INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

A Novel 3D Multi-Layer Convolutional Neural Networks for Lung Cancer Segmentation in CT Images

P L Chithra[1]*, **P Bhavani**[2]

1 Professor, Department of Computer Science, University of Madras, Chennai, 600025, Tamil Nadu, India
2 Research Scholar, Department of Computer Science, University of Madras, Chennai, 600025, Tamil Nadu, India

OPEN ACCESS

Received: 16-08-2023

Accepted: 27-02-2024

Published: 28-03-2024

Citation: Chithra PL, Bhavani P (2024) A Novel 3D Multi-Layer Convolutional Neural Networks for Lung Cancer Segmentation in CT Images. Indian Journal of Science and Technology 17(13): 1368-1380. https://doi.org/10.17485/IJST/v17i13.2081

* **Corresponding author**.

chitrasp2001@yahoo.com

Funding: University of Madras

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Background/Objectives: A novel three-dimensional efficient Multi-Layer Convolutional Neural Network (3D-MLCNN) is proposed for detecting lung tumors accurately using Computerized Tomography (CT) lung tumor images. The proposed K-means segmentation algorithm for labeling the tumor region automatically. This proposed K-means segmentation algorithm automatically labels the tumor regions to process with the 3D MLCNN model to predict tiny tumors and extract tumor regions accurately. **Methods**: The proposed 3DMLCNN network goal is to extract the tumor region in CT lung images to classify the lung tumor volume by pixel-wise segmentation model. **Findings**: The proposed 3D MLCNN segmentation model for detecting the segmenting tumors produces outperforming results for predicting even tiny tumors in the lung images. Experimental results demonstrated with lung cancer CT images in TCIA datasets show that the proposed model 3D-MLCNN achieved a dice coefficient (9.6%), Intersection over Union (IoU) (80%), F1-Score (9.33%), Sensitivity (17.11%), and Accuracy (98%) respectively. However, the proposed model 3D MLCNN was evaluated and compared with the existing state-of-the-art segmentation methods, which shows a 10% improvement in the segmentation process. **Novelty**: A novel 3D MLCNN model enhances the tumor region and predicts the tumor accurately by labeling the tumor using K-means labeling techniques.

Keywords: 3D Convolutional Neural Networks; Lung Cancer CT Images; K - Means Labeling; Feature Visualization; Deep Learning

1 Introduction

Automatic lung cancer detection, diagnosis, and treatment planning are the various processes that improve the patient's life span. In early detection stages, lung nodules measure from 3mm to 30mm. Furthermore, lung tumors can be either benign or malignant. A benign lung tumor is generally considered non-cancerous, while malignant tumors are cancerous. Lung tumors smaller than 14mm are benign 95% of the time, while 84% of the time, nodules more significant than 16mm are malignant [(1)].

The Convolutional Neural Networks (CNN) are based on traditional approaches for accurate and automated tumor segmentation from Computerized Tomography (CT) images, which are crucial for clinical assessments such as diagnosis of tumor detection, monitoring the tissues, and treatment in a 3D view, which generates better segmentation by learning the deep networks[2]. For automatic lung tumor delineation utilizing CT scans, 2D CNN and 3D CNN were used. 3D CNN extracted tumor context information from CT sequence pictures using the V-Net model. 2D CNN employed an encoder-decoder structure based on a dense connection technique to increase information flow and feature propagation. Then, using a hybrid modulation, 2D and 3D features are combined[3].

3D segmenting the tumor regions from CT is challenging because the nodule type, shape, size, and intensity vary across images. The task has been processed in a multi-view 3D perspective. CNN structure has reduced false positives that extract 3D features. These automated CNNs use deep learning techniques with image patches for segmenting the model in many ways, such as fuzzy Markov random field segmentation, computing the Gaussian distribution of pixels present in the image, and K-Means[4,5]. A malignant tumor is more problematic as it can spread to other organs. Therefore, early-stage detection can lead to better treatment for the patient and avert the spread of cancer at the right time, leading to an increase in the survival rate[6]. Doctors can benefit significantly from a wealth of imaging data to interpret lesions[7] accurately. Multi-parameter Magnetic Resonance Imaging (Mp-MRI), which consists of diffusion-weighted imaging (DWI), KTrans imaging, and dynamic contrast-enhanced imaging (DCE). MRI is frequently employed in segmentation and identification because it can record lesion's structural and pathological information[8]. To enhance lung tumor segmentation in a 3D segmentation network that can more effectively use spatial and volumetric information while requiring little to no computational modification is being developed[9].

The standard conventional method for identifying lung nodules is through computer tomography (CT). In contrast, accurate differentiation of benign from malignant nodules requires patient follow-up, biopsy, or FDG uptake calculated by PET/CT systems. Furthermore, Positron Emission Tomography (PET) imaging accurately defines the nodule suspiciousness rating[10].

Distraction-Sensitive U-Net (DSU-Net) was a cascaded two-stage U-net model that explicitly incorporates information from uncertain regions (distraction regions), containing local segmentation data for volume patches to distinguish tumor regions further. To achieve this, a Distraction Attention Module (DAM) is suggested and implemented at each level of the U-Net in stage II, improving the ability to distinguish between features[11]. Segment the lung tumor using a variational technique on PET-CT images by first applying a 3D Fully Convolutional Network (FCN) on the CT image to build a probability map, then presenting a fuzzy variational model to include the probability map and the PET intensity picture for accurate segmentation[12].

PSP Dense U-Net, proposed by Mohammed Yusuf Ansari et al., segments the tumor region without comparing existing architecture and increases the segmentation accuracy[13]. Deep learning segmentation techniques can automatically learn relevant feature hierarchies from the trained data in an end-to-end process[14]. According to Bin Li et al.[15] and Peng Li et al.[16], the U-Net comprises a decoder that does information fusion and an encoder that performs feature extraction. To utilize the information at various levels of resolution and prevent feature loss, U-Net uses the skip connection for feature fusion[17,18]. Tumor segmentation can be considered a pixel-by-pixel binary classification problem for the U-Net-based deep learning algorithms[19], where each pixel is labeled as a tumor or non-tumor pixel[20].

