IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 June 2024, accepted 21 July 2024, date of publication 30 July 2024, date of current version 15 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3435774

RESEARCH ARTICLE

End-to-End Fully Automated Lung Cancer Screening System

PUSHKAR SATHE[1], ALKA MAHAJAN[2], (Senior Member, IEEE), DEEPAK PATKAR[3], AND MITUSHA VERMA[3]

[1]Department of Electronics and Telecommunications Engineering, Mukesh Patel School of Technology Management and Engineering, SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deemed-to-University, Mumbai 400056, India
[2]Institute of Engineering and Technology, JK Lakshmipat University, Jaipur 302026, India
[3]Nanavati Superspeciality Hospital, Mumbai 400056, India

Corresponding author: Pushkar Sathe (pushkarmsathe@gmail.com)

ABSTRACT The computer aided diagnosis of lung cancer is majorly focused on detection and segmentation with very less work reported on volume estimation and grading of cancerous nodule. Further, lung cancer segmentation systems are semi automatic in nature requiring radiologists to demarcate cancerous portions on every slice. This leads to subjectivity and delayed diagnosis. Further, these techniques are based on standard convolution leading to inaccurate segmentation in terms of actual boundary retention of the cancerous nodule. Also, there is a need of automatic system that not only grades the lung cancer based on actual parameters but also enables early warning for flagging of anomalies in periodic screening. This research work reports the design of a fully automated end-to-end screening system that consists of 5 major models with an improved performance on cancer detection, segmentation, volume estimation, grading, and an early warning system. The traditional convolutional technique is modified to allow for retention of actual shape of cancerous nodule. The simultaneous segmentation of cancer, lymph nodes and trachea is also achieved through a focus module and a modified loss function to remove redundancy and achieve an accuracy of 92.09%. The volume estimation model is developed using GPR interpolation to give an improved accuracy of 94.18%. A grading model based on the TNM classification standard is developed to grade the detected cancerous nodule to one of the six grades with an accuracy of 96.4%. The grading model is further extended to develop an early warning system for changes in the CT scans of lung cancer patients under treatment. The research is undertaken in collaboration with Nanavati Hospital, Mumbai, and all the models are validated on a real dataset obtained from the hospital.

INDEX TERMS Cancer segmentation, cancer grading, deep learning, lung cancer volume estimation, early warning system.

I. INTRODUCTION

Lung cancer is the most widespread type of cancer and is the second most common cancer after prostate cancer in males and breast cancer in females. The early detection of lung cancer plays an effective role in diagnosis and leads to an early treatment increasing the likelihood of patient survival rate [1], [2]. Chemotherapy, which is the main treatment for lung cancer requires knowledge of accurate location of the cancer along with its spread along all the 3 axes. Accurate volume estimation is required for determining the stage of the cancer. Generally, 2D data is used to estimate the slice wise area of cancer spread and then volume is obtained using these area values. It is very difficult to get accurately annotated data as manual segmentation is usually performed by radiologists. The human intervention leads to errors arising out of fatigue and subjectivity. To overcome this problem, cancer segmentation models based on deep learning were proposed but they require a large dataset along with corresponding annotations of every CT scan for training to get acceptable segmentation accuracy that leads to precise area calculation. Also, majority of cancer segmentation algorithms

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Zuo.

fail to clearly define the boundaries of cancerous nodules. While many software's are available for calculating volume automatically, they are semi-automatic in nature as they need demarcations of cancerous portions to be done by radiologists in every or at most on alternate slices. The segmentation accuracy plays a crucial role in deciding the accuracy of the overall cancer grading system. The grade of the cancer helps oncologists determine the aggressiveness of the cancer and tailor treatment plans accordingly. It is an important factor in predicting the likely outcome of the disease. Monitoring changes in the grade of the cancer over time can help assess how well the treatment is working. If the grade increases despite treatment, it may indicate that the cancer is becoming more aggressive or resistant to therapy. Similarly, performing timely screening of lung CT scans for the same patient is critical for monitoring disease progression, detecting recurrence, assessing treatment response, identifying new abnormalities and providing long-term surveillance. So this work proposes a fully automatic end-to-end lung cancer screening system that includes cancer detection, cancer segmentation, grading and an early warning module.

The paper is organized into several sections for easy reading and understanding. Section II offers a thorough literature review. Section III introduces the proposed methodology, details of image pre-processing, the proposed segmentation and grading algorithm, and the concept of an early warning system along with the developed mathematical models. Section IV includes the implementation details, experimental results of the proposed techniques, and a comparison with recent state-of-the-art methods. Lastly, Section V provides a conclusion and outlines future research directions.

II. LITERATIURE REVIEW

A. ABNORMALITY AND CANCER DETECTION

In recent times CT scans have been dominantly used for cancer detection and a lot of research on developing AI (Artificial Intelligence) based cancer detection algorithms is reported in the literature. Cancer detection using CAD (Computer Aided Diagnosis) is in use since long and its performance in terms of accuracies achieved has got settled to a good extent. So, this section majorly focuses on the work reported about the abnormality detection. Daykin et al. used a One-Class Support Vector Machine (OCSVM) based model to detect lung abnormality [1]. The method, however, does not use complete images but patches of lung CT scans majorly to increase the size of the data set. Automatic feature extraction or the use of deep learning techniques is also not explored fully. Other reported abnormality detection algorithms [2], [3] majorly dealt with cancers other than lung cancer. Irigoien et al. applied OCC (One Class Classification) to medical data for the abnormality detection of various diseases like breast cancer, liver disorders, leukemia, etc. [2]. The authors compared the performance of four different algorithms namely Gaussian, mixtures of Gaussian, Parzen, and typicality approach based on their average AUC (Area Under Curve). The paper reported the best results with an average AUC of 77.4% with the typicality approach. Tarassenko et al. [3] also applied OCC to investigate normality using a large number of available mammograms which do not show any evidence of mass-like structures. The recent advancement in deep learning techniques allows for automatically extracting features from the images, thereby improving the overall performance as compared to conventional CAD (Computer Aided Diagnosis). The use of transfer learning techniques is also reported in literature where fixed feature extraction is done using a pre-trained network. Ardimento et al. [4] applied three transfer learning models viz; VGG (Visual Geometry Group), Xception, and ResNet for feature extraction and combined the results using ensemble architecture to classify the scans as cancerous and noncancerous. Similarly, the state-of-the-art transferable architectures such as VGG-16, VGG-19, GoogLeNet, Inception-V3, ResNet-18, ResNet-50, ResNet-101, InceptionResNet-V2 and 3D multipath VGG like network have also been used for lung cancer feature extraction [5], [6], [7], [8], [9] and the performance of SVM (Support Vector Machine) and AdaBoostM2 classifier is analyzed on the deep features extracted from publicly available datasets. In [10] the performance of various machine learning algorithms was evaluated on a lung cancer detection task. The fusion of DenseNet201 with color histogram techniques was used to extract a hybrid feature set. Similarly, an innovative deep-learning model for lung cancer detection by integrating markers from mRNA, miRNA, and DNA methylation was developed [11]. The principal components analysis (PCA) was implemented to streamline features and the synthetic minority over-sampling technique (SMOTE) algorithm was applied to ensure class balance. The PCA-SMOTE model achieved F1 score of 0.97.

B. CANCER SEGMENTATION

Lung nodule segmentation methods can be categorized into traditional approaches and deep learning-based techniques. Traditional methods encompass threshold and region-growing methods, clustering methods, active contour models, and mathematical models [12], [13], [14], [15], [16], [17], [18], [19], [20]. On the other hand, deep learning-based methods for segmentation can be further divided into 2D and 3D segmentation networks [21]. While traditional segmentation methods do not necessitate a substantial amount of labeled data for model training, they heavily rely on human intervention and are more focused in nature. These methods primarily rely on shallow image features such as grayscale and texture. In contrast, 2D segmentation networks utilize 2D convolution to extract features from images. The original Fully Convolutional Neural Network (FCN) [22], has found extensive use in various image segmentation fields but it lacks global context information which is very important in the case of medical image segmentation. Ronneberger et al. [23] introduced a U-Net network based on the FCN architecture for medical image segmentation. U-Net incorporates both

low-resolution and high-resolution information through skip connections, which is particularly advantageous for segmenting medical images with blurred boundaries. In this approach, low-resolution information is used for target identification, while high-resolution information aids in the localization of the segmentation. Consequently, U-Net has served as a foundation for several improved algorithms. The average IoU of 77.5% ia achieved but the algorithm failed to consider the multiscale information required for accurate segmentation. So, to further improve segmentation results, some researchers have explored the integration of Atrous Spatial Pyramid Pooling (ASPP) [24] instead of the intermediate or output layer of U-Net. This approach enables the extraction of multi-scale image information through different perceptual fields [25], [26], [27] but lacks in retention of exact shapes. Though the average IoU of 81.3 was achieved, the algorithm lacked complete recovery of spatial information. Some 3D segmentation networks utilize 3D convolution to extract features from volumetric images, allowing them to better capture the spatial relationship between nodules and surrounding tissues compared to 2D segmentation networks. Duo et al. [28] introduced a 3D fully convolutional neural network capable of automatically segmenting the liver and cardiac great vessels. Milletari et al. [29] incorporated residual connections inspired by ResNet [30] and employed 3D convolution in their V-Net architecture for prostate volume segmentation, departing from the 2D convolution used in U-Net. Deepseed [31] proposed 3D-squeeze-and-excitation (SE) networks, incorporating dynamic scaling with cross-entropy loss [32] to address the sample imbalance problem. Their method was evaluated on the LIDC [33] and LUNA16 [34] datasets, demonstrating promising results for lung nodule segmentation. However, it is worth noting that 3D networks suffer from longer training times and may not necessarily outperform 2D networks in terms of evaluation metrics. Zhou et al. [35] proposed UNet++, an extension of U-Net that incorporates dense skip connections to enable multi-scale fusion and feature acquisition at different levels. The model achieves an average IoU of 81.4% but as the algorithm involves non-integral convolution, it fails in maintaining the exact shape of nodules. In [36] the attention module is used with the Unet algorithm having residual connections for fast backpropagation. But model failed to maintain the exact shape of cancerous nodules. Similarly, different deep learning segmentation models were evaluated on different datasets, and the effects of different preprocessing methods were examined [37]. Though the TransUet model achieved the highest segmentation accuracy with an average dice coefficient of 0.81, it failed to maintain the exact shape of the cancerous nodule. Being a transformer-based model, the decision-making process and interpreting predictions was challenging. Also, In [38] a method to integrate feature information through a dual-branch network framework and multi-dimensional fusion module is proposed. By training and validating with multiple data sources and different data qualities, the method demonstrated leading performance on the LUNA16, Multi-thickness Slice Image dataset, LIDC, and UniToChest, with an average dice similarity coefficient of 82%. This method failed to maintain the exact shape of the cancer nodule leading to inaccurate segmentation. Accurate segmentation of cancerous nodules is of high importance as the accuracy of volume estimation and grading is directly depending on the same.

