International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Enhancing Lung Cancer Detection with Deep Learning: A CT Image Classification Approach

Jeevika K S[1], Dr. Savitha S K[2]

[1]PG Student, Department of Computer Science Engineering, Bangalore Institute of Technology, Bengaluru, India

[2]Professor, Department of Computer Science Engineering, Bangalore Institute of Technology, Bengaluru, India

Abstract: ***Lung cancer is a highly perilous illness ranking as one of the primary causes of disease and death, particularly when diagnosed in its initial stages. It presents significant challenges, as it is often only discernible after it has already diffused. This study proposes a lung cancer prognostication framework that uses deep learning to enhance the accuracy of cancer forecasting and disease determination, thereby enabling personalized treatment approaches based on disease severity. It consists of various steps, including image preprocessing and segmentation of lung CT image features extracted from the segmented images. Three different models, namely a DCNN model, a DCDNN model, and an ANN model, were employed for image classification, and a deep convolutional neural network (DCNN) was employed to detect lung diagnosis based on the extracted feature evaluation results showing the best accuracy of 99.41% in accurately discerning the presence or absence of lung cancer. The GAN model generates realistic lung CT scan images by training a generator to produce authentic images, and a discriminator to distinguish between real and fake images. The outcome of the system depends on the quality of the data, and a well-trained DCNN through training, validation, and testing on diverse datasets is crucial to ensure the reliability and generalizability of the model.***

Keywords: Lung cancer detection, Deep learning, Deep convolutional neural network

1. Introduction

Lung cancer is a complex and heterogeneous ailment characterized by a rise in the number of cells in lung tissues. This stands as the main reason for most cancer-related deaths worldwide, accounting for a sizable proportion of cancer-associated morbidity and mortality. Detection assumes a crucial role in identifying the existence of lung cancer at an early stage and facilitating timely treatment. This study focuses on the development and improvement of detection methods using different techniques. This study provides an input CT image for image preprocessing, which includes grayscale conversion for noise removal, histogram calculation, and image quality enhancement for more clearly visible images after image enhancement. The second step is image segmentation, which detects the edge using Canny edge detection, and lung segmentation techniques using K-means clustering to separate the background and foreground. Morphological operations to refine the lung mask. The original image binary threshold image eroded, dilated color-labeled image-segmented lung mask, and segmented lung area in the original image are displayed. After the segmentation lung feature extraction process, which is not displayed in the system, the model directly underwent analysis assessment. In the third step, the classification model was trained and evaluated. Collect a dataset of lung images with labels normal or cancer and split the dataset into three sets. In the evaluation process, it provides access to the training history to retrieve the training and validation accuracy, and the loss values are plotted using Matplotlib. The models used in this study were the DCNN,DCDNN, and ANN. Every model exhibited good accuracy and loss percentage and plotted graphs. Further details are provided in the following sections. In the fourth step, the GAN model predicts, the structure of the image batch, class names, labels, and filenames in this understood.

During the ultimate step, the Streamlit app for lung cancer detection allows users to upload images and predict cancer. User-friendly interface that interacts with an application on a web browser.

1.1 Aims and Objectives

a) The aim of this project is to enhance the precision and effectiveness of lung cancerdiagnosis by advanced deep learning methodologies.
b) To develop an application that detects and properly classifies Lung cancer in CT scan images using a DCNN.