The Ioan-Daniel et al. proposed clustering algorithms of Fuzzy C-Means (FCM) and K-Means (KM) are used for segmentation to set the seed point selection in the clustering process; they suffer initial point setting and handling big datasets. For this reason, linearly, it increases proportionally the data size and performs the cluster operations[21]. An improved encoder and decoder CNN architecture to extract the hierarchy of tumor features from the brain imageswasproposed by Chithra et al.[22]. Then, these average values of attributes are deconvolved, getting the original image size. Finally, the extracted outcome can be classified in softmax classification to segment the tumor region, which was implemented by Chithra et al.[23]. Ansari, M.Y., Yang, et al. proposed the backbone of Res-PAC-UNet architecture is used to provide more effective segmentation accuracy. The weight of the model size reduces the disc consumption[24].

The improvement module has produced positive outcomes when used in the Random Walk segmentation procedure. Compared to Gaussian or DoG, DroLoG has been developed as a superior contender for the weighting function role in Random Walk, implemented by Sarada Prasad Dakua[25]. Dakua and Sarada introduced a deterministic and time-efficient to provide a better solution of a 3D semi-automatic graph-based segmentation algorithm based on chaos theory to recreate objects using volumetric data from segmented slices[26].

In order to highlight the edges in the image where various objects are just marginally different, the algorithm uses anisotropic diffusion filtering. The LV is next segmented using the AnnularCut approach[27]. Clinicians with extensive training have evaluated the volumetric left ventricle contour acquired using this method, which was developed by Dakua and Sarada, as highly excellent.

3D U-net-based encoding and decoding architectures have been used for pixel-wise prediction of classes, which have been explored in the CNN model, but they need more computations. This model also handles two pathways of the feature map[28].

Norman and Berk et al.[29] implemented the 2D U-Net model lung tumor segmentation produces one tumor slice from CT scans, typically dispersed among continuous CT slices. Qi Dou et al.[30] proposed volumetric medical image segmentation is an essential but complex issue in medical image analysis. Several 3D volume-to-volume segmentation networks have been proposed, such as 3D CNN and 3D U-Net[31].

To overcome these drawbacks, an automated 3D Multi-Layer Convolutional Neural Network architecture is trained with 3 x 3 and 1 x 1 to produce separate feature maps has been proposed for 3DMLCNN lung cancer segmentation controls the feature map in one way, labels the tumor region in the entire Nifty dataset one by one, and sets the initial point to perform the K-Means labeling algorithm, which improves the left, root, and right sub-tree with the help of the encoding and decoding process in 3DMLCNN architecture and predicts the accurate segmentation regions from lung cancer images.

Challenges of the Existing Methods:

- In the existing literature, mapping a 3D data set to process with a 3D convolutional neural network is challenging.
- Automatically targeting lung tumor shapes is challenging due to the high variance in lung tumor segmentation.
- The disparity between classes can cause overfitting, resulting in poor performance.
- An extensive set of data can be labeled for training, which makes it challenging to train and test.
- Inadequate performance metrics outcomes while mapping 3D input Shape.

Significance of the Proposed 3D-MLCNN Method:

The proposed strategy goal is to extract whole-segment tumors, particularly in the tiny tumor regions. The main contribution is as follows.

- This method can predict the slice-by-slice tumor regions by segmenting faster than the patch-based segmentation model.
- The proposed 3D model is used to segment lung tumors in K-means segmentation in lung CT images to save K-means segment (lung mask) images and process with the proposed 3D MLCNN network.
- To train the lung images and lung label (mask) images, a pixel-to-pixel or end-to-end 3DMLCNN model for pixel or slice-wise prediction of whole tumors and specifically for hidden tumors has been proposed.
- In this proposed method, increasing data capacity during network training in encoding and decoding node block features are tracked and visualized in all intermediate layers. It has a high outcome of segmenting regions compared with the existing encoding and decoding model.
- This proposed model has been evaluated and validated using datasets from The Cancer Imaging Archive (TCIA), showing that it achieves a dice score (9.6%), Intersection over Union (IoU) (80%), F1-Score (9.33%), Sensitivity (17.11%), and Accuracy (98%) respectively. These values have been compared with the existing state-of-art segmentation models.

2 Methodology

The proposed 3DMLCNN network process entails five key steps, including pre-processing of lung CT images along with feature extraction using the left sub-tree node, reconstructed images of feature extraction using the right sub-tree node, using softmax classification and calculating Dice coefficient, Intersection over Union (IoU) and loss functions to perform training, validating, testing and visualization could be evaluated. Each node input convolved separately with kernel size, 3 X 3 and 1 X 1 filters, and softmax classifier features based on proposed training and validating the 3D lung tumor images to detect the tumor accurately. The proposed workflow process is presented in Figure 1and portrayed in the following subsections.

2.1 Pre-processing of Lung CT images in K-Means

3D K-Means segmentation algorithm and image augmentation are applied to the lung cancer images to correct the input image's tumor noise and intensity value. Furthermore, an efficient way to rotate the angle of the tumor shape removes noise, which helps to trace the tumor to calculate the distance of tumors and extract the tissue shapes as mentioned in the Algorithm (1), and the resultant tissue is segmented in k means threshold as shown in Figure 2.