C. CANCER GRADING AND EARLY WARNING

No constructive work either on the grading of lung cancer or flagging of anomalies in timely screening is found in the literature. In [39] grade classification of pulmonary nodules is performed using multiscale DenseNet but these grades dealt with the probability of malignancy indicating how severe the disease is and not the actual TNM classification [40] of lung cancer. Some basic image processing techniques and feature extraction methods are used to find histogram, area, perimeter, and centroid of cancerous nodules and grade is predicted based on this information [41], [42]. But all other important parameters required for grading are not covered leading to incomplete analysis. CNN (Convolution Neural Network) is used for feature extraction [43] of cancerous nodule and they are categorized in following into four grades, namely benign, primary lung cancer, malignant, and suspicious malignant.

The primary contributions of our proposed research are as follows:

1) We have proposed a two-stage cancer detection system. The first stage of abnormality detection segregates the scans as normal and abnormal, placing the abnormal scans at top of the stack facilitating faster treatment for other lung diseases. As normal scans are removed from the system in first go, it reduces the burden of radiologists
2) Designed an improved segmentation algorithm based on modified convolution and focus module that performs simultaneous segmentation of three classes namely cancer, trachea and lymph nodes. We achieved an improved segmentation accuracy in terms of well-defined nodule boundaries and the retention of actual nodule shape.
3) We have developed the cancer volume estimation system using the interpolation technique.
4) Proposed the grading system that estimates grade of lung cancer using actual parameters specified in the TNM classification.
5) We have developed an early warning system that monitors the periodic CT scans for changes in parameters and recurrence of cancer. This system is useful for monitoring changes and raising a concern after follow-up periodic scanning.

III. PROPOSED METHODOLOGY

A. IMAGE AUGMENTATION

A significant obstacle in training deep learning models for various tasks is the abundance of data required. This

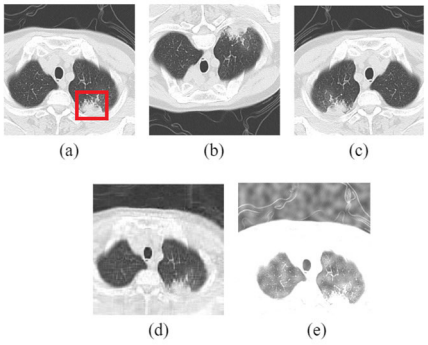


FIGURE 1. (a) Original CT scan (b) Vertical flipping (c) Horizontal flipping (d) PCA with 20 principal components (e) image overlay.

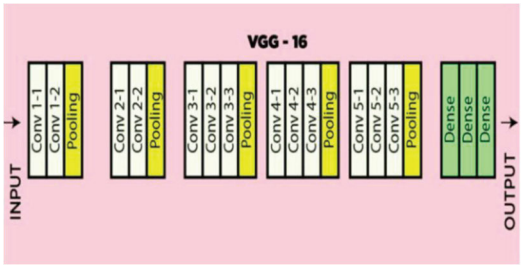


FIGURE 2. Layered architecture of VGG-16.

challenge is particularly pronounced in the field of medicine, where limited access to costly imaging resources or a scarcity of study subjects can hinder progress. Additionally, the requirement of annotations for every cancerous scan puts a limit on data collection. Consequently, there is a growing trend of employing data augmentation techniques in research involving small data sets [44], [45]. The demand for extensive data in deep learning networks has spurred the development of various strategies. Only those strategies are used in this work that provide a different visual impact compared to an original image. Vertical flipping, horizontal flipping, PCA (Principal Component Analysis) and image overlay techniques are used in our work for data augmentation. Figure 1 shows the images generated using corresponding data augmentation methods for a sample CT scan.

Figure 2(a) shows the original CT scan. The irregularly shaped cancerous nodule is highlighted using the red bounding box. Figures 2(b) and 2(c) show the results of the flipping operation performed on the original CT scan. After vertical and horizontal flipping the position of the cancerous nodule changed, making it appear as a new image for a deep-learning model. Figure 2(d) shows the result obtained by performing Principal Component Analysis (PCA) with 20 components. Similarly, figure 2(e) shows the new image generated using the image overlay technique. Image augmentation was performed only on the cancerous scans. A total of 596 cancerous scans (of 97 patients in total) were augmented to 2384. Thereafter, features are extracted using VGG (Visual Geometry Group) and applied to the abnormality detection and cancer detection models as explained below. In the proposed system, only the first 13 convolutional layers out of the 16 layers of VGG were used for feature extraction as the remaining 3 layers do not contribute to feature extraction. They are used to reduce extracted feature size and hence not used. Figure 2 shows the layered architecture of VGG-16.

B. ABNORMALITY DETECTION

The Figure 3(a) shows the block diagram of an existing system of cancer detection. Raw CT scans are applied to the image augmentation and feature extraction model and then extracted features are applied to the cancer detection system. The existing cancer detection model is trained to classify the applied scans as cancerous or normal. So the radiologists need to exclusively assess the segregated normal scans to check for any other abnormality. So we modified the existing system by adding an abnormality detection block to form a proposed cancer detection system as shown in the Figure 3(b). In our case, applied scans are applied to the abnormality detection model which is trained to classify applied scans as normal or abnormal. This successfully removes all normal scans from the further checking. Only abnormal scans are applied to the cancer detection model that classifies them as cancerous or non cancerous but abnormal scans giving radiologist a small set to investigate further for any other abnormality.

In the proposed two stage system, the very first step of prescreening removes all normal scans. Only the identified abnormal scans are applied to the cancer detection model. It segregates the applied scans as cancerous and non-cancerous scans. The rejected lot of this model are flagged off as non-cancerous but abnormal scans, which may belong to any other lung abnormality like Tuberculosis, Covid-19, Pneumonia, etc.

The example given in Figure 4 above indicates that with the proposed system radiologists will need to check just one-fourth of the scans as compared to the current practical scenario where a radiologist manually and sequentially checks scans of approximately a minimum of 50 patients on an average, every day. While the proposed system uses one additional module for abnormality detection, it needs only normal CT scans for training that are easier to obtain as compared to cancerous scans. Also, there is no need of using data augmentation methods as sufficient normal CT scans are available.

The abnormality detection model was trained using only normal lung CT scans. 3250 normal lung CT scans were used to train the model. The model learned the distribution of normality when many instances of normal samples were used to train it. Once the model was trained, test samples were applied to see how much they differed from normality to classify them as abnormal and normal. The abnormality detection model not only separated abnormal scans but also helped establish the priority of processing for radiologists. A comparative study was conducted using two classifiers;

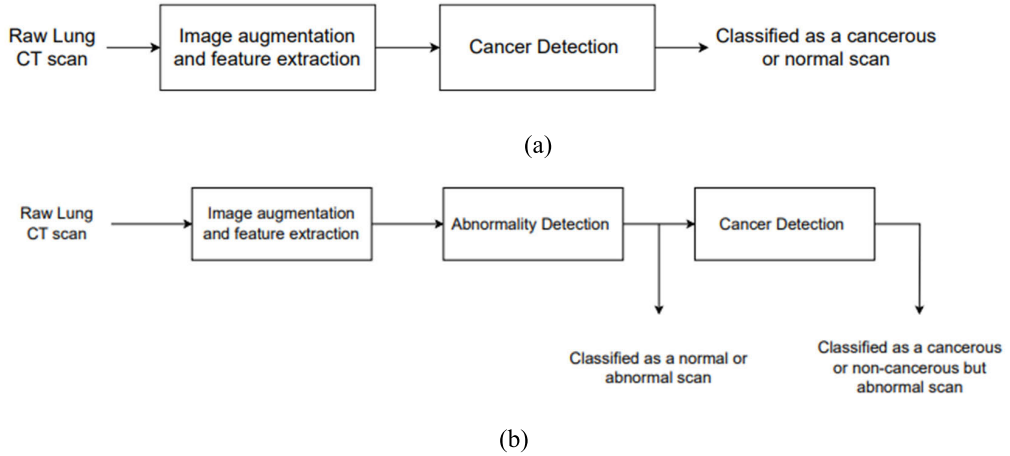


FIGURE 3. Block diagram of (a) The existing cancer detection system (b) The proposed cancer detection system.

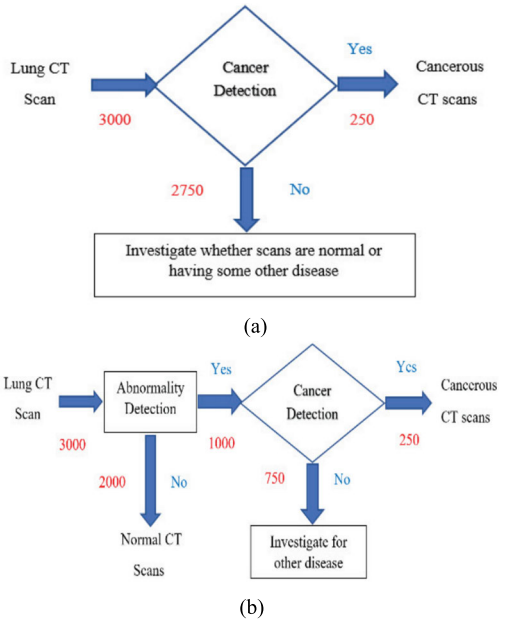


FIGURE 4. (a) Flowchart of the existing system (b) Flowchart of the proposed system.

Isolation Forest and OCSVM (One Class Support Vector Machine).

