

# Residual edge dense enhanced module network: a deep learning approach with multi-class SVM for lung tumor stage classification

Prabakaran Jayaraman<sup>1</sup>, Pandiaraj Selvaraj<sup>2</sup>, Ashwini Elango<sup>3</sup>

<sup>1</sup>Department of Networking and Communications, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, India

<sup>2</sup>Department of Computing Technologies, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, India

<sup>3</sup>Department of Anaesthesia, Velammal Medical College Hospital and Research Institute, Madurai, India

## Article Info

### Article history:

Received Jun 6, 2024

Revised Jun 25, 2025

Accepted Jul 13, 2025

### Keywords:

Computed tomography

Deep learning

Lung cancer

Machine learning

Multi-class support vector machine

Positron emission technology

REDEM-NET

## ABSTRACT

Lung cancer segmentation with positron emission tomography (PET) and computed tomography (CT) images plays a critical role to accurately detect lung cancer. Nevertheless, lung tumor segmentation in PET/CT images were extremely difficult due to the movement caused by respiration. Despite this fact, the lung tumor images shown large number of variations mostly in PET images and CT images. As PET-CT images are acquired concurrently the shape and size of lung tumor varies according to modality. To address these issues, we developed a residual edge dense enhanced module network (REDEM-NET) framework for lung tumor stage classification. The proposed REDEM-NET can process PET and CT images as inputs. In addition, the dense residual convolutional network (DRCN) collects both inputs and extracts high-dimensional features concurrently. The extracted features from both imaging modalities were fed into UNet+++ to obtain multi-level decoded features. The extracted decoded features are concurrently supplied to the pixel level learning module (PELM) and edge level learning module (E2LM) which resulting in two outputs for subsequent learning. The outputs were merged to provide a very precise lung tumor segmentation. Furthermore, segmented tumor was fed to multi-class support vector machine (MC-SVM) for lung tumor stage classification. Moreover, it was able to identify three stages and its substages namely primary tumor, region lymph node and distant metastasis.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



## Corresponding Author:

Pandiaraj Selvaraj

Department of Computing Technologies, Faculty of Engineering and Technology

SRM Institute of Science and Technology

Kattankulathur-603203, Chengalpattu District, Tamil Nadu, India

Email: selvarap@srmist.edu.in

## 1. INTRODUCTION

Lung cancer endures one of the most extensive and deadly diseases that spreads its mortality annually. Earlier diagnosis of lung cancer and identifying the respective stages is more crucial. Lung tumor stages inclusively referred to as the tumor, node, metastasis (TNM) classification system such as the size and extent of the T, the presence of distant M and the involvement of lymph N [1]–[3]. Advanced medical imaging technologies have a major improved ability to identify and estimate lung tumors such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET). The

advances in artificial intelligence (AI), machine learning (ML), and deep learning (DL) enable more precise tumor staging [4]–[6]. AI could tackle complex and huge datasets with precise classification. The most modern developments in imaging technologies apply AI based methods for classifying the tumor stages. To avoid human error and medical modalities must be equipped with precise AI driven solutions. AI driven lung tumor classification can also be able to consider the individual patient's features in cancer cell segmentation. The DL suffers with the curse of over fitting and reduced the generalizability. However, features from the CT and PET images have varying resolutions and properties making it difficult to incorporate and arrange them properly [7]. Especially, Lung tumor is difficult to detect and categorized as the tumor stages varies with respect to size, texture and shape. In real world clinical situation supporting the accuracy of the model and reliability specifies assuring the toughness of these variations. The major contribution of this work is mentioned as follows:

- We adopt parallel feature extractors named deep residual convolutional network (RCN) for modality (i.e. CT and PET) specific feature extraction with shared weights for extracting high dimensional spatial features.
- We design UNet+++ for performing segmentation by examining high dimensional spatial features with better feature processing accuracy. The designed UNet+++ contains convolutional block attention segment (CBAS) to reduce the unwanted computational complexity. The adopting of pixel level learning module (P2LM) and edge level learning module (E2LM) enhance the tumor segmentation accuracy by effectively processing the multi scale decoded features.

The rest of the study is organized as follows; section 2 demonstrates the related works. Section 3 emphasises the material and methods needed to proposed research design. Section 4 explains the proposed residual edge dense enhanced module network (REDEM-NET) model with appropriate mathematical equations and diagrams. Section 5 implements the proposed work with existing works. Section 6 concludes the proposed work.

## 2. RELATED WORKS

Hamdi *et al.* [8] have developed a lung cancer classification method by using a multi output convolutional neural network (CNN) tool to assist lung cancer patient's stages. It refers to the TNM staging method and histologic subtypes classification. Furthermore, VGG-16 network has been incorporated with PET/CT images to extract relevant features from images. According to Kasinathan and Jayakumar [9], a cloud-based lung tumor detector and stage classifier (cloud-LTDSC) method were proposed to classify and validate lung tumor stages by utilizing DNN and cloud-based data collection. This method was validated with the benchmark lung image database consortium - image database resource initiative (LIDC-IDRI) dataset and CT digital imaging and communications in medicine (DICOM) images.

The research in [10], [11] have presented a framework that automatically localizes lung cancer from PET/CT images. Rose *et al.* [12] proposed a framework for the cancer detection with fuzzy C-means (FCM). Xiang *et al.* [13] proposed a modality-specific segmentation network (MoSNet) technique for lung tumor segmentation to yield a modality-specific map. Fu *et al.* [14] proposed a lung tumor segmentation model with multimodal spatial attention module (MSAM). Xie *et al.* [15] proposed a novel method to identify the preoperative lymph node staging in non-small cell lung cancer. Moreover, retrospective examination of 263 abnormally verified lymph nodes from 124 non-small cell lung cancer (NSCLC) patients was performed.

Xia and Zhang [16] proposed a novel DL-based graph model for tumor segmentation was. Their method has exploited the CT's spatial resolution and PET's higher contrast for multi-scale fusion and co-segmentation. Nawreen *et al.* [17] proposed a novel method for pre-processing with image enhancement and smoothing. Rehman *et al.* [18] demonstrated the idea of feature fusion with patch base local binary pattern (LBP) and discrete cosine transform. Yadav *et al.* [19] proposed a framework for chest CT and X-ray images with generative adversarial network (GAN).

Raza *et al.* [20] proposed Lung-EffNet for lung cancer by utilizing a EfficientNet from CT-scan images. This method has additional top layers for the classification head and it was evaluated by using five EfficientNet variations (B0-B4). Moreover, it was experimented on the IQ-OTH/NCCD benchmark dataset to classify lung cancer as benign or malignant. Venkatesh *et al.* [21] proposed a lung cancer lesions identification method with Otsu thresholding and CNN based cuckoo search algorithm. This framework was validated with the scaling, rotation and contrast modification of the images taken from LIDC-IDRI database [22]. Faruqui *et al.* [23] proposed a hybrid deep-CNN model named Lung Net with 22-layer hybrid deep-CNN. They have trained the model with CT scans and wearable sensor-based MIoT data. Nazir *et al.* [24] proposed a lung segmentation framework with LP decomposition and adaptive sparse representation (ASR) and validated with the images taken from LIDC-IDRI database.

Ashwaty *et al.* [25] proposed a model to detect lung tumor using Nano-segmented CT image. This method was enhanced with Gabor filter and color-based histogram equalization techniques. Moreover, this

lung cancer images were segmented by utilizing the guaranteed convergence particle swarm optimization (GCP SO) algorithm. Additionally, the tumor regions are classified using a graphical user interface tool and bag-of-visual-words (BoVW) a convolutional recurrent neural network (CRNN) was utilized for image classification and feature extraction. Crasta *et al.* [26] proposed a classification framework using cosine sail fish optimization-based generative adversarial network. This framework merges the sine cosine algorithm through the sailfish optimizer. Furthermore, the process includes pre-processing, feature extraction, lung cancer detection, severity classification and lung nodule segmentation. Additionally, CT images are segmented to detect abnormal regions.

### 3. MATERIALS AND METHODS

For our proposed research, we retrospectively collect (n=120) stage I lung tumor patients have a favorable diagnosis when compared to the patients with stage II and III respectively. To be clearer, it is difficult to victimize the stage II and III patients with good and worse prognosis. So that this study firmly focused on lung tumor patients with stage II and III respectively.

- Patient demographics: for our research, we utilize about 140 lung cancer patients with stage II and III were adopted from the Centre Hospitalier Universitaire (CHU), France in which the patients are subjected with curative chemoradiotherapy. Those patients' details were collected between the years 2017 to 2019 respectively. The inclusion criteria include patients with NSCLC with both PET and CT images and stage-II or III with subjected to radiotherapy treatment. Table 1 shows the demographics of patient characteristics. Table 2 shows the tumor staging of lung cancer.
- PET/CT image acquisition: all the considered patients must undergo PET/CT acquisition before staging and treatment approach. Entrenched on the clinical routine protocol, a biograph with 22.7cm axial view was taken on mCT 50 ToF. After 7h of fasting and  $65 \pm 7$  minutes of 3.5 MBq/kg of FDG ( $423 \pm 98$  MBq, range 225-700 MBq) PET/CT image acquisition was started.

The CT image with non-contrast enhanced, non-respiratory gates were acquired with 4 mm thickness and  $0.942 \times 0.942$  mm<sup>2</sup> in-plane thickness using 130 kVp modulation system. Furthermore, the PET images can be acquired based on the bed position arrangement with 3.5 min. The acquired PET and CT images are reconstructed using spatial resolution modelling and time of flight with 22 subsets, four iterations, voxel size of  $5 \times 5 \times 5$  mm<sup>2</sup>, and 6 mm 3D gaussian filtering approaches respectively.

Table 1. Demographics of patient

Characteristics		# of Patients (%)	Test (N=55) (%)	Train (N=90) (%)
Stage	I	3	3	3
	II	45(33)	19(32)	28(28)
	III	98(71)	36(69)	63(72)
Treatment	CHE	72(53)	25(47)	59(68)
	RAD	70(51)	30(57)	32(36)
Age	Mean±SD	73.6±9.4	73.8±12	73.6±9.4
	Range	48-96	48-91	48-96
Gender	Female	34(25)	9(16)	27(31)
	Male	106(79)	46(88)	64(73)

Table 2. Lung tumor stages

Primary tumor (TU)	Regional lymph nodes (LN)	Distant metastasis (DM)
TU0–No tumor	LN0–No regional node metastasis	DM0–No distant metastasis
TU1–Tumor≤3 cm	LN1–Ipsilateral	DM1–Malignant effusion or contralateral nodule
TU1a–Tumor≤1 cm	peribronchial/perihilar/ intrapulmonary nodes	DM2–Distant metastasis detected
TU1c–Tumor>1 cm to≤2 cm	LN2–Ipsilateral mediastinal or subcarinal nodes	
TU2–Tumor>2 cm to≤5 cm	LN3–Contralateral	
TU2a–Tumor>3 cm to≤4 cm	mediastinal/perihilar/supraclavicular nodes	
TU2b–Tumor>4 cm to≤5 cm		
TU3–Tumor>5 cm to≤7 cm		
TU4–Tumor>7 cm		

### 4. REDEMNET-MODEL

Figure 1 represents the architecture of the proposed REDEM-NET. The proposed REDEM-NET is composed of parallel feature extractors named dense residual convolutional network (DRCN) for extracting high dimensional features from both the PET and CT modalities. Note that the feature extractors shared weights among themselves to examine the bispatial features. Both produce high dimensional-spatial pixelated

features and provide them as input to the UNet<sup>+++</sup>. The UNet<sup>+++</sup> extracts multi-level decoded features and resolve the redundant computation problems by including CBAS. The output of UNet<sup>+++</sup> is depicted in form of multi-level feature maps. From the multi-level feature maps, P2LM and E2LM output the precise lung tumor segmentation result. Finally, the detected tumor is then fed to the multi class-support vector machine (MC-SVM) for multi class lung tumor stage classification with peak accuracy and precision respectively.