Algorithm 1. 3D K- Means Segmentation labeling (Mask)

Input: 3D lung cancer CT images

Output: 3D K-Means lung cancer labeling (mask)

1: Read the tissue shapes and tissue position (Ts, Tp)

2: Normalize the CT input data images (Pivot point (Pt), max, min)

Normalize_ Pt = (Pt - Pt.min()) / (Pt.max() - Pt.min())

3: Extract the tissue shapes fixed with Seed point(Sp)

4: Ts = k-means (Sp(Pt.max()))
5: Ts, Tp = sorted (tissues)
return (Pt -Tp)< center (Pt – Ts)
6: Reshaped the K-means Pt values and generated the slices
7: out_seg = k-means(in_seg)
8: **for** i, slice in k-means (in_seg):
Slice=label(slice)
for Pivote_index in range(1, slice_label.max () +1):
mask = slice_label == Pivote_index)
max_dist=center (mask).max()
out_seg[i] = mask * mask_dist
end for
end for
9: Display the segmented mask images
10: Segmented mask tissue saved as .nii.gz

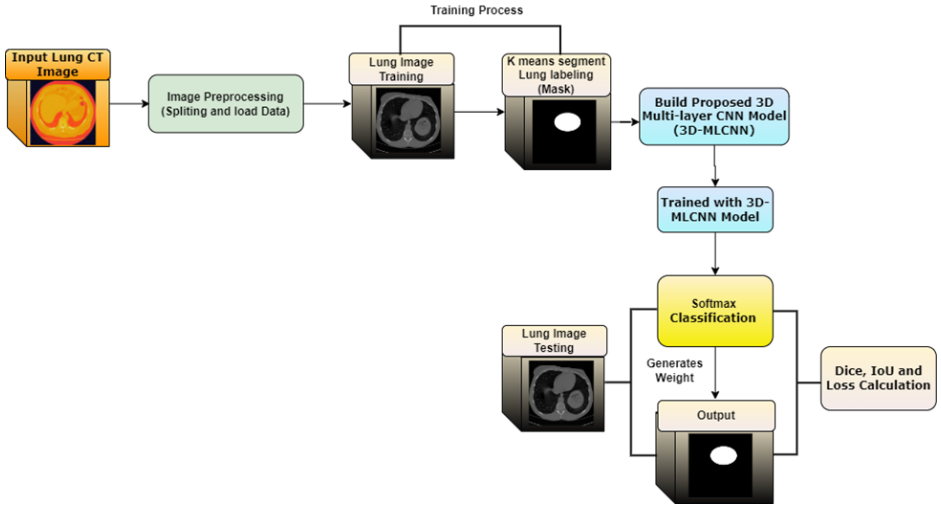

Fig 1. The workflow of the proposed 3DMLCNN

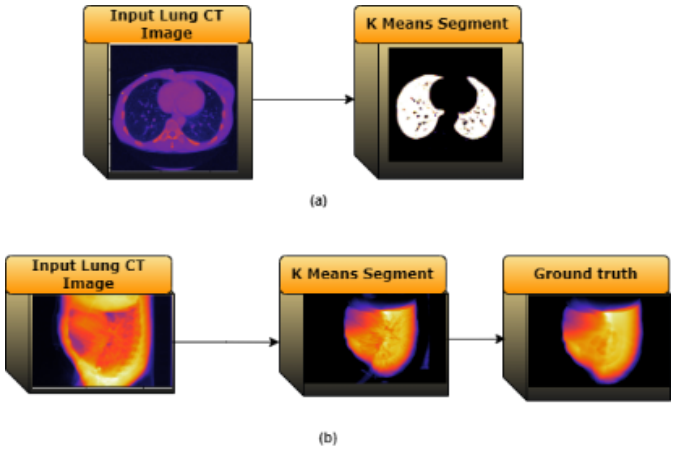

Fig 2. K- Means segmentation labeling (mask) (a) input lung CT image to produce K-means mask (b) input image to produce K-means mask with ground truth points

2.2 Multi-Layer Convolutional Neural Network

The proposed 3DMulti-Layer Convolutional Neural Network (3DMLCNN) is a standard architecture, and it processes the encoding and decoding techniques through the left sub-tree Node and right sub-tree node; these two nodes are connected between the rooted sub-trees at the bottom of the node to process the convolution node (block), end to end manner and produce the segmentation results effectively.

2.2.1 Left Sub-tree Node

The left sub-tree node is called the encoding layer, which extracts the feature region in the ladder form from lung cancer images. In the first node, the block is convolved separately in each input image with two different kernels in the 3D format of 3 x 3 and 2 x 2 to produce the 3D efficient feature maps. D_{y1} and D_{y2} are the two feature map outputs that have been combined by creating input D_{xa} with weight D_{wa} through the intervals(a=1, b=1), and bias is defined in Equations (1) and (2).

$$D_{y1} = f\left(\sum_{a=1}^{n}\left(D_{xa} * D_{wa}\right) + D\right) \tag{1}$$

$$D_{y2} = f\left(\sum_{b=1}^{n}\left(D_{xb} * D_{wb}\right) + D\right) \tag{2}$$

Here, f is one of the activation functions for the convolution layer of non-linear transformation. The non-linear transformation consists of a sigmoid, hyperbolic target, and ReLU (Rectifier Linear Unit) functions. ReLu activations produce outstanding results compared to other activation functions in neural networks [23]. The activation function of (a,b) ReLU has been calculated in the given Equation (3).

$$f\ (a,b)\ =\ max(0,(a,b)) \tag{3}$$

$$f\ (a,b) = \begin{cases} 0\ if(a,b) < 0 \\ x\ otherwise \end{cases}$$

The non-linear activation function is down-sampled using max-pooling from two outputs of the feature map D_{y1} and $D_{y2,}$ generating the feature maps D_{z1} and $D_{z2,}$ respectively. Max pooling is used to incorporate the nearby features and reduce the computational weight in the network in the left sub-tree node; then, the Dyout is calculated using the max-pooling outputs of D_{z1} D_{z2} as shown in Equation (4).

$$D_{yout}\ =\ Mean\ (D_{z1},\ D_{z2}) \tag{4}$$

The feature map's output from the left sub-tree first node is processed into the upsampled nodes in a left sub-tree node of the encoding node, and their corresponding feature map value is calculated.

2.2.2 Root Sub-tree Node

The root sub–tree node is the bottom of the node in 3DMLCNN, which is used to store, monitor, and control the encoding and decoding process of feature map outputs combined in D_{y1} and D_{y2} denotes as Dx, the computational value calculated by the left sub-tree of D_{yout} indicated as Dy and in the given Equation (5).

$$D\ =\ Dx\ +\ Dy \tag{5}$$

2.2.3 Right sub–tree Node

The right sub-tree node is also called the decoding layer or expanding layer, which is used to develop the feature map dimension to get the size of the original image. This node contains three decoder nodes in the deconvolutional layer, where each node includes convolutional upsampling and Batch Normalization (BN). The decoder node upsampled the encoder node endpoint and performed the convolution to generate the feature map. This outcome is transmitted with the max-pooling feature map from the second convolution node link into an upcoming second decoder node to create an expanded feature. These features are sent from the first node of convolutional to produce the feature map with the original image size.