C. CANCER DETECTION

The CT scans classified as abnormal by the abnormality detection model were applied to the cancer detection model in the second stage. The cancer detection model was trained using both normal and cancerous CT scans. 3250 normal and 2384 cancerous CT scans leading to a total of 5634 scans were used in training. This model was designed to separate cancerous scans from abnormal scans. The extracted features were applied to three different classifiers namely SVM (Support Vector Machine), Decision Tree and Random Forest and their comparative study follows in the sections below.

D. CANCER SEGMENTATION

Currently existing segmentation algorithms are based on standard convolution or its variants. The standard convolution can not extract multi-scale features because of the fixed size of the receptive field [46]. Also, it can not locate the exact location of the feature as the convolution operation takes place at the fixed locations. Because of these reasons, the segmentation of medical images is challenging in terms of boundary detection and retention of the exact shape of irregularly shaped objects like cancer. This paper proposes a modified convolution that overcomes the problems of the standard convolution.

In this work, we have proposed the CLT (Cancer-Lymph-Trchea) multiclass segmentation model that simultaneously segments three classes namely cancerous nodules, lymph nodes and the trachea. Instead of drawing annotations for all 3 classes in the same mask, we generated three different annotation masks for three classes. We know that, depending on the stage of cancer there is a high chance of overlap of cancer on lymph or trachea. So if we mark annotations for all three classes in the same mask, learning of the model becomes difficult because of the overlapping between classes. This leads to the poor segmentation of the classes in terms of perfect shape and boundary extraction. We can not take any risk of imperfect segmentation as the grade of the cancer heavily depends upon the highest dimension of cancer and the reach of cancer to the lymph or trachea. So in the training of images of size (512, 512, 3), we used three masks each of size (512, 512, 1) containing annotation of single class. The simultaneous segmentation of three classes led to reduced redundancy and computation time of the model. Also, the modified convolution, focus module and modified loss function are proposed in the CLT segmentation model to overcome the segmentation constraints of the irregularly shaped objects resulting into the better performance.

The proposed CLT segmentation model is as shown in Figure 5. Various parts of the architecture are explained in the sections to follow.

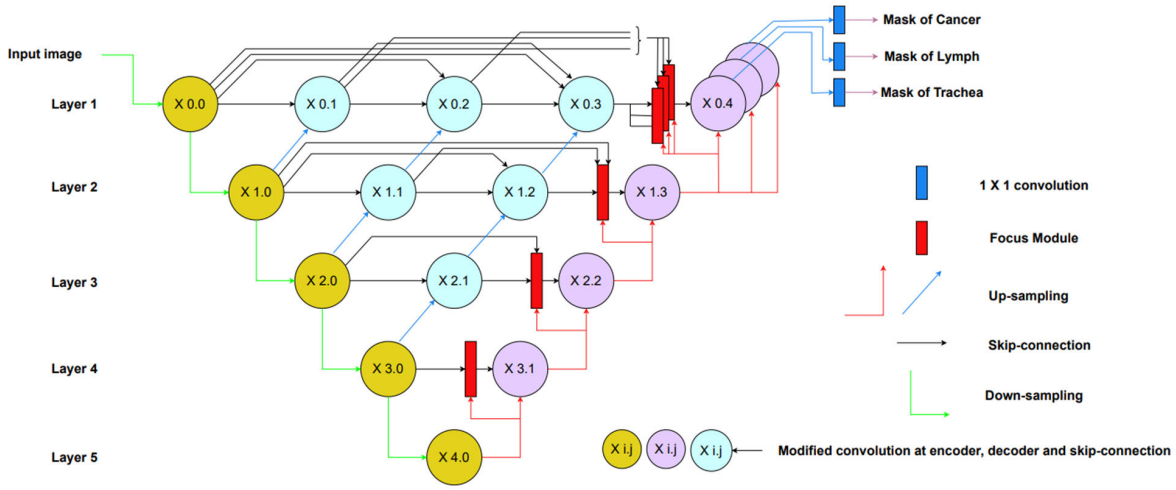
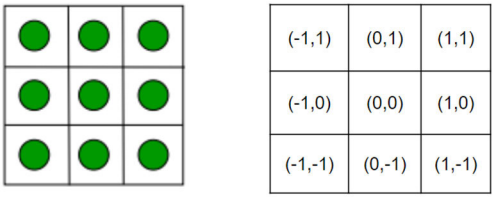


FIGURE 5. The proposed segmentation model.



(-1,1)	(0,1)	(1,1)
(-1,0)	(0,0)	(1,0)
(-1,-1)	(0,-1)	(1,-1)

FIGURE 6. (a) Sample standard grid of size 3 × 3 used in the standard convolution (b) Corresponding fixed integral locations.

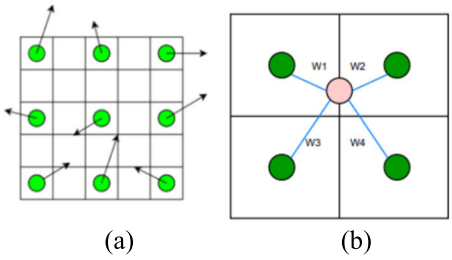


FIGURE 7. Grid of proposed modified convolution.

1) MODIFIED CONVOLUTION

The main problem with the standard convolution is that the receptive field remains constant regardless of the size of the object being analyzed. However, this approach is not ideal when dealing with objects of varying sizes. In the context of medical image target segmentation, the lesions that need to be segmented often have irregular shapes and sizes. Also, it is incapable of capturing exact features spatially. In standard convolution, at any pixel the weighted addition with a particular convolution mask of a fixed size is performed for surrounding pixels to extract the features. The sample grid of size 3 × 3 used in the standard convolution is as shown in Figure 6(a). Corresponding fixed integral locations at which convolution takes place are shown in Figure 6(b). Because of the fixed grid size and fixed integral locations, the standard convolution is incapable of capturing multi-scale features and their exact locations.

Mathematically the standard convolution is represented as

$$y(P_0) = \sum_{Pn \epsilon R} w(P_n).x(P_0 + P_n) \quad (1)$$

Here, as shown in equation (1) pixel value at P_0 is replaced by weighted addition performed between image pixels and convolution mask at locations P_n where

$$P_n \in R \quad (2)$$

$$\text{and } R = \{(-1,-1), (-1,0), ..., (0,1), (1,1)\} \quad (3)$$

as we have considered an example of grid of size 3 × 3. We have proposed a modified convolution that allows convolution grid to change its receptive field to capture local and global features collectively and to perform convolution process at best possible locations within the feature map to capture exact features spatially.

Figure 7(a) shows the grid of the proposed modified convolution with dilation rate of 2 and arrows indicate the non-integral locations at which convolution takes place to capture exact location of features. As the model is capable of extracting edges of the cancerous nodules, it leads to better segmentation in terms of retaining the exact shape of it. It is important to have a higher accuracy of the segmentation of cancerous nodule as the accuracy of volume estimation and grading is depending on it.

The mathematical model of the proposed convolution is given below.

$$y_i(P_0) = \sum_{Pn \epsilon R} w(P_n).x(P_0 + D_i.P_n + \Delta P_n) \quad (4)$$

$$P = \frac{1}{j} \sum_j W_j.P_j \quad (5)$$

$$W_j \alpha \frac{1}{Dj} \quad (6)$$

$$Y(P_0) = \frac{1}{i} \sum_i y_i(P_0) \quad (7)$$

Mathematically the proposed modified convolution can be represented as shown in equation (4) where D_i represents the i^{th} dilation rate and ΔP_n indicates the learnable distance offset for the best location to perform convolution. The image information at non-integral location is required to perform convolution operation. It is obtained using bilinear interpolation operation as shown in Figure 7(b). If as per the model a point marked in a pink color is the most appropriate place to perform convolution and extract feature, then the pixel value is obtained for that point first. It is obtained as a weighted sum of four adjacent pixels as shown in equation (5). The weights are inversely proportional to the distance from the pixel as indicated by equation (6). It allows the convolutional kernel to sample input features from non-integer locations, effectively enabling more flexible and adaptive sampling of the input feature map based on the predicted offsets. Finally, to combine local and global features, features extracted using modified convolution with different dilation rate are combined together as shown in equation (7). The proposed modified convolution allows the model to focus more on the relevant regions of the input feature map by dynamically adjusting receptive field and performing convolution at the most appropriate location.

2) ENCODER

Every encoder layer extracts the contextual information by performing a modified convolution operation on a feature map received from the encoder of the previous layer. It is followed by ReLU and max pooling operations. The mathematical model of the proposed encoder is as given below.

Let $x^{i,j}$ be the output of X i.j convolution block.

$$x^{i,0} = \mathrm{Y}(x^{i-1,0}) \tag{8}$$

$$\mathrm{n}(x^{i,0}) = \max(0, \mathrm{n}\,(x^{i,0})), \text{ for every } \mathrm{n} \epsilon \mathrm{I} \tag{9}$$

$$\mathrm{MP(r, s)} = \max(\mathrm{I[r{:}r + t, s : s + t]}) \tag{10}$$

Here n indicates all number of pixels in a feature map I. The location of a point in a feature map is denoted by (r,s). The equations (8) and (9) indicate modified convolution and ReLU operation respectively. The max pooling operation is shown in equation (10) that halves the height and width of feature map as value of 't' used is 2. The hierarchical nature of the encoder enables the segmentation model to capture fine details from the applied CT scans. These features are further used by the decoder to generate accurate segmentation mask.

3) SKIP PATHWAYS

Skip pathways are used to club the features from different levels together and provide spatial information to the decoder to contribute to better segmentation. All the modified convolutions performed in skip pathways, maintain the dimension of the feature map of that layer intact and only the number of filters gets changed. Mathematically, functions of skip pathway is represented in equation (11).

$$x^{i,j} = \mathrm{Y}([[x^{i,k}]_{k=0}^{j-1}, \quad \mathrm{U}(x^{i+1,j-1})]) \tag{11}$$

where Y indicates the modified convolution operation followed by an activation function, U denotes an up-sampling operation, and [ ] denotes the concatenation. Basically, nodes at level j = 1 receives two inputs, both from the encoder sub-network but at two consecutive levels; and nodes at level j > 1 receive (j + 1) inputs, of which j inputs are the outputs of the previous j nodes in the same skip pathway and the last input is the up-sampled output from the lower skip pathway. So features belonging to different levels are combined first and then modified convolution is performed over it.

