M. Kusumoto, S.I. Watanabe, R. Tsuchiya, H. Asamura "Radiologic classification of small adenocarcinoma of the lung: radiologic-pathologic correlation and its prognostic impact" *The Annals of Thoracic Surgery*, 81 (2) (2006), pp. 413-419
- [3] Su CC, et al "Impact of low-dose computed tomography screening for primary lung cancer on subsequent risk of brain metastasis" *J Thorac Oncol* 2021; DOI: 10.1016/j.jtho.2021.05.010.
- [4] P. B. Sangamithraa and S. Govindaraju, "Lung tumour detection and classification using EK-Mean clustering," 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2016, pp. 2201-2206, doi: 10.1109/WiSPNET.2016.7566533.
- [5] Selin Uzelaltinbulat, Buse Ugur, Lung tumor segmentation algorithm, *Procedia Computer Science*, Volume 120, 2017, Pages 140-147.
- [6] G. Niranjana and M. Ponnaivaikko, "A Review on Image Processing Methods in Detecting Lung Cancer Using CT Images," 2017 International Conference on Technical Advancements in Computers and Communications (ICTACC), 2017, pp. 18-25, doi: 10.1109/ICTACC.2017.16.
- [7] 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, pp. 88-91, doi: 10.1109/ISCSIC.2017.22.
- [8] Suren Makaju, P.W.C. Prasad, Abeer Alsadoon, A.K. Singh, A. Elchouemi, Lung Cancer Detection using CT Scan Images, *Procedia Computer Science*, Volume 125, 2018, Pages 107-114.
- [9] Pang, Shanchen; Zhang, Yaqin; Ding, Mao; Wang, Xun; Xie, Xianjin (2020). A Deep Model for Lung Cancer Type Identification by Densely Connected Convolutional Networks and Adaptive Boosting. *IEEE Access*, 8(), 4799–4805.
- [10] Agarwal, Aman & Patni, Kritik & Devarajan, Rajeswari. (2021). Lung Cancer Detection and Classification Based on Alexnet CNN. 10.1109/ICCES51350.2021.9489033.
- [11] Li, Zirong; Li, Lian (2017). [IEEE 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) - Kansas City, MO (2017.11.13-2017.11.16)] 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) - A novel method for lung masses detection and location based on deep learning. , (), 963–969.
- [12] Shi, Hui; Peng, Zhenwei; Wan, Honglin (2019). [IEEE 2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT) - Jinan, China (2019.10.18-2019.10.20)] 2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT) - Pulmonary Nodules Detection Based on CNN Multi-scale Feature Fusion. , (), 86–90.
- [13] M. Vas and A. Dessai, "Lung cancer detection system using lung CT image processing," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), 2017, pp. 1-5, doi: 10.1109/ICCUBEA.2017.8463851.
- [14] W. Abdul, "An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network," 2020 13th International Conference on Developments in eSystems Engineering (DeSE), 2020, pp. 443-446, doi: 10.1109/DeSE51703.2020.9450778.
- [15] D. Sharma and G. Jindal, "Computer aided diagnosis system for detection of lung cancer in CT scan images," *Int. J. Comput. Elect. Eng.*, vol. 3, no. 5, pp. 714–718, Sep. 2011.
- [16] T. Yoshiya, T. Mimae, and Y. Tsutani, "Prognostic role of subtype classification in small-sized pathologic N0 invasive lung adenocarcinoma," *Ann. Thoracic Surg.*, vol. 102, no. 5, pp. 1668–1673, 2016
- [17] R. K. Samala, H.-P. Chan, C. Richter, L. Hadjiiski, C. Zhou, and J. Wei, "Analysis of deep convolutional features for detection of lung

- nodules in computed tomography,” in *Medical Imaging 2019: Computer-Aided Diagnosis*, vol. 10950, p. 109500Q, International Society for Optics and Photonics, 2019.
- [18] D. Kumar, A. Wong, and D. A. Clausi, “Lung nodule classification using deep features in CT images,” in *Conference on Computer and Robot Vision*, pp. 133–138, 2015.
- [19] M. Tan, R. Deklerck, B. Jansen, M. Bister, and J. Cornelis, “A novel computer-aided lung nodule detection system for CT images,” *Medical Physics*, vol. 38, no. 10, pp. 5630–5645, 2011.
- [20] A. A. A. Setio, A. Traverso, T. de Bel, M. S. N. Berens, C. van den Bogaard, P. Cerello, H. Chen, Q. Dou, M. E. Fantacci, B. Geurts, R. van der Gugten, P. A. Heng, B. Jansen, M. M. J. de Kaste, V. Kotov, J. Y.-H. Lin, J. T. M. C. Manders, A. S’oñora-Mengana, J. C. Garc’ia- Naranjo, E. Papavasileiou, M. Prokop, M. Saletta, C. M. Schaefer- Prokop, E. T. Scholten, L. Scholten, M. M. Snoeren, E. L. Torres J. Vandemeulebroucke, N. Walasek, G. C. A. Zuidhof, B. van Ginneken, and C. Jacobs, “Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge,” *Medical Image Analysis*, vol. 42, pp. 1–13, 2017.