2.2.4 Softmax Classification

The final extracted features are classified and allocated to the class label using a softmax classifier for segmentation purposes. The softmax classifier is used to classify the tumor and non tumor lung images. Here, to predict the lung tumor images to calculate the loss value by comparing the ground truth label values, these loss values are simulated, and network configurations are controlled in the left and right sub-tree nodes in the order of encoding and decoding layer for getting accurate results. The ultimate goal of this proposed network is to detect the segment class labels accurately through the minimized loss function. The proposed 3DMLCNN architecture is elaborately mentioned in Figure 3, and its factorization (model) is shown in Table 1.

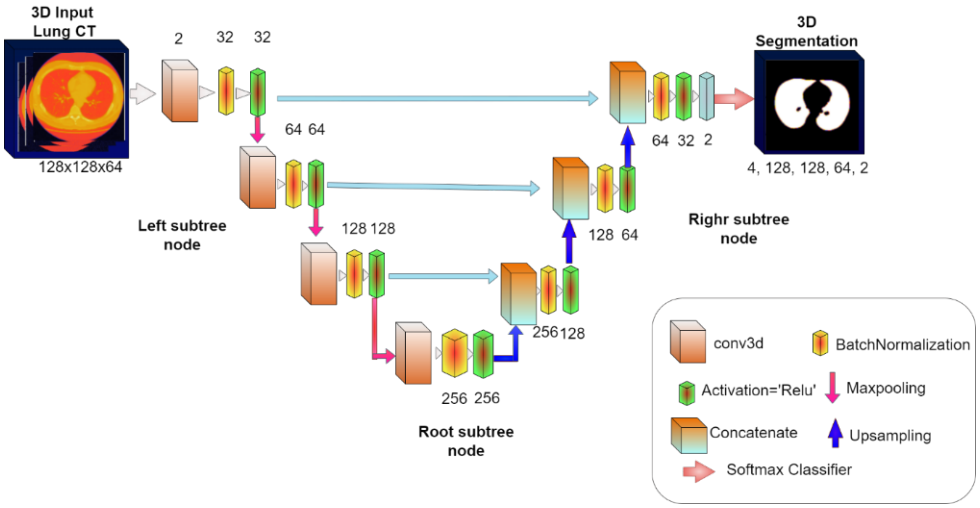

Fig 3. Proposed 3DMulti-Layer Convolutional Neural Network Architecture

Table 1. Left Sub-tree Node and Right Sub-tree Node layers model

Left Sub-tree Node (Encoding layer)				**Root Sub-tree Node**		**Right Sub-tree Node (Decoding layer)**		
Input: 128x128x64 + Normalization						**Output:4, 128, 128, 64, 2 Softmax with loss + IoU +Accuracy**		
Conv3d node (block -1)	Conv1	3x3,32, 2x2 Max-pooling	Batch normalization, Relu	Conv7	3x3,128, 2x2 Max-pooling	DeConv3d node (block -1)	3x3,128, 2x2 un-pooling	Batch normaliza-tion, Relu, UpSam-pling3D
	Conv2	3x3,32, 2x2 Max-pooling	Batch normalization, Relu				3x3,128, 2x2 un-pooling	Batch normaliza-tion, Relu, UpSam-pling3D
Conv3d node (block -2)	Conv3	3x3,64, 2x2 Max-pooling	Batch normalization, Relu			DeConv3d node (block -2)	3x3,64,2x2 un-pooling	Batch normaliza-tion, Relu, UpSam-pling3D
	Conv4	3x3,64, 2x2 Max-pooling	Batch normalization, Relu				3x3,64,2x2 un-pooling	Batch normaliza-tion, Relu, UpSam-pling3D

Continued on next page

Table 1 continued

Conv3d node (block -3)	Conv5	3x3,128, 2x2 Max-pooling	Batch normalization, Relu	DeConv3d node (block -3)	3x3,32,2x2 un-pooling	Batch normaliza-tion, Relu, UpSam-pling3D
	Conv6	3x3,128, 2x2 Max-Pooling	Batch normalization, Relu		3x3,32,2x2 un-pooling	Batch normaliza-tion, Relu, UpSam-pling3D

2.3 3DMLCNN Training and Conceptualizing Lung Cancer Image

This proposed experiment is trained, tested, and validated with 3Dimensional CT lung cancer image datasets, namely DICOM and NIfTI. 3D medical images are more complicated to map the 3D network than 2D medical images. However, 3D networks contain height, weight, and depth. Fitting those parameters in the network is very difficult, so only a small set of data is taken to annotate the k-means algorithms saved in the nifty dataset format. Lung CT images are used for training in nearly 80% of cases, while validation and testing are performed on the remaining 20%. The entire Left sub-tree Node and Right sub-tree node network parameters were trained using training images to extract cancer tumoral features. Softmax classification labels these final feature outcomes as tumor and non-tumor (cancerous and non-cancerous) pixels. After training, all other validation images are processed concurrently to evaluate the proposed model.

The proposed network is capable of identifying significant features in an input. The proposed network's first convolutional left sub-tree encoder blocks learn the edges and corners of an input image, preserving the most critical data in the lung image. The third convolutional left sub-tree encoder block creates an abstract visual representation. The right sub-tree decoder blocks increase the encoder feature dimension and generate high-level features that focus more on the tumor and non-tumor class labels and less on picture data. These samples of elements taken from every left and right sub-tree through the encoder and decoder block are visualized in Figure 4.