4) FOCUS MODULE

It is proposed to enhance the segmentation model's ability to selectively focus on the important features leading to better segmentation accuracy. It receives feature information from the lower layer and spatial information from the skip connection of the same layer. The focus module adds these informations and adjusts the weights accordingly leading to strengthening of the important features.

Further, three separate focus modules are introduced in the uppermost layer. They use class-wise multiplication factors to adjusts the weights so as to have a perfect segmentation of all the three classes. Figure 8 shows the block diagram of the proposed focus module. Also, the mathematical model of the proposed focus module is presented below.

Spatial information coming from multiple convolution blocks of a particular layer is first concatenated as shown in equation (12). Then 1 X 1 convolution is performed on it with stride 2 to match the size with feature information so as to perform addition. Parallely, 1 X 1 convolution with stride 1 is performed on the feature information received from the lower layer (equation 13) and is then multiplied by the class-specific multiplication factors (equation 14). These factors are set to 1 for all focus modules except the uppermost layer (equation 15). Then the addition of spatial features and information features is performed as shown in equation (16). The resulting addition will be higher for a particular pixel only if both the feature maps are strong at that pixel. So, the higher addition is the indication of the important pixel and it should be assigned with higher weight and vice-versa. The addition operation is followed by ReLU activation and 1 X 1 convolution to adjust the weight tensor size (equation 17). The sigmoidal activation is then applied to bring all weight values within the range of 0 to 1. Then, the weights are upsampled to match with the tensor size of spatial information (equation 18). Finally, the multiplication of weights with the spatial information is performed to increase the importance of important features (equation 19).

$$s^{i,j} = \varphi_2([x^{i,k}]_{k=0}^{j-1}) \tag{12}$$

$$t^{i,j} = \varphi_1(x^{i+1,j-1}) \tag{13}$$

$$w^{i,j}(\mathrm{h, k}) = \mathrm{c(h, k)} * t^{i,j}(\mathrm{h, k}) \quad \text{if} \quad \mathrm{j} = 4 \tag{14}$$

$$\mathrm{c(h, k)} = 1; \mathrm{j} < 4 \tag{15}$$

$$m^{i,j}(\mathrm{h, k}) = w^{i,j}(\mathrm{h, k}) + s^{i,j}(\mathrm{h, k}) \tag{16}$$

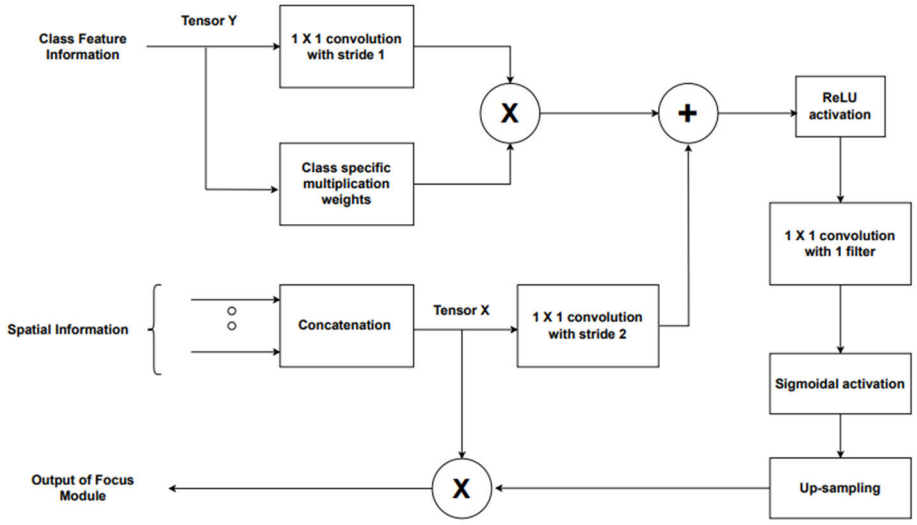


FIGURE 8. The proposed focus module.

$$m^{i,j}(\mathrm{h},\mathrm{k}) = \varphi_1(max(0, m^{i,j}(\mathrm{h},\mathrm{k}))) \tag{17}$$

$$\mathrm{M} = \mathrm{U}(\sigma(m)) \tag{18}$$

$$F^{i,j} = s^{i,j} * M^{i,j} \tag{19}$$

In the uppermost layer, the focus module allows the segmentation model to focus its attention to the class to be segmented. Three separate focus modules are used leading to the better segmentation of all the three classes.

5) DECODER

The decoder plays the crucial role of generating accurate segmentation masks. It upsamples the low resolution feature map obtained from the lower layer decoder and concatenates upsampled feature information with the scaled spatial information received from the focus module. Then it performs modified convolution to refine the feature map that helps model learn more abstract representation and nuance required for a perfect semantic segmentation. Mathematically the role of decoder can be represented as shown in equation (20).

$$x^{i,j} = \mathrm{Y}([F^{i,j},\ \mathrm{U}(x^{i+1,j-1})]) \tag{20}$$

where $F^{i,j}$ is the output of focus module of the same layer.

6) LOSS FUNCTION

The loss function quantifies the discrepancy between the predicted output of the model and the ground true values. The primary goal during the training phase of a deep learning model is to minimize the loss function. By minimizing the loss function, the model learns to make better predictions and improve its performance on the segmentation task.

For a mask of any class, we expect to have the foreground white and the background black. Binary cross-entropy is the most commonly used loss function in the literature. But, it can be problematic in scenarios with a severe class imbalance between foreground and background classes, as in our case where background pixels significantly outnumber foreground pixels. In such cases, the model may focus more on correctly classifying the abundant background pixels since they contribute more to the overall loss calculation. As a result, the loss value can become dominated by the background class, potentially leading to misguidance in the training process as we mainly aim to have a perfect classification of the foreground class.

In this work, we used the loss function as a combination of dice loss and focal loss instead of binary cross-entropy loss. The focal loss is beneficial to address the class imbalance issue. Focal loss is just an extension of the cross-entropy loss function that down-weights easy segmentation portion and assigns more weightage to hard segmentation portion i.e. foreground pixels.

The Focal Loss can be represented as shown in equation (21)

$$\mathrm{FL}(P_i) = \frac{1}{N}\sum_{i=1}^{N} -\alpha(1-P_i)^{\gamma}.\log(P_i) \tag{21}$$

where

P_i is the predicted probability of the true class of i^{th} pixel.

N is total number of pixels.

α is the balancing factor to mitigate class imbalance.

γ is the focusing parameter to down-weight easy examples

The dice Loss is a metric used in segmentation tasks where the actual delineation of the region of interest is essential for decision-making. It is well suited for applications involving irregular structures like cancer.

The equation of dice loss is given below (equation 22)

$$\mathrm{DL} = 1 - \{(2 * |\mathrm{X} \cap \mathrm{Y}|)/(|\mathrm{X}| + |\mathrm{Y}|)\} \tag{22}$$

Here, X and Y represent the sets of pixels that belong to predicted segmentation and the ground truth respectively. So the loss function used for a single class looks like as shown in equation (23).

$$\mathrm{TL} = \mathrm{FL} + \mathrm{DL} \tag{23}$$

The proposed Final Loss Function (FLF) is obtained as a weighted sum of the total losses of all the 3 classes as shown in equation (24). The weights are directly proportional to the total loss values of the classes (equation 25). So class with higher loss gets penalized with higher weight resulting into the further higher weighted loss. As FLF is dominated by the poor performing class, model is forced to pay attention to that class in the next epoch. This will ensure that the segmentation performance of all the classes is balanced.

$$\mathrm{FLF} = \sum_{j=1}^{3} (W_j)\mathrm{x}(TL_j) \tag{24}$$

where W_j and TL_j are the weight value and the total loss associated with the j^{th} class. The weight values are obtained as

$$W_j = \mathrm{K\ x}\ TL_j \tag{25}$$

where K is the controlling parameter that determines how severely we want to reduce the imbalance between the three classes. We have used K=3.

The proposed algorithm provides the segmentation approach that uses not only modified convolution to club features from different resolutions but also an focus module that scales the feature map based on spatial and feature information leading to more accurate segmentation.

E. VOLUME ESTIMATION

When a series of CT scans of the patient is applied to the cancer segmentation model, it provides corresponding masks having cancerous portions highlighted for every CT scan. These annotated masks are applied to the volume estimation model that calculates the area of the cancerous portion of every scan and uses the interpolation method to find the volume of the tumor. From the segmented mask, a number of white pixels are obtained. The area of the cancerous portion is calculated by multiplying the number of white pixels by the area of each pixel which is obtained from pixel height and pixel width functions of the Python library. Likewise, the area is calculated for all consecutive CT scans and then the volume is obtained from these area values using the interpolation method. Figure 9 shows the consecutive cancerous scans for one sample patient case. Corresponding area values are 0.435, 2.966, 4.104, 3.8 and 1.755 respectively.

Each slice is 5 mm thick. This means the real area of cancer spread is known only at intervals of 5 mm each and not in between. Considering the same cancerous area throughout the slice thickness will lead to wrong volume calculation. For better volume estimation, interpolation is used to predict the

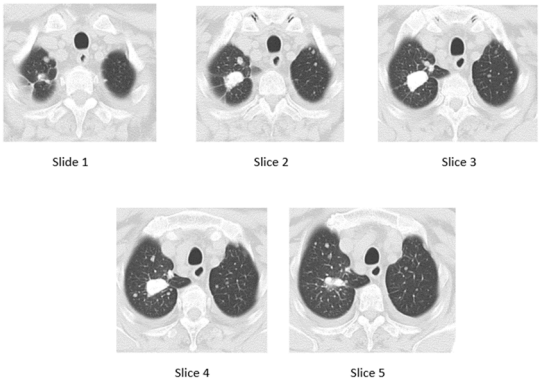


FIGURE 9. Consecutive cancerous scans of a sample patient.

area values of the cancerous portion 0.01 mm apart slice-wise. Various types of interpolations like linear, bilinear, cubic, Lagrange, GPR were tried to determine the volume of cancer, and their results are compared with the actual volumes obtained from radiologists. GPR interpolation is found to be the best approach.