1.2 Flow Diagram

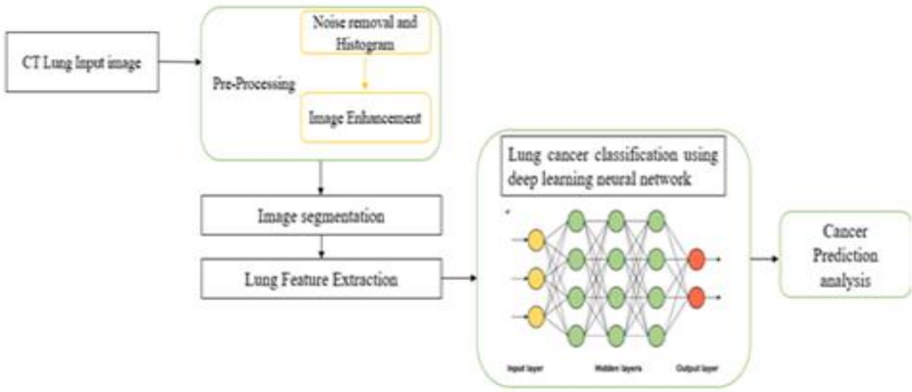

Figure 1: Proposed system

2. Proposed Methodology

A. Image Preprocessing

Image preprocessing plays a fundamental role in the accurate analysis of medical images, particularly in detecting malignancy using CT scans. This section focuses on the key preprocessing steps: grey conversion, histogram analysis, and image quality enhancement. Gray conversion simplifies the image by converting it to grayscale, thereby facilitating the subsequent processing. Histogram analysis provides insights into the pixel intensity distribution, aiding in the threshold determination for image segmentation. Image quality enhancement techniques enhance visibility and reduce noise and artifacts. The cumulative distribution function is normalized and scaled to the range [0,25]. By implementing these preprocessing steps, CT images were optimized for accurate lung cancer detection, ensuring improved analysis and more reliable results.

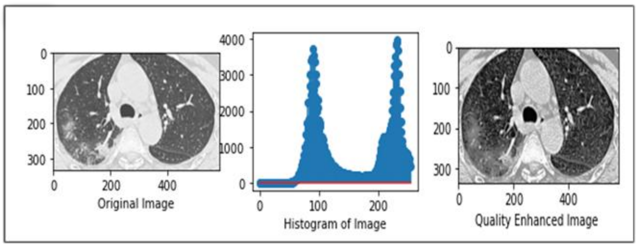

Figure 2: Input image and Quality enhanced image

B. Image Segmentation

Image segmentation is an important step in this study. It involves precise identification and separation of the cancerous region within the enhanced tomography images. To achieve this, the k-means algorithm is typically employed.

The segmentation process begins by examining the pixel similarity in the lung images. The picture waspartitioned into multiple subregions, allowing the prediction and localization of the affected area. This is accomplished by analyzing the similarity of pixel values and grouping superpixels that exhibit similar characteristics. The segmentation algorithm effectively identifies the outline of the malignant region by detecting abrupt changes in pixel values. Pixel similarity is a fundamental attribute in image segmentation and critical analysis of visual data in accurately shaping clusters. Through quantitative evaluation, the algorithm determines the level of similarity between pixels, facilitating the correctdivision of the cancerous region. This evaluation considered the enhanced quality of the image pixels, ensuring that the segmentation process was performed effectively.

An undirected graph was constructed to represent the abnormal regions in the lung image. This graph provides a visual representation of the segmented areas, aiding the identification and analysis of cancerous regions.

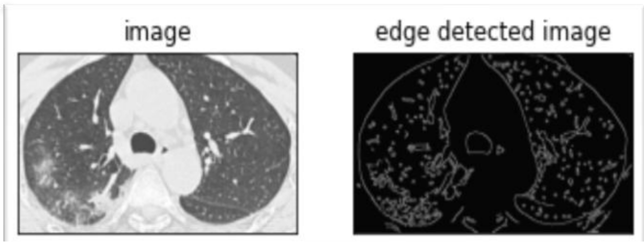

Figure 3: Edge-detected image

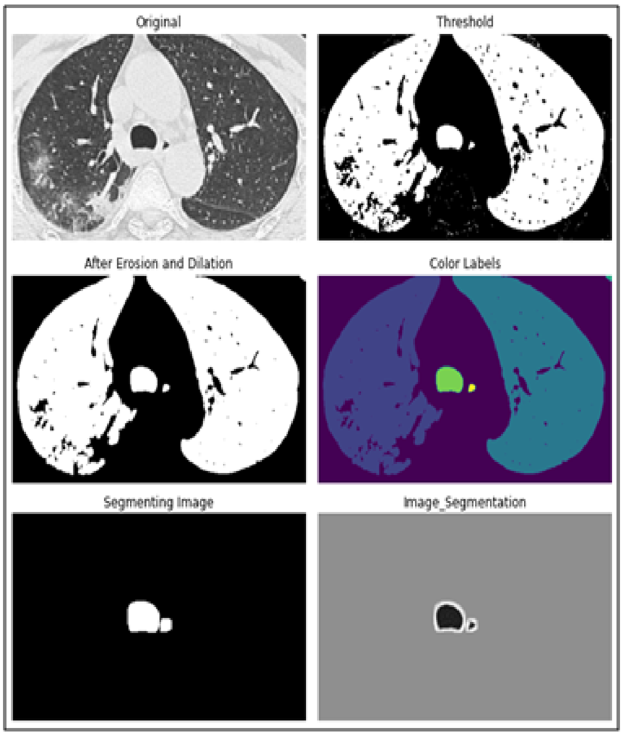

Figure 4: Segmented images

C. Image Classification

The final stage involves the recognition of malignant. Detection was performed utilizing algorithms with CT scan image input. Dataset of lung images with labels indicating cancer or not. After splitting the dataset into training, validation, and test sets as a folder. The algorithms used are deep convolutional neural network, double convolutional deep neural network, and artificial neural network.