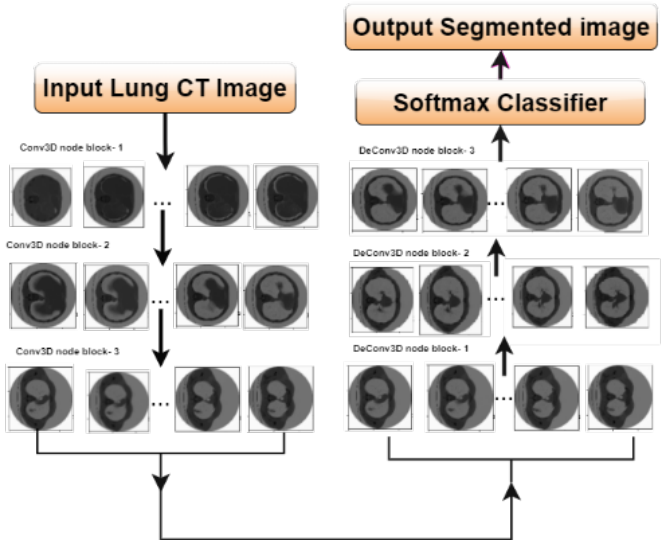

Fig 4. Feature Map Visualization of NIfTI Lung CT image Sample in the Proposed 3DMLCNN Network

2.4 Performance Evaluation Methods and Implementation

The outcome of the process can be categorized as labels generated by the proposed approach that has been compared to the ground truth, and their performance is assessed using Accuracy, Dice Score Coefficient (DSC), Sensitivity, F1 score, and IoU

as indicated in the Equations (6), (7), (8), (9), (10), (11) and (12).

$$Accuracy = \frac{(AP+AN)}{(Ap+AN+IP+IN)} \tag{6}$$

$$DSC = \frac{2*AP}{(2*(TP)+IP+IN)} \tag{7}$$

$$Sensitivity = \frac{AP}{(AP+IN)} \tag{8}$$

$$F1\ score = \frac{2*(Precision\ X\ Recall}{(Precision\ +\ Recall)} \tag{9}$$

$$Precision = \frac{AP}{(AP+IP)} \tag{10}$$

$$Recall = \frac{AP}{(AP+IN)} \tag{11}$$

$$IoU = \frac{Area\ of\ Overlap\ (AO)}{Area\ of\ Union\ (AOU)} \tag{12}$$

Where, AP (Accurate True Positive) refers to the precise count of positively detected pixels. The number of negative pixels discovered is called AN (Accurate True Negative). A group of positively determined positive pixels is called IP (Inaccurate False Positives). Incorrectly calculated negative pixels are counted as IN (Inaccurate False Negative).

2.5 Loss Function

The Loss function is an evaluation measure for detecting discrepancies between the targeted and predicted labels. Categorical Cross-Entropy is used for loss function for Convolutional neural network applications. It has two types of distribution: accurate distribution denotes a(x), and predicated distribution represents b(x) in class labels. The loss function is derived by correlating the actual distribution with the predicted distribution, given in Equation (13).

$$L(a,b) = -\sum_{i=1}^{n} a(x_i)\, log(b(x_i)) \tag{13}$$

3 Result & Discussion

3.1 Datasets and Experimental Scheme

The proposed experimental images collected from the TCIA (The Cancer Imaging Archive) data repository consist of 492 non-small lung cancer (NSCLC) patients. For these kinds of patients, preliminary CT scans are automated delineation of the 3D volume of the overall tumor volume by radiation oncology specialists. Clinical outcome data are available in the TCIA portal. Each patient image consists of two model sequences (tumor and non-tumor), and their tumor region is labeled manually using the k means algorithm and verified with experience. Training patches and specification details as shown in Table 2.

Table 2. Specification of training patches for the Train set, Validate set, and Test set in the dataset

Set	Train set	Validate set	Test set	Total
Healthy (non-tumor)	8104	506	832	9442
Lung cancer (tumor)	8104	506	832	9442
Total	16,208	1012	1664	18,884

3.2 The Proposed Delineation Performance Metrics for the Pre-processed K-Means Algorithm

The pre-processed images are more processed with the proposed K-means algorithm for segmenting the complete tumor region to label those regions and save them into the lung label (mask) images in 3D format. The training and validation outcome of the K- Means algorithm calculated by the accuracy, dice coefficient, loss, and Intersection union (IoU) of pre-processed images is discussed in Figure 5.

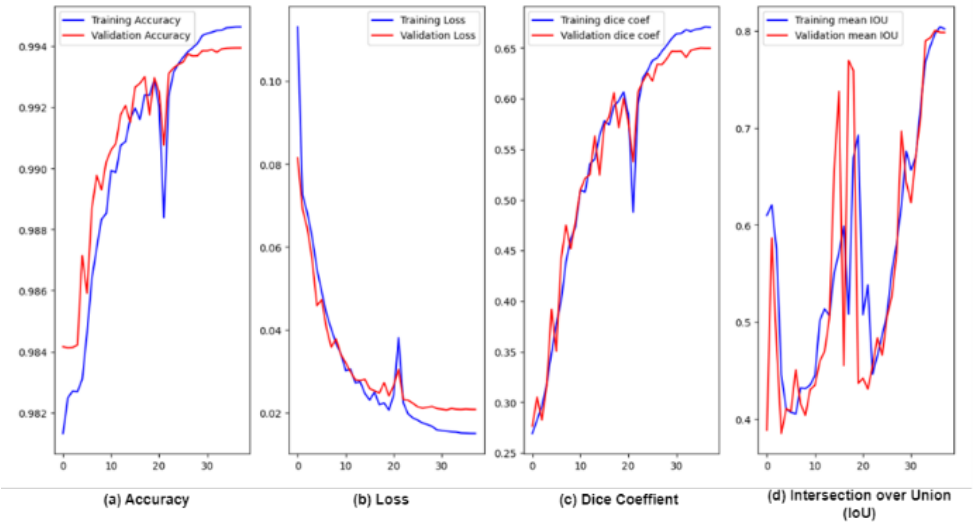

Fig 5. K-means lung labeling (Masking) with proposed results

3.3 The Proposed Performance Metrics Compared with an Existing Methodology for Lung Cancer Ground Truth Prediction Images

The proposed left and right sub-tree nodes in a network are trained and validated with lung cancer images TCIA datasets. The outcome of the proposed delineation labels is contrasted with the original ground truth labels, and their performance is assessed using Dice, Sensitivity, and F1 scores compared with the existing methodology [28]. The predicted sideway ground truth value montage view and a single image view are shown in Figure 6, and the whole ground truth image is illustrated in Figure 7. Finally, the sideway prediction of lung image results is demonstrated in Tables 3, 4 and 5.