1) GPR INTERPOLATION

The basic idea of GPR is to model the relationship between the input data and the output values as a Gaussian process. A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution. In GPR, we assume that the output values follow a Gaussian process with a mean function $\mu(\mathrm{x})$ and a covariance function k(x, x') that captures the similarity between given data points and interpolation points. The most commonly used covariance function is the Radial Basis Function (RBF) kernel. This interpolation is the most suitable for irregular data as it can better capture underlying trends and patterns. This property makes GPR a most suitable interpolation method in cancer volume estimation. The key idea of GPR interpolation is to use Bayesian inference to compute the posterior distribution over the unknown value $\hat{\mathrm{y}}$ conditioned on the observed data. The posterior distribution is a Gaussian distribution with a mean $\hat{\mu}$ and a covariance $\hat{\Sigma}$. The mean $\hat{\mu}$ represents the estimated value of $\hat{\mathrm{y}}$, and the covariance $\hat{\Sigma}$ quantifies the uncertainty associated with the estimation.

Mathematically GPR interpolation can be derived as follows:

First, the mean function $\mu(\mathrm{x})$ and the covariance function k(x, x') were defined. Then the covariance matrix K(X, X) between the observed input points in X and the covariance vector k(X, $\hat{\mathrm{x}}$) between the observed input points in X and the new input point $\hat{\mathrm{x}}$ were computed. The covariance scalar k($\hat{\mathrm{x}}$, $\hat{\mathrm{x}}$) between the new input point $\hat{\mathrm{x}}$ and itself was then calculated. Finally, the mean vector $\hat{\mu}$ and the covariance matrix $\hat{\Sigma}$ of the posterior distribution were obtained using the formulae [31] given in equations (26) and (27)

$$\hat{\mu} = \mathrm{k(X, \hat{x})^T[K(X, X) + \sigma^2 I]^{-1}.y} \tag{26}$$

$$\hat{\Sigma} = \mathrm{k(\hat{x}, \hat{x}) - k(X, \hat{x})^T[K(X, X) + \sigma^2 I]^{-1}k(X, \hat{x})} \tag{27}$$

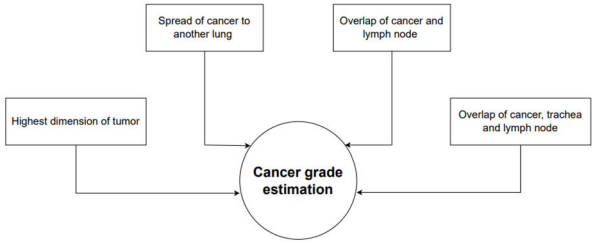


FIGURE 10. Lung cancer grade estimation factors.

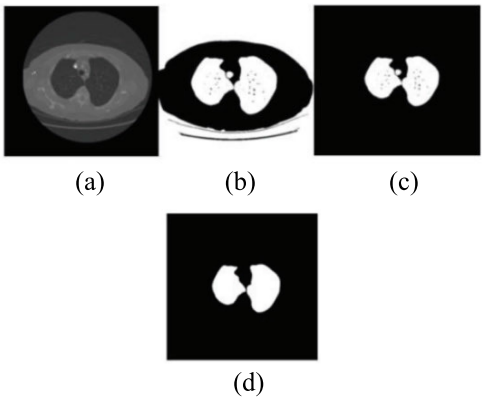


FIGURE 11. (a) Raw DICOM CT scan (b) binarized image (c) boundary cleared image (d) result of erosion and dilation.

where σ^2 is the noise variance parameter and I is the identity matrix.

The estimated output value $\hat{y}$ at the new input point $\hat{x}$ was obtained by the mean $\hat{\mu}$, and the uncertainty associated with the estimation was quantified by the covariance $\hat{\Sigma}$. Various types of interpolations like linear, bilinear, cubic, Lagrange, GPR (Gaussian Process Regression) are checked in this work and their results are compared.

F. CANCER GRADING

Accurate grading is required for the correct treatment of lung cancer in terms of dosage and duration. Figure 10 shows the factors on which grade of the lung cancer majorly depends upon.

1) HIGHEST DIMENSION ESTIMATION

The highest dimension of the tumor is a key factor in cancer staging which helps determine the severity of the cancer. The cancer segmentation model provides the cancer annotations forming a 2D data. The third dimension information is added from the slice width. The highest dimension of the nodule is obtained from the longest distance between the 3D cancerous voxels.

2) CHECK OF CANCER SPREAD TO OTHER LUNG

Though grading mainly focuses on the characteristics of the primary tumor, one factor deciding grade still requires the presence of the secondary nodules to be checked in the other lung. The annotated mask obtained from the cancer segmentation module is checked against the masked lung portion for overlap. Initially, the raw DICOM lung CT scan as shown in Figure 11(a) was converted into Hounsfield units (HU), a standardized quantitative scale for representing radio density. Each tissue type possesses a distinct HU range, which remains consistent across different individuals. Lung tissues typically fall within the range of -400 to -600 HU. Subsequently, the image was binarized using a threshold value of -600, as depicted in Figure 11(b). Boundary clearing was then conducted on Figure 11(b) to isolate the lung area, resulting in Figure 11(c). Boundary clearing includes edge detection and noise reduction using median filtering. To further refine the image, an erosion followed by dilation process was implemented to eliminate pixels on the object boundaries morphologically. The complete lung portion demarcated in white is shown in Figure 11(d). Then overlap is checked between this lung portion and the annotation mask obtained from the cancer segmentation model. The ANDing operation is performed to obtain an overlap image. If it has white demarcation in both lungs, it indicates that cancer has reached another lung as well.

3) CHECK OF CANCER SPREAD UPTO LYMPH NODE

In the lung, several groups of lymph nodes play crucial roles in the immune system and lymphatic drainage. Lymph nodes contain immune cells such as lymphocytes (B cells and T cells) and macrophages, which play crucial roles in recognizing and attacking pathogens (such as bacteria, viruses, and other foreign substances) that enter the lungs. Lymph nodes also act as filters for lymphatic fluid, removing pathogens and cellular debris before returning the fluid to the bloodstream. They help prevent the spread of infection and facilitate the removal of waste products from the lung tissue.

When cancer cells spread to nearby lymph nodes, it signifies a more advanced stage of cancer and can have several harmful implications requiring additional surgery to remove affected lymph nodes or radiation therapy to target cancer cells in the lymph nodes. The risk of metastasis increases when the cancer cells spread up to the lymph nodes as cancer cells can travel through the lymphatic system to reach distant parts of the body. Involvement of lymph nodes by cancer can lead to complications such as lymphedema, impaired lymphatic drainage, and compromised immune function, further impacting the patient's quality of life and overall health. So it is of high importance to find whether cancer cells have reached lymph nodes and especially tracheobronchial lymph nodes to decide on grade and associated further treatment of the cancer. Our segmentation algorithm provides separate annotated masks of cancer, trachea and lymph. So by ANDing method, we found overlap between classes as shown in Figure 12.

4) CHECK OF CANCER SPREAD UPTO TRACHEA

The trachea, also known as the windpipe, is the main airway that carries air to and from the lungs. When lung cancer reaches the trachea, it can physically obstruct the airway,

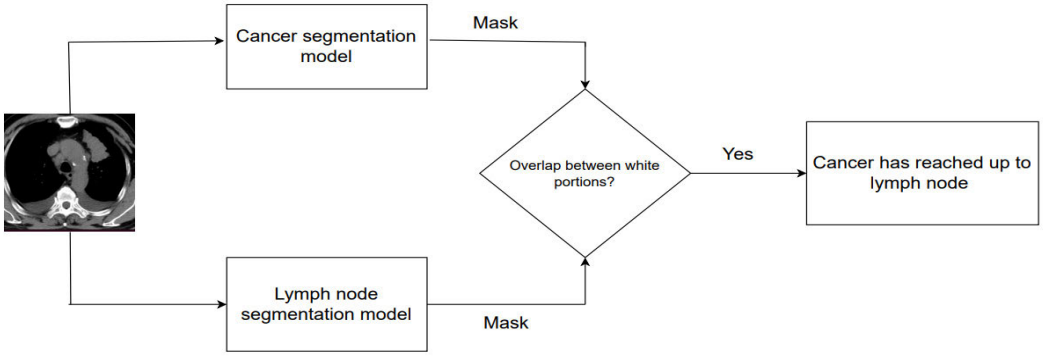


FIGURE 12. Block diagram of lymph node detection.

leading to difficulty breathing, shortness of breath, wheezing, and potentially life-threatening respiratory distress. This obstruction can significantly impair the exchange of oxygen and carbon dioxide, leading to hypoxia (low oxygen levels) and hypercapnia (high carbon dioxide levels) in the blood. Lung cancer reaching the trachea often indicates advanced disease progression and metastasis (spread) beyond the lungs. So, using the same logic as shown in Figure 16, we found overlap between trachea and cancer.

G. EARLY WARNING SYSTEM

In case of lung cancer, there is always a risk of recurrence and so performing timely screening of lung CT scans for the same patient is important. For individuals at high risk of lung cancer, such as current or former smokers, periodic lung CT scans allow healthcare providers to monitor changes in the lungs over time. Comparing scans taken at different intervals can help detect any new or growing abnormalities, indicating possible progression of disease. It helps in early detection of recurrence and an identification of new abnormalities. The CT scans belonging to periodic screening for a sample patient are shown in Figure 13.

We have modified the grading model that incorporates the red flagging in periodic screening. Four red flags are defines as follows

i) RF1 - distance between trachea and cancer is less than 5 cm

ii) RF2 - Cancer spread to lymph and $3 < H < 5$

iii) RF3 - Cancer spread to lymph and $5 < H < 7$

iv) RF4 - $H > 7$

At present, there is no automatic system for monitoring changes and raising a concern after follow-up periodic scanning. As the follow-up may not indicate a change in stage of the cancer, a change in the volume or spread may go undetected. The proposed grading model was further modified to develop an early warning system that monitors the periodic CT scans for changes in parameters and recurrence of cancer. Four relevant warning flags were designed to give clarity on a particular irregularity detected. The Figure 14 shows the Flowchart of early warning system.