At this stage, The Generative Adversarial Network (GAN) model acts as a discriminator, tasked to distinguish between real and synthetic images, with the ultimate goal of producing life-like images capable of misleading discriminators Through training the GAN model enhances the fidelity of chest CT scan images which were then saved and displayed. The saved model is readily deployable for further analysis, evaluation, or application-specific tasks.

After the user uploads an image of the lung tissue, the image is preprocessed and passed through a pre-trained model. The model predicts the probability of an image belonging to each class: "CANCER" or "NORMAL." The class with the highest probability was considered the predicted class for the input image. The predicted class was then displayed on the web interface to indicate whether the image was classified as cancerous or normal lung tissue.

It should be emphasized that the accuracy and reliability of cancer identification depend on the excellence of the pre-trained model, the dataset on which it was trained, and the diversity and quality of the images used for testing and evaluation. The performance can vary based on these factors, and a well-validated and reliable model is recommended for the accurate identification of lung cancer.

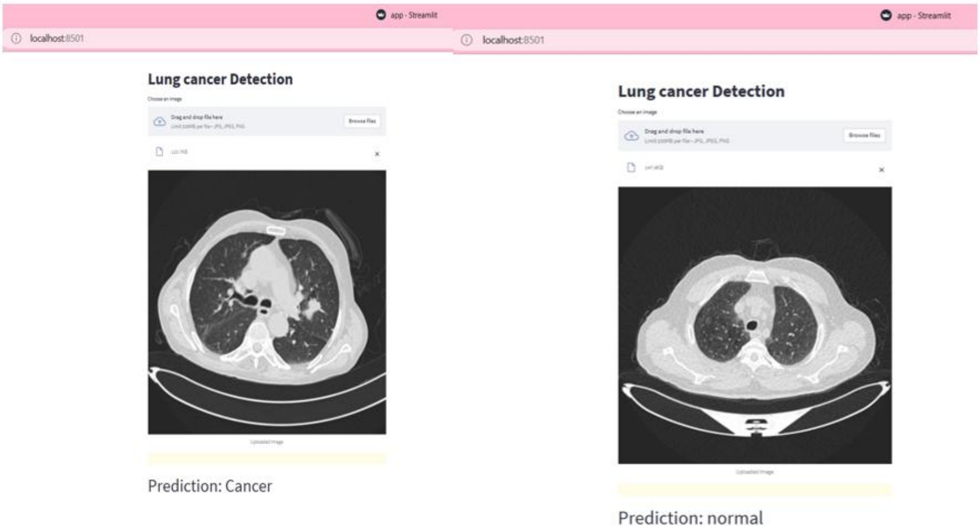

Figure 5: Detection of cancer or not

3. Results

This paper work finds that image preprocessing shows the results in Figure (2) shows the image quality enhancement, Figure (3) shows the edge detection using the canny edge detection, Figure (4) shows the lung segmentation using k-means clustering, Figure (5) shows the detected cancer or not, Figure (6) DCNN (7) DCDNN (8) ANN shows the Accuracy and loss graph for training and validation.

Table 1: Accuracy analysis

Methods	Final Accuracy	Best Accuracy
DCNN	96.6	99.41
DCDNN	93.2	95.15
ANN	57.4	80.0

B. Graph for Accuracy and Loss (training and validation)

A Deep convolutional neural network model consists of multiple convolutional, pooling, and dense layers. After training, validation, and testing, the data were obtained. Training the model saves it, evaluates its performance using accuracy metrics, and plots graphs.