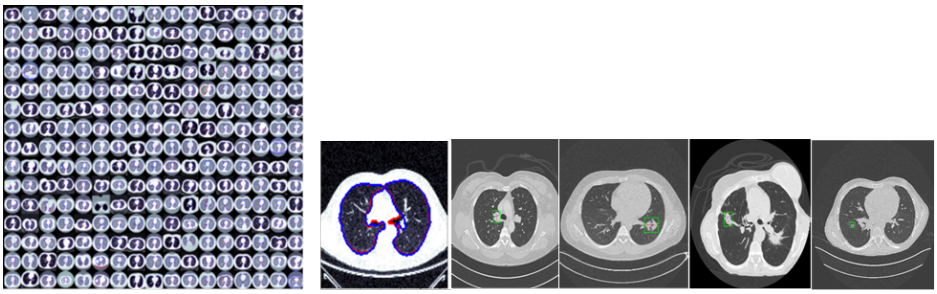

Fig 6. Predict the sideway ground truth value (a) Entire Lung Image (b) Single view lung image

The average segmentation outcomes of proposed 3D MLCNN with various sides of the count. The 3D MLCNN outperformed the proposed method and was evaluated with a student's t-test with a 0.05 significance level and compared with the existing MSDS-UNet [28], as shown in Table 3.

The average segmentation outcomes of the proposed 3D MLCNN with three sides of the count. The 3D MLCNN outperformed the proposed method and was evaluated with a student's t-test with a 0.05 significance level, as shown in Table 4.

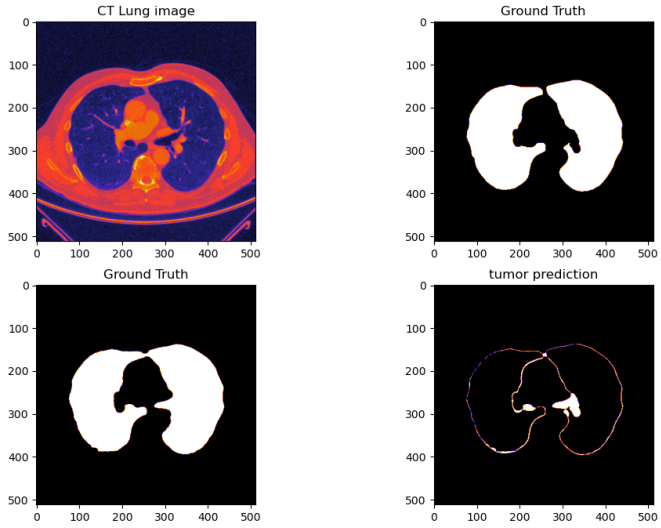

Fig 7. Prediction of whole ground truth lung images

Table 3. Comparing Dice, Sensitivity, and F1 score values of various existing methodologies of MSDS [28] w ith proposed results with ground truth sides

	(4)-side	(34)-side	(234)-side	Jinzhu Yang et al. [28] (1 234) side	Proposed (1234) side
Dice	0.654	0.657	0.673	0.664	**0. 678**
Sensitivity	0.718	0.680	0.743	0.728	**0. 765**
F1 score	0.650	0.653	0.669	0.667	**0. 808**

Table 4. Comparing Dice, Sensitivity, and F1 score values of various existing methodologies [28–30] with proposed results with three sides of ground truth

	(234)-side [29]	(134)-side [30]	(124)-side [28]	Proposed (123) side
Dice	0.654	0.657	0.673	**0. 781**
Sensitivity	0.718	0.680	0.743	**0. 782**
F1 score	0.650	0.653	0.669	**0. 877**

Table 5. Comparing Dice, Sensitivity, and F1 score values of various existing methodologies [28–30] with proposed results with ground truth sides CT image side

Model	Dice	Sensitivity	F1 score
2D U-net [29]	0.673	0.645	0.679
3D U-net [30]	0.735	0.745	0.730
3D MSDS-Unet [28] with(1234)	0.743	0.764	0.739
3D MSDS-Unet [28] with(124)	0.758	0.790	0.753
Proposed 3D MLCNN with (1234) ground truth CT image side	0.921	0.869	0.815
Proposed 3D MLCNN with(124) ground truth CT image side	0.913	0.894	0.840

The average segmentation outcomes of large-scale tumors by different model ground truth values. The proposed 3D MLCNN outperformed and was evaluated with a student's t-test with a 0.05 significance level, as shown in Table 5.

3.4 Comparison with the benchmark methods

Experimental results have been compared with an existing MSDS- U-Net [28] for the proposed network 3DMLCNN three cross-validation method. Table 6 shows the results of models concerning dice score, sensitivity, and F1 score. The results prove that the multi-layer strategy helps to identify lung tumors with higher segmentation performance and segmentation results from Nifti dataset1 and dataset2, as shown in Figure 8, and Dicom datasets are displayed in Figure 9.

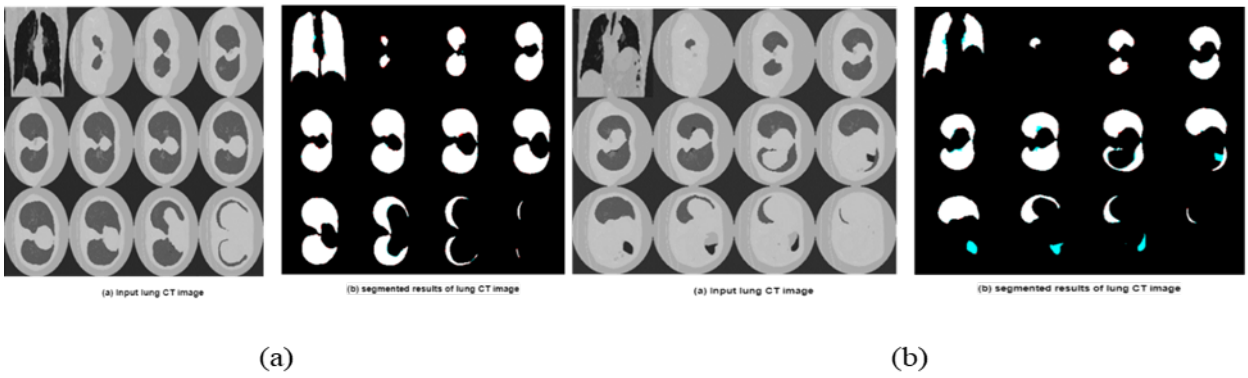

Fig 8. Segmentation Results of Lung CT Sample Image in the Nifty Dataset (a) Dataset 1 (b) Dataset 2