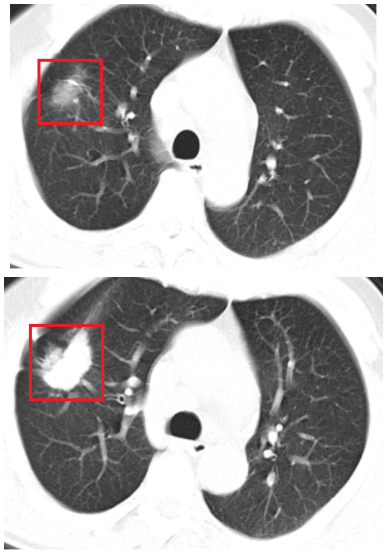


FIGURE 13. CT scans belonging to timely screening for a sample patient.

The Figure 15 shows the schematic of the developed end-to- end system that performs abnormality detection, cancer detection, cancer segmentation, volume estimation, Cancer grading and early warning as shown in the schematic below. (Figure 15)

IV. RESULTS AND DISCUSSION

A. ABNORMALITY AND CANCER DETECTION

The abnormality detection model was validated on a test dataset of 160 normal and 160 abnormal scans. The classification results of Isolation Forest and OCSVM are summarized in the table 1 part (a) and (b). Table 1 part (c) shows the comparison of evaluation metric values obtained for isolation forest and OCSVM classifiers.

As the recall (or sensitivity) obtained is 100%, it is clear that all abnormal scans were correctly classified by OCSVM. Also, the AUC plot for each classifier is shown in Figure 16. It is seen that the ROC curve of OCSVM is closer to the ideal ROC curve.

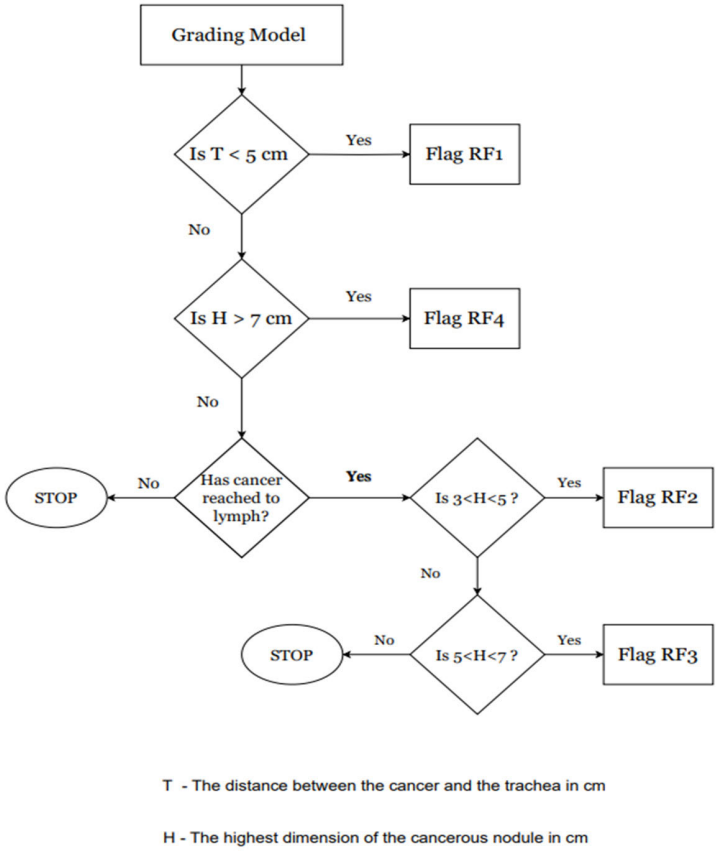


FIGURE 14. Flowchart of early warning system.

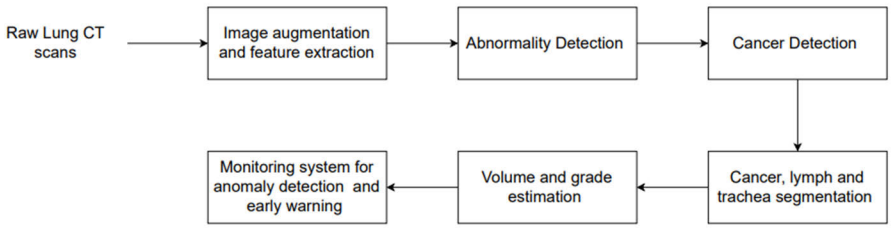


FIGURE 15. The schematic of the proposed system.

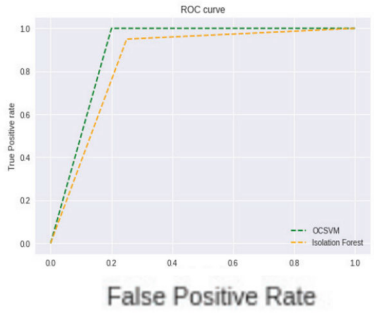


FIGURE 16. AUC plot of abnormality detection classifiers.

The model was validated on a test dataset of 200 normal and 200 cancerous scans. Tables 2(a), (b) and (c) show confusion matrices of SVM, Decision Tree and Random Forest algorithms respectively. It was found that classification results are optimum for random forest classifier when compared to other classifiers. It's confusion matrix shows that all 200 cancerous scans and 199 scans out of 200 non-cancerous scans were correctly classified. Table 2(d) shows a comparison of evaluation metric values obtained for the three classifiers. Again, the evaluation metrics for random forest classifier outperformed the other two. As all the cancerous scans were correctly classified by Random Forest, recall (or sensitivity) obtained is 100%. The AUC plot (Figure 17) for each classifier shows that the random forest classifier gives the best performance.

Based on the comparative study of classifiers for the two models, OCSVM was used for abnormality detection and Random Forest was used for cancer detection.

TABLE 1. Classification results of abnormality detection classifiers (a) Confusion matrix for isolation forest classifier (b) Confusion matrix for OCSVM classifier (c) Comparison of evaluation parameters.

		Predicted class	
		Abnormal	Normal
Actual Class	Abnormal	152	8
	Normal	38	122

(a)

		Predicted class	
		Abnormal	Normal
Actual Class	Abnormal	160	0
	Normal	30	130

(b)

Algorithm / Metric	Accuracy	Precision	Recall	Specificity	F1 - score
Isolation forest	85.62 %	80 %	95 %	76.25 %	86.85 %
OCSVM	**90.62 %**	**84.21 %**	**100 %**	**81.25 %**	**91.42 %**

(c)

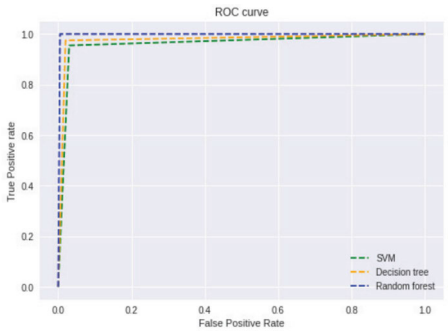


FIGURE 17. AUC plot of cancer detection classifiers.

TABLE 2. Classification results of cancer detection classifiers (a) Confusion matrix for SVM classifier (b) Confusion matrix for decision tree classifier (c) Confusion matrix for random forest classifier (d) Comparison of evaluation parameters.

		Predicted class	
		Cancerous	Non-cancerous
Actual Class	Cancerous	191	9
	Non-cancerous	6	194

(a)

		Predicted class	
		Cancerous	Non-cancerous
Actual Class	Cancerous	195	5
	Non-cancerous	4	196

(b)

		Predicted class	
		Cancerous	Non-cancerous
Actual Class	Cancerous	200	0
	Non-cancerous	1	199

(c)

Algorithm / Metric	Accuracy	Precision	Recall	Specificity	F1 - score
SVM	96.25 %	96.95 %	95.5 %	97 %	96.22 %
Decision tree	97.75 %	97.98 %	97.5 %	98 %	97.73 %
Random forest	**99.75 %**	**99.5 %**	**100 %**	**99.5 %**	**99.74 %**

(d)

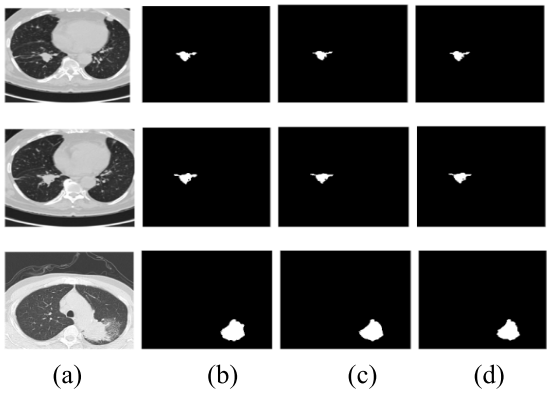


FIGURE 18. (a) Original CT scan (b) Ground truth (c) result of U-net++ (d) result of proposed algorithm.

B. CANCER SEGMENTATION

In this work, the Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Relative Volume Difference (RVD), Average Surface Distance (ASD) and Hausdorff Distance (HD) were used to evaluate the image segmentation results. Results of segmentation of three classes using proposed algorithm are as discussed below. The performance metrics for 3 patients CTP1, CTP2 and CTP3 for cancer segmentation are highlighted in Table 3.

Figure 18 shows the results of segmentation obtained using U-net++ and the proposed algorithm for patients. It is seen clearly that the proposed algorithm retains the shape of a nodule by correctly demarcating its boundaries and provides more accurate segmentation compared to U-net++. The spider chart showing the comparison of U-net++ and the proposed algorithm for various segmentation metrics is shown in Figure 19. The larger difference in the areas covered by the two algorithms in the spider chart indicates the difference in their performance.

The average value of segmentation, an IoU of 0.9035, obtained with the proposed segmentation algorithm is nearly 5% higher than that given by the U-net++ algorithm.

C. VOLUMETRIC ESTIMATION AND GRADING

Table 4 shows the results of the interpolation methods applied to CT scans of 3 patients. VP1, VP2 and VP3 are the volumes estimated for 3 patients P1, P2 and P3 respectively. These volumes are compared with ground truth

TABLE 3. Comparison of various segmentation metrics for Unet++ and the proposed segmentation algorithm.