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

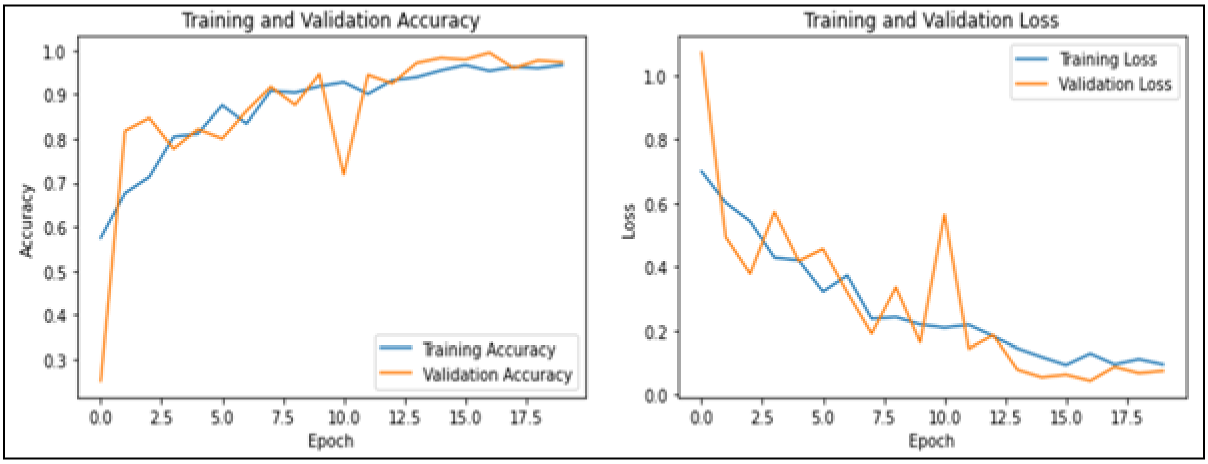

Figure 6: DCNN Graph

A Double convolutional deep neural network it consists of convolutional and pooling layers, followed by fully connected layers. After training, validation, and testing, the data were analyzed. Training the model saves the model, evaluates its performance using accuracy metrics, and plots the training and validation accuracy and loss curves.

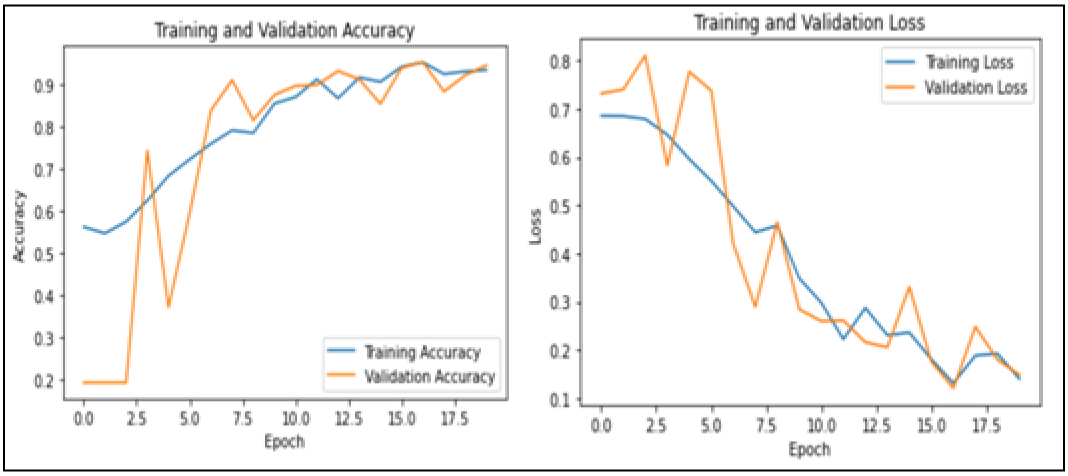

Figure 7: DCDNN graph

An Artificial neural network in this model consists of a flattened input layer, a dense hidden layer, and an output layer. After training, validation, and testing data. Training the model saved evaluates its performance using the accuracy metrics, and plots it is shown in below figure 8.