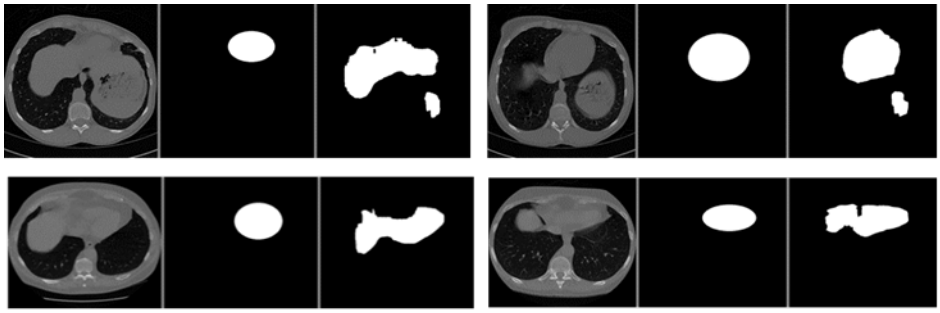

Fig 9. Segmentation Results of Lung CT Sample Image in the Dicom Dataset

The segmentation outcomes of the proposed 3D MLCNN compared with different U-Net-based segmentation models. The 3D MLCNN outperformed the proposed method and was evaluated with a student's t-test with a 0.05 significance level. It predicted the segment mask 0.5 threshold value, as shown in Table 6.

Table 6. Comparing Dice, Sensitivity, F1 score, and Mean IoU values of various existing methodologies [28] with proposed results

	2D [29] U-Net	3D U-Net [30]	3D U-Net [31] with ResNet	3D MSDS-U-Net [28]	3D MSDS-U-Net-GM [28]	Proposed 3D MLCNN (t-test)	Proposed 3D MLCNN (threshold value)
Dice	0.578	0.643	0.649	0.664	0.675	0.683	0.724
Sensitivity	0.566	0.693	0.705	0.728	0.731	0.784	0.812
F1 score	0.586	0.639	0.641	0.667	0.682	0.703	0.712
Mean IoU						0.802	0.842

4 Conclusion

This proposed network enhances the left and right sub-trees through the encoding and decoding process for segmenting lung tumors from CT images. The hierarchy of tumoral features is extracted from the lung images using three convolutional nodes. The last left sub-tree node output is processed into the following three right nodes for enlarging feature dimensions to obtain the original image size. The final extracted feature outcome is classified and labeled, and it has been compared with ground truth labels for estimating performance. This proposed network has been evaluated using TCIA datasets, demonstrating that it produces the desired model outcomes with higher achievement of 9.6% dice, 17.11% sensitivity, and 9.33% F1 score, respectively, compared to the existing state-of-the-art methods. It outperforms the current state-of-the-art methods with a 10% improvement outcome based on the proposed segmentation algorithm.

Acknowledgment

This Research work is supported by MHRD RUSA 2.0 Biomaterials – Research Innovation and Quality Improvement - Nano informatics on the role of Nano-HAp in cancer diagnostics and therapy: Schema and System Design under Functional Synthetic Material for Biomedical Applications Theme-2 which has received funding from the University of Madras.