Patients	Algorithm/parameters	IoU	DSC	RVD	ASD	HD
CTP1	Unet++	0.8751	0.933	0.1426	0.1554	2.8284
	CLT Segmentation model	**0.945**	**0.972**	**0.0582**	**0.0656**	**2.2361**
CTP2	Unet++	0.8754	0.934	0.1423	0.2546	6.7082
	CLT Segmentation model	**0.9287**	**0.963**	**0.0767**	**0.123**	**4.4721**
CTP3	Unet++	0.8698	0.93	0.1496	0.1747	2.8284
	CLT Segmentation model	**0.9368**	**0.967**	**0.0674**	**0.0817**	**2.00**
Average for 58 patients	Unet++	0.8716	0.9152	0.1446	0.1982	2.828
	CLT Segmentation model	**0.9209**	**0.9586**	**0.0682**	**0.0972**	**2.2471**

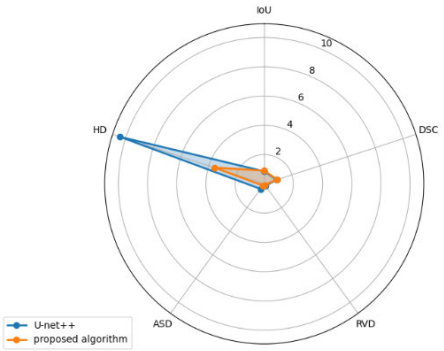


FIGURE 19. The spider chart of comparison of U-net++ with proposed algorithm.

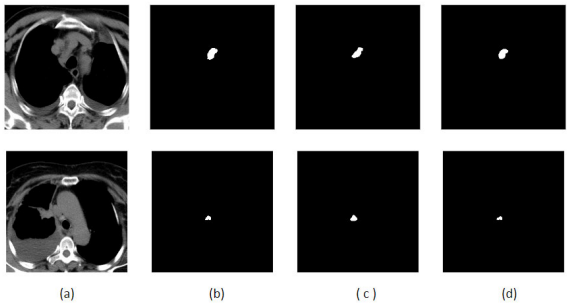


FIGURE 20. (a) Original image (b) ground truth (c) segmentation results of Unet++ (d) Segmentation results of the proposed model.

volumes calculated by the radiologists. It is clearly seen that there is a huge difference in calculated volume and ground truth volume when interpolation is not used. In the case of VP3, volume estimation accuracies with quadratic, cubic, spine, lagrange and GPR interpolations are 88.13%, 88.91%, 90.27%, 91.63% and 94.36% respectively. Similarly, for the other patients it is found that GPR gives the best volume estimation accuracy with an average value of 94.18%.

For comparison purpose, we applied GPR interpolation on the segmentation masks obtained using the U-net++ algorithm. Table 5 indicates volumes obtained using GPR interpolation applied to the Unet++ algorithm and our proposed algorithm for VP1, VP2 and VP3. It is seen that the accuracy of the proposed algorithm is approximately 4% higher than that of U-net++. For all volumes VP1, VP2 and VP3 as well as for average it is seen that the volume estimated using the proposed algorithm is closer to the ground truth compared to that of the U-net++ algorithm.

The table 6 shows the comparison of highest dimension values obtained using standard unet++ and that by the proposed model.

Though cancer grading depends on multiple factors, highest dimension of the tumor is considered to be the most important factor. So if that goes wrong, it may lead to the inaccurate prediction of the stage of the cancer and hence the corresponding dosage and the treatment. A similar case is observed in the case of patient P1 of Table 6. The ground truth value of the highest dimension is 3.27 cm. It belongs to the cancer stage 2A as per the TNM classification[32] as it is in the range of '3 to 5 cm'. With the original U-net++, the obtained dimension value is 2.96. Though it is 90% accurate it falls in the category of 'less than 3 cm' and hence belongs to cancer stage 1. This scenario will lead to the wrong prediction and incorrect treatment. On the other hand, the highest dimension value obtained by the proposed algorithm is 3.13 cm and it correctly belongs to stage 2A.

The results of lymph node segmentation for sample cases are shown in Figure 20 and corresponding IoU values are specified in Table 7.

The results of the trachea segmentation for sample cases are shown in Figure 21 and corresponding IoU values are specified in Table 8. It is clearly seen that performance of the proposed model is much better than standard unet++.

TABLE 4. Volume estimations (in cubic cm) for various interpolation methods.

Nodule volume estimation / Type of interpolation	Ground truth volume	Volume without interpolation	Quadratic	Cubic	Spine	Lagrange	GPR	Volume estimation accuracy in % for GPR
VP1	5.79	6.53	6.37	6.36	6.39	6.35	6.192	93.06
VP2	4.49	4.93	4.13	4.09	4.12	4.18	4.21	93.77
VP3	5.14	5.97	5.75	5.71	5.64	5.57	5.43	94.36
Average for 32 patients	5.17	5.82	5.78	5.81	5.69	5.63	5.47	94.18

TABLE 5. Comparison of volume estimations (in cubic cm) for U-net++ and proposed algorithm.

Nodule volume estimation	Ground truth volume	Volume :GPR interpolation on U-net++	Volume accuracy in % for GPR on Unet++	Volume :GPR interpolation on proposed algorithm	Volume accuracy in % for GPR on proposed algorithm
VP1	5.79	6.422	89.08	6.192	93.06
VP2	4.49	4.039	89.95	4.21	93.77
VP3	5.14	5.62	90.66	5.43	94.36
Average for 32 patients	5.17	5.66	90.51	5.47	94.18

TABLE 6. Comparison of highest dimension estimations (in cm) for U-net++ with standard convolution and proposed algorithm.

Nodule highest dimension estimation	Ground truth value of highest dimension (cm)	U-net++ with standard convolution		Proposed algorithm	
		Value (cm)	% accuracy	Value (cm)	% accuracy
P1	3.27	2.96	90.52	3.13	95.72
P2	2.49	2.08	83.54	2.271	91.17
P3	2.78	2.492	89.65	2.54	91.37
Average for 74 patients	2.821	2.506	88.84	2.63	93.23

Further, the grading model is tested on 56 patients out of which model predicts 54 grades accurately leading to an accuracy of 96.4%.

V. CONCLUSION

In this paper an end-to-end fully automated lung cancer screening system is designed that includes modules on

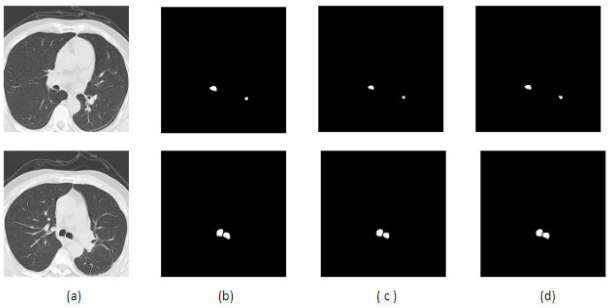


FIGURE 21. (a) Original image (b) ground truth (c) segmentation results of Unet++ (d) Segmentation results of the proposed model.

TABLE 7. Comparison of IoU values for standard Unet++ versus for the proposed model for lymph node segmentation (average for 188 samples).

Parameter	IoU with standard Unet++	IoU with CLT segmentation model
Patient 1	0.683	0.822
Patient 2	0.762	0.883
Average for 188 samples	0.698	0.864

TABLE 8. Comparison of Unet++ results for different types of convolutions for trachea segmentation (average of 54 samples).

Parameter	IoU with standard Unet++	IoU with CLT segmentation model
Patient 1	0.722	0.854
Patient 2	0.903	0.95
Average for 107 samples	0.811	0.916

cancer- detection, segmentation, volume estimation, grading, and a built-in early warning system. The cancer detection module consists of two stages; the first of abnormality detection followed by second stage classifying abnormal scans into cancerous scans and non-cancerous scans indicating other lung disease. This facilitates not only priority attention to the cancerous scans but also ensures further investigation of other lung diseases. An accuracy of 90.5% and 99.7% was achieved for abnormality detection and cancer detection models respectively.

The segmentation model was then designed to capture multiscale information at non-integral locations within the feature map. The standard convolution is incapable of extracting multi-scale features at correct spatial location because of the fixed size of the receptive field leading to inaccurate segmentation in terms of capturing the actual shape of the cancerous nodule. With the proposed modification in the convolution technique, an improved accuracy of segmentation in terms of well-defined nodule boundaries and the retention of actual nodule shape is reported. Further, the proposed model was designed to segment cancer nodules, lymph nodes and trachea simultaneously. This led to reduced computation time and redundancy of the process. A focus module was designed additionally to include learnable class-wise multiplication weights for separating three annotation masks. The loss function was modified to balance the segmentation performance of three classes by penalizing poor-performing classes. The overall accuracy of the proposed segmentation was improved to 92.09% from the reported 87% in the literature.

With a view to improve an accuracy of volume estimation, interpolation was used and the best interpolation technique was identified through a comparative analysis of several techniques. After a thorough analysis, GPR was found to be the best technique, improving the accuracy of the volume estimation to 94.18%.

In the work reported on cancer grading, TNM classification norms are not used leading to grading done at a very superficial level without checking for any of the specified parameters in TNM. The proposed grading system in this research work is designed to detect whether cancer has reached to the second lung, lymph nodes and trachea. It also estimates the highest dimension of the cancerous nodule for determining the grade of the cancer. The proposed grading system graded the cancerous nodule into one of the six grades as per the TNM classification with an accuracy of 96.4%.

Currently, there is no automatic system for monitoring changes and raising a concern after follow-up periodic scan of patients treated for cancer. While the follow-up may not indicate a change in stage of the cancer but a change in the volume or spread may go undetected. The designed grading system was further extended to develop an early warning system that monitors the periodic CT scans for changes in parameters and recurrence of cancer. The module is designed to raise one of the four relevant warning flags to clearly indicate a particular irregularity in the detected stage, helping the doctors initiate a quick informed treatment plan.

REFERENCES

[1] M. Daykin, M. Sellathurai, and I. Poole, "A comparison of unsupervised abnormality detection methods for interstitial lung disease," *Commun. Comput. Inf. Sci.*, vol. 894, pp. 287–298, Jul. 2018, doi: 10.1007/978-3-319-95921-4_27.

[2] I. Irigoien, B. Sierra, and C. Arenas, "Towards application of one-class classification methods to medical data," *Scientific World J.*, vol. 2014, no. 1, 2014, Art. no. 730712, doi: 10.1155/2014/730712.

[3] L. Tarassenko, P. Hayton, N. Cerneaz, and M. Brady, "Novelty detection for the identification of masses in mammograms," in *Proc. 4th Int. Conf. Artif. Neural Netw.*, Jun. 1995, pp. 442–447, doi: 10.1049/cp:19950597.