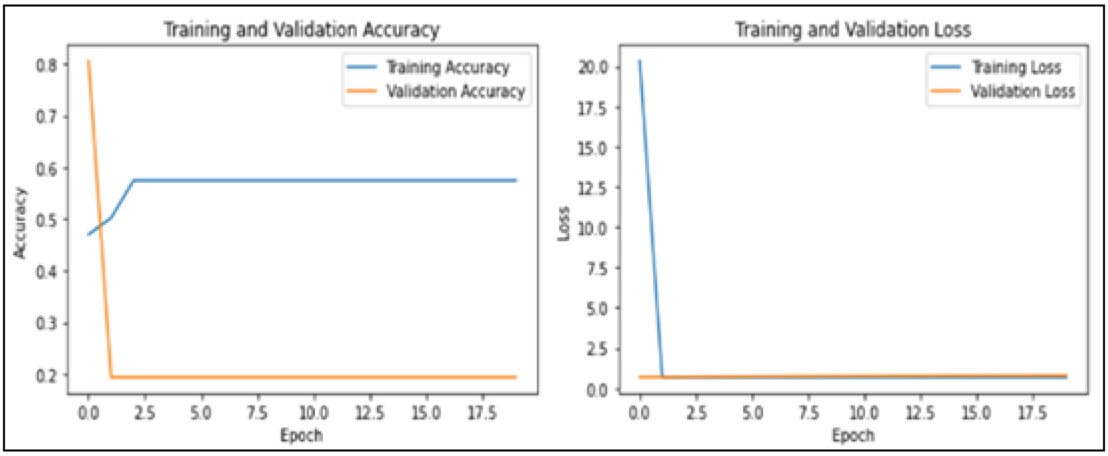

Figure 8: ANN graph

4. Conclusion

In conclusion, this paper developed using deep learning techniques showed promising results. The main prediction of the presence or absence of lung cancer is achieving high accuracy and enabling personalized treatment approaches based on disease severity. The ability of the system to classify CT images provides a valuable tool for detection.

The system development begins with image preprocessing, segmentation, and image classification using different algorithms or models, including DCNN, DCDNN, and ANN. In this model, the evaluation results show that the best accuracy is achieved by the DCNN, which yields 99.41% accuracy and detects lung cancer.

Future research should prioritize the early detection and exploration of the stage of cancer. Integration with advanced medical imaging technologies and clinical data could also be explored to augment the predictive capabilities of the system and enable more comprehensive lung cancer analysis.

References

[1] Heng Yu, Zhiqing Zhou, And Qiming Wang: "Deep Learning Assisted Predict of Lung Cancer on Computed Tomography Images Using the Adaptive Hierarchical Heuristic Mathematical Model" In the IEEE Access, Volume 08 2020, pp 86401-86411, Digital Object Identifier 10.1109/ACCESS.2020.2992645

[2] Mohammad A. Alzubaidi, MwaffaqOtoom ,(Senior Member, Ieee), And Hamza Jaradat, "Comprehensive and Comparative Global and Local Feature Extraction Framework for Lung Cancer Detection Using CT Scan Images", In the IEEE Access, Volume 9, 2021, pp 158140- 158154, Digital Object Identifier 10.1109/ACCESS.2021.3129597

[3] Wafaa Alakwaa, Mohammad Nassef, Amr Badr, "Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)", in Proceedings of (IJACSA) International Journal of Advanced Computer Science and Applications, Volume. 8, 2017, pp 409-417 https://www.ijacsa.thesai.org/

[4] PrathyushaChalasani, S Rajesh.: "Lung CT Image Recognition using Deep Learning Techniques to Detect Lung Cancer" .in International Journal of Emerging Trends in Engineering Research**.**, Volume 8. No. 7, July 2020, ISSN 2347–3983, https://doi.org/10.30534/ijeter/2020/113872020

[5] Rasika N. Kachore, Kivita Singh: "Lung Nodule Detection", in the International Journal of Scientific Research in Science, Engineering and Technology, IJSRSET, Volume 3, Issue 2, ISSN: 2395-1990, Online ISSN : 2394-4099.

[6] Yaozhi Lu, Shahab Aslani, Mark Emberton, Daniel C. Alexander, And Joseph Jacob: "Deep Learning-Based Long Term Mortality Prediction in the National Lung Screening Trial", in the proceedings of IEEE Access, Year – 2022, VOLUME 10, 2022, pp: 34369-34378, Digital Object Identifier 10.1109/ACCESS.2022.3161954

[7] B. Keerthi Samhitha, SujaCherukullapurath Mana, Jithina Jose, R.Vignesh, D. Deepa: "Prediction of Lung Cancer Using Convolutional Neural Network (CNN)", in Proceedings of the International Journal of Advanced Trends in Computer Science and Engineering, Volume 9, No.3, 2020,ISSN 2278-3091, https://doi.org/10.30534/ijatcse/2020/135932020

[8] Surbhi Vijh, Prashant Gaurav, Hari Mohan Pandey: "Hybrid bio-inspired algorithm and convolutional neural network for automatic lung tumor detection", in the springer, Year of publication 19 September 2020, https://doi.org/10.1007/s00521-020-05362-z

[9] Dr. M. Sangeetha, P Sangeetha, P. Pavithra: "Classification of Lung Cancer using Deep Learning Algorithm", in the Conference proceedings International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published by, www.ijert.org RTICCT- 2020.

[10] S. Perumal, T. Velmurugan, "Lung cancer detection and classification on CT scan images using enhanced artificial bee colony optimization", in the International Journal of Engineering & Technology Website: www.sciencepubco.com/index.php/IJET, Volume 7, Year-2018, pp 74-79.

[11] Onur Ozdemir , Member, IEEE, Rebecca L. Russell, and Andrew A. Berlin, Member, IEEE "A 3D Probabilistic Deep Learning System for Detection and Diagnosis of Lung Cancer Using Low-Dose CT Scans", in the IEEE TRANSACTIONS ON MEDICAL IMAGING, Volume. 39, NO. 5, MAY 2020, pp: 1419-1429

[12] KhevnaVasani, Ayushi Shah: "Lung Cancer Detection Using CT Scan Images", in the International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056 Volume: 08 Issue: 04 Apr 2021 www.irjet.net p-ISSN: 2395-0072

[13] Longxi Zhou, Zhongxiao Li, Juexiao Zhou, Haoyang Li, Yupeng Chen, Yuxin Huang, Dexuan Xie, Lintao Zhao, Ming Fan , Shahrukh Hashmi, Faisal Abdelkareem , Riham Eiada , Xigang Xiao, Lihua Li, "Adaptive Local Ternary Pattern on Parameter Optimized-Faster Region Convolutional Neural Network for Pulmonary Emphysema Diagnosis", in the IEEE Access, Volume 9,2021, pp:114135-114152, DOI: 10.1109/ACCESS.2021.3105114

[14] ChanunyaLoraksa, SirimaMongkolsomlit, NitikarnNimsuk, MeenutUscharapong, and Piya Kiatisevi: "Development of the Osteosarcoma Lung Nodules Detection Model Based on SSD-VGG16 and Competency Comparing with Traditional Method", in the IEEE Access, VOLUME 10, 2022, pp: 65496-65506, Digital Object Identifier 10.1109/ACCESS.2022.3183604

[15] Aishwarya Kalra, Brijmohan Singh, Himanshu Chauhan, "An Approach for Lung Cancer Detection using Deep Learning", in the International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056 Volume: 07, Issue: 09 Sep 2020 www.irjet.net p-ISSN: 2395-0072

[16] P. Mohamed Shakeel, M. A. Burhanuddin, Mohammad Ishak Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier", published in the springer pp:349579–9592 https://doi.org/10.1007/s00521-020-04842-6

[17] Zhang Li, Jiehua Zhang, Tao Tan, Xichao Teng, Xiaoliang Sun, Hong Zhao, Lihong Liu, Yang Xiao, "Deep Learning Methods for Lung Cancer Segmentation in Whole-Slide Histopathology Images-The ACDC@LungHP Challenge 2019", in the IEEE

Journal Of Biomedical And Health Informatics, Volume. 25, No. 2, February 2021
[18] Lulu Wang: "Deep Learning Techniques to Diagnose Lung Cancer", in the MDPI, Volume 14 , Issue 22 ,DOI: 10.3390/cancers14225569
[19] Imdad Ali, Muhammad Muzammil, (Member, IEEE), IHSAN UL HAQ, (Member, IEEE), AMIR A. Khaliq, (Member, IEEE), "Efficient Lung Nodule Classification Using Transferable Texture Convolutional Neural Network", In the IEEE Access, Volume 8, 2020, pp: 175859- 175870, Digital Object Identifier 10.1109/ACCESS.2020.3026080, Year of publication – 2020
[20] Longxi Zhou, Zhongxiao Li, Juexiao Zhou, Haoyang Li, Yupeng Chen, Yuxin Huang, DexuanXie,Lihua Li , Member,IEEE, Zhaowen Qiu, Member,IEEE, and Xin Gao, Member,IEEE, "A Rapid, Accurate and Machine-Agnostic Segmentation and Quantification Method for CT-Based COVID-19 Diagnosis", in the IEEE Transactions On Medical Imaging, Volume. 39, NO. 8, AUGUST 2020, pp: 2638- 2652