References

1) Zhao D, Liu Y, Yin H, Wang Z. An attentive and adaptive 3D CNN for automatic pulmonary nodule detection in CT image. *Expert Systems with Applications*. 2023;211:118672. Available from: https://doi.org/10.1016/j.eswa.2022.118672.
2) Shakeel PM, Burhanuddin MA, Desa MI. Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks. *Measurement*. 2019;145:702–712. Available from: https://doi.org/10.1016/j.measurement.2019.05.027.
3) Gan W, Wang H, Gu H, Duan Y, Shao Y, Chen H, et al. Automatic segmentation of lung tumors on CT images based on a 2D & 3D hybrid convolutional neural network. *The British Journal of Radiology*. 2021;94(1126):1–9. Available from: https://doi.org/10.1259/bjr.20210038.
4) Chakravarthy SRS, Bharanidharan N, Rajaguru H. Deep Learning-Based Metaheuristic Weighted K-Nearest Neighbor Algorithm for the Severity Classification of Breast Cancer. *IRBM*. 2023;44(3):100749. Available from: https://doi.org/10.1016/j.irbm.2022.100749.
5) Nishio M, Muramatsu C, Noguchi S, Nakai H, Fujimoto K, Sakamoto R, et al. Attribute-guided image generation of three-dimensional computed tomography images of lung nodules using a generative adversarial network. *Computers in Biology and Medicine*. 2020;126:104032. Available from: https://doi.org/10.1016/j.compbiomed.2020.104032.
6) Song E, Long J, Ma G, Liu H, Hung CC, Jin R, et al. Prostate lesion segmentation based on a 3D end-to-end convolution neural network with deep multi-scale attention. *Magnetic Resonance Imaging*. 2023;99:98–109. Available from: https://doi.org/10.1016/j.mri.2023.01.015.
7) Cifci MA. SegChaNet: A Novel Model for Lung Cancer Segmentation in CT Scans. *Applied Bionics and Biomechanics*. 2022;2022:1–16. Available from: https://doi.org/10.1155/2022/1139587.
8) Najeeb S, Bhuiyan MIH. Spatial feature fusion in 3D convolutional autoencoders for lung tumor segmentation from 3D CT images. *Biomedical Signal Processing and Control*. 2022;78:103996. Available from: https://doi.org/10.1016/j.bspc.2022.103996.
9) Apostolopoulos ID, Papathanasiou ND, Panayiotakis GS, Panayiotakis. Classification of lung nodule malignancy in computed tomography imaging utilizing generative adversarial networks and semi-supervised transfer learning. *Biocybernetics and Biomedical Engineering*. 2021;41(4):1243–1257. Available from: https://doi.org/10.1016/j.bbe.2021.08.006.
10) Xu R, Wang C, Xu S, Meng W, Zhang X. Dual-stream Representation Fusion Learning for accurate medical image segmentation. *Engineering Applications of Artificial Intelligence*. 2023;123(Part B):106402. Available from: https://doi.org/10.1016/j.engappai.2023.106402.
11) Zhao J, Dang M, Chen Z, Wan L. DSU-Net: Distraction-Sensitive U-Net for 3D lung tumor segmentation. *Engineering Applications of Artificial Intelligence*. 2022;109:104649. Available from: https://doi.org/10.1016/j.engappai.2021.104649.
12) Li L, Zhao X, Lu W, Tan S. Deep learning for variational multimodality tumor segmentation in PET/CT. *Neurocomputing*. 2020;392:277–295. Available from: https://doi.org/10.1016/j.neucom.2018.10.099.
13) Ansari MY, Yang Y, Meher PK, Dakua SP. Dense-PSP-UNet: A neural network for fast inference liver ultrasound segmentation. *Computers in Biology and Medicine*. 2023;153:1–11. Available from: https://doi.org/10.1016/j.compbiomed.2022.106478.
14) Chi J, Li Z, Sun Z, Yu X, Wang H. Hybrid transformer UNet for thyroid segmentation from ultrasound scans. *Computers in Biology and Medicine*. 2023;153:106453. Available from: https://doi.org/10.1016/j.compbiomed.2022.106453.
15) Li B, Keikhosravi A, Loeffler AG, Eliceiri KW. Single image super-resolution for whole slide image using convolutional neural networks and self-supervised color normalization. *Medical Image Analysis*. 2021;68:101938. Available from: https://doi.org/10.1016/j.media.2020.101938.
16) Li P, Ma W. OverSegNet: A convolutional encoder–decoder network for image over-segmentation. *Computers and Electrical Engineering*. 2023;107:108610. Available from: https://doi.org/10.1016/j.compeleceng.2023.108610.
17) Zunair H, Hamza AB. Sharp U-Net: Depthwise convolutional network for biomedical image segmentation. *Computers in Biology and Medicine*. 2021;136:104699. Available from: https://doi.org/10.1016/j.compbiomed.2021.104699.
18) Lembhe A, Motarwar P, Patil R, Elias S. Enhancement in Skin Cancer Detection using Image Super Resolution and Convolutional Neural Network. *Procedia Computer Science*. 2023;218:164–173. Available from: https://doi.org/10.1016/j.procs.2022.12.412.
19) Ibtehaz N, Rahman MS. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*. 2020;121:74–87. Available from: https://doi.org/10.1016/j.neunet.2019.08.025.
20) Afshar P, Oikonomou A, Naderkhani F, Tyrrell PN, Plataniotis KN, Farahani K, et al. 3D-MCN: A 3D Multi-scale Capsule Network for Lung Nodule Malignancy Prediction. *Scientific Reports*. 2020;10(1):1–11. Available from: https://doi.org/10.1038/s41598-020-64824-5.
21) Borlea ID, Precup RE, Borlea AB, Iercan D. A Unified Form of Fuzzy C-Means and K-Means algorithms and its Partitional Implementation. *Knowledge-Based Systems*. 2021;214:106731. Available from: https://doi.org/10.1016/j.knosys.2020.106731.

22) Dheepa G, Chithra PL. An Efficient Encoder-Decoder CNN for Brain Tumor Segmentation in MRI Images. *IETE Journal of Research*. 2022;p. 1–12. Available from: https://doi.org/10.1080/03772063.2022.2098182.
23) Chithra PL, Dheepa G. Di-phase midway convolution and deconvolution network for brain tumor segmentation in MRI images. *International Journal of Imaging Systems and Technology*. 2020;30(3):674–686. Available from: https://doi.org/10.1002/ima.22407.
24) Ansari MY, Yang Y, Balakrishnan S, Abinahed J, Al-Ansari A, Warfa M, et al. A lightweight neural network with multiscale feature enhancement for liver CT segmentation. *Scientific Reports*. 2022;12(1):1–12. Available from: https://doi.org/10.1038/s41598-022-16828-6.
25) Dakua SP, Abi-Nahed J. Patient-oriented graph-based image segmentation. *Biomedical Signal Processing and Control*. 2013;8(3):325–332. Available from: https://doi.org/10.1016/j.bspc.2012.11.009.
26) Dakua SP. Use of chaos concept in medical image segmentation. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*. 2013;1(1):28–36. Available from: https://doi.org/10.1080/21681163.2013.765709.
27) Dakua SP. AnnularCut: a graph-cut design for left ventricle segmentation from magnetic resonance images. *IET Image Processing*. 2014;8(1):1–11. Available from: https://doi.org/10.1049/iet-ipr.2013.0088.
28) Yang J, Wu B, Li L, Cao P, Zaiane O. MSDS-UNet: A multi-scale deeply supervised 3D U-Net for automatic segmentation of lung tumor in CT. *Computerized Medical Imaging and Graphics*. 2021;92:101957. Available from: https://doi.org/10.1016/j.compmedimag.2021.101957.
29) Norman B, Pedoia V, Majumdar S. Use of 2D U-Net Convolutional Neural Networks for Automated Cartilage and Meniscus Segmentation of Knee MR Imaging Data to Determine Relaxometry and Morphometry. *Radiology*. 2018;288(1):177–185. Available from: https://doi.org/10.1148/radiol.2018172322.
30) Dou Q, Yu L, Chen H, Jin Y, Yang X, Qin J, et al. 3D deeply supervised network for automated segmentation of volumetric medical images. *Medical Image Analysis*. 2017;41:40–54. Available from: https://doi.org/10.1016/j.media.2017.05.001.
31) He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 2016;p. 770–778. Available from: https://doi.org/10.1109/CVPR.2016.90.