[4] P. Ardimento, L. Aversano, M. L. Bernardi, and M. Cimitile, "Deep neural networks ensemble for lung nodule detection on chest CT scans," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–8, doi: 10.1109/IJCNN52387.2021.9534176.

[5] M. Muzammil, I. Ali, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Pulmonary nodule classification using feature and ensemble learning-based fusion techniques," *IEEE Access*, vol. 9, pp. 113415–113427, 2021, doi: 10.1109/ACCESS.2021.3102707.

[6] I. Ali, M. Muzammil, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Deep feature selection and decision level fusion for lungs nodule classification," *IEEE Access*, vol. 9, pp. 18962–18973, 2021, doi: 10.1109/ACCESS.2021.3054735.

[7] D. Sethi, K. Arora, and S. Susan, "Transfer learning by deep tuning of pre-trained networks for pulmonary nodule detection," in *Proc. IEEE 15th Int. Conf. Ind. Inf. Syst. (ICIIS)*, Nov. 2020, pp. 168–173, doi: 10.1109/ICIIS51140.2020.9342686.

[8] R. Tekade and K. Rajeswari, "Lung cancer detection and classification using deep learning," in *Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA)*, Aug. 2018, pp. 1–5, doi: 10.1109/ICCUBEA.2018.8697352.

[9] Md. S. Rahman, P. C. Shill, and Z. Homayra, "A new method for lung nodule detection using deep neural networks for CT images," in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2019, pp. 1–6, doi: 10.1109/ECACE.2019.8679439.

[10] N. F. Noaman, B. M. Kanber, A. Al Smadi, L. Jiao, and M. K. Alsmadi, "Advancing oncology diagnostics: AI-enabled early detection of lung cancer through hybrid histological image analysis," *IEEE Access*, vol. 12, pp. 64396–64415, 2024, doi: 10.1109/ACCESS.2024.3397040.

[11] T. I. A. Mohamed and A. E.-S. Ezugwu, "Enhancing lung cancer classification and prediction with deep learning and multi-omics data," *IEEE Access*, vol. 12, pp. 59880–59892, 2024, doi: 10.1109/ACCESS.2024.3394030.

[12] S. K. Fazilov, O. R. Yusupov, and K. S. Abdiyeva, "Mammography image segmentation in breast cancer identification using the OTSU method," *Web Sci.*, vol. 3, no. 8, pp. 196–205, Aug. 2022.

[13] K. V. Shiny and N. Sugitha, "Automatic brain tumor segmentation on pre-operative MRI using region growing algorithm," in *Proc. Inf. Commun. Technol. Competitive Strategies (ICTCS)*. Cham, Switzerland: Springer, 2021, pp. 653–665.

[14] C. J. J. Sheela and G. Suganthi, "Morphological edge detection and brain tumor segmentation in magnetic resonance (MR) images based on region growing and performance evaluation of modified fuzzy C-means (FCM) algorithm," *Multimedia Tools Appl.*, vol. 79, nos. 25–26, pp. 17483–17496, Jul. 2020.

[15] H. Deng, J. P. Fitts, and C. A. Peters, "Quantifying fracture geometry with X-ray tomography: Technique of iterative local thresholding (TILT) for 3D image segmentation," *Comput. Geosci.*, vol. 20, no. 1, pp. 231–244, Feb. 2016.

[16] Y. Akbari, H. Hassen, S. Al-Maadeed, and S. M. Zughaier, "COVID-19 lesion segmentation using lung CT scan images: Comparative study based on active contour models," *Appl. Sci.*, vol. 11, no. 17, p. 8039, Aug. 2021.

[17] P. M. Bruntha, D. J. Rose, A. T. Shruthi, K. G. Juliet, and M. Kanimozhi, "Application of selective region growing algorithm in lung nodule segmentation," in *Proc. 4th Int. Conf. Devices, Circuits Syst. (ICDCS)*, Mar. 2018, pp. 319–322.

[18] P. B. Sangamithraa and S. Govindaraju, "Lung tumour detection and classification using EK-mean clustering," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2016, pp. 2201–2206.

[19] M. Keshani, Z. Azimifar, F. Tajeripour, and R. Boostani, "Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system," *Comput. Biol. Med.*, vol. 43, no. 4, pp. 287–300, May 2013.

[20] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke, "Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," *IEEE Trans. Med. Imag.*, vol. 22, no. 10, pp. 1259–1274, Oct. 2003.

[21] H. Xie, D. Yang, N. Sun, Z. Chen, and Y. Zhang, "Automated pulmonary nodule detection in CT images using deep convolutional neural networks," *Pattern Recognit.*, vol. 85, pp. 109–119, Jan. 2019.

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[24] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*.

[25] X. Luo, T. Song, G. Wang, J. Chen, Y. Chen, K. Li, D. N. Metaxas, and S. Zhang, "SCPM-Net: An anchor-free 3D lung nodule detection network using sphere representation and center points matching," *Med. Image Anal.*, vol. 75, Jan. 2022, Art. no. 102287.

[26] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "CE-Net: Context encoder network for 2D medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2281–2292, Oct. 2019.

[27] S. Bose, R. S. Chowdhury, R. Das, and U. Maulik, "Dense dilated deep multiscale supervised U-network for biomedical image segmentation," *Comput. Biol. Med.*, vol. 143, Apr. 2022, Art. no. 105274.

[28] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, and P.-A. Heng, "3D deeply supervised network for automated segmentation of volumetric medical images," *Med. Image Anal.*, vol. 41, pp. 40–54, Oct. 2017.

[29] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[31] Y. Li and Y. Fan, "DeepSEED: 3D squeeze-and-excitation encoder–decoder convolutional neural networks for pulmonary nodule detection," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1866–1869.

[32] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4685–4694.

[33] S. G. Armato et al., "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, Feb. 2011.

[34] A. A. A. Setio et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," *Med. Image Anal.*, vol. 42, pp. 1–13, Dec. 2017.

[35] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham, Switzerland: Springer, 2018, pp. 3–11.

[36] J. Hou, C. Yan, R. Li, Q. Huang, X. Fan, and F. Lin, "Lung nodule segmentation algorithm with SMR-UNet," *IEEE Access*, vol. 11, pp. 34319–34331, 2023.

[37] W. Chen, Y. Wang, D. Tian, and Y. Yao, "CT lung nodule segmentation: A comparative study of data preprocessing and deep learning models," *IEEE Access*, vol. 11, pp. 34925–34931, 2023, doi: 10.1109/ACCESS.2023.3265170.

[38] W. Jiang, L. Zhi, S. Zhang, and T. Zhou, "A dual-branch framework with prior knowledge for precise segmentation of lung nodules in challenging CT scans," *IEEE J. Biomed. Health Informat.*, vol. 28, no. 3, pp. 1540–1551, Mar. 2024, doi: 10.1109/JBHI.2024.3355008.

[39] W. Wang, G. Chakraborty, and B. Chakraborty, "3D multi-scale DenseNet for malignancy grade classification of pulmonary nodules," in *Proc. 11th Int. Conf. Awareness Sci. Technol. (iCAST)*, Dec. 2020.

[40] (2018). *Radiology Assistant*. [Online]. Available: https://radiology assistant.nl/chest/lung-cancer/tnm-classification-8th-edition

[41] S. Shafiq, M. A. Asghar, M. E. Amjad, and J. Ibrahim, "An effective early stage detection of lung cancer using fuzzy local information cMean and GoogLeNet," in *Proc. 16th Int. Conf. Open Source Syst. Technol. (ICOSST)*, Dec. 2022.

[42] M. F. Abdullah, S. N. Sulaiman, M. K. Osman, N. K. A. Karim, I. L. Shuaib, and M. D. I. Alhamdu, "Classification of lung cancer stages from CT scan images using image processing and k-nearest neighbours," in *Proc. 11th IEEE Control Syst. Graduate Res. Colloq. (ICSGRC)*, Aug. 2020, pp. 68–72.

[43] D. Gupta and S. Dawn, "Detection and staging of lung cancer from CT scan images by deep learning," in *Proc. Int. Conf. Disruptive Technol. (ICDT)*, May 2023, pp. 274–278.

[44] F. Liao, M. Liang, Z. Li, X. Hu, and S. Song, "Evaluate the malignancy of pulmonary nodules using the 3-D deep leaky noisy-OR network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3484–3495, Nov. 2019, doi: 10.1109/TNNLS.2019.2892409.

[45] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[46] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, "Deformable convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 764–773, doi: 10.1109/ICCV.2017.89.

PUSHKAR SATHE received the bachelor's and master's degrees from Mumbai University. He is currently pursuing the Ph.D. degree in electronics and telecommunication engineering with the Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai. His research interests include image processing, neural networks, and embedded systems.

ALKA MAHAJAN (Senior Member, IEEE) received the degree in Electronics and Electrical Engineering from Andhra University College of Engineering, Waltair, in 1983, the M.Tech. degree in electronics design technology from CEDTI, Aurangabad, in 1994, and the Ph.D. degree in electronics and communication engineering in the broad area of soft computing from Delhi College of Engineering, Delhi University, in 2007. She is currently a Professor and the Pro Vice-Chancellor with JK Lakshmipat University, Jaipur. Prior to joining the JK Lakshmipat University, in May 2024, she was a Professor and the Dean with the Mukesh Patel School of Technology Management and Engineering, NMIMS University, Mumbai, India. Her research interests include digital signal processing, AI and machine learning, soft computing, and power electronics.

DEEPAK PATKAR received the D.M.R.D., D.M.R.E., and Doctor of Medicine (M.D.) degrees in radiology. He is currently the Director of medical services and the Head of the Department of Imaging, Nanavati Max Super Speciality Hospital, Mumbai. He is an experienced Radiologist with a demonstrated history of working in the outsourcing/offshoring industry. He is skilled in MRI, clinical research, medical education, medical devices, and medicine. He was a Fellow of ICR.

MITUSHA VERMA received the bachelor's degree from the Government Medical College, Nagpur, and the master's degree in radio diagnosis from the National Board of Examination, Ministry of Health and Family Welfare. She is currently a Consultant Radiologist of MRI, PET-CT, and CT with the Department of Imaging, Nanavati Superspeciality Hospital, Mumbai. She has a keen interest in teaching and research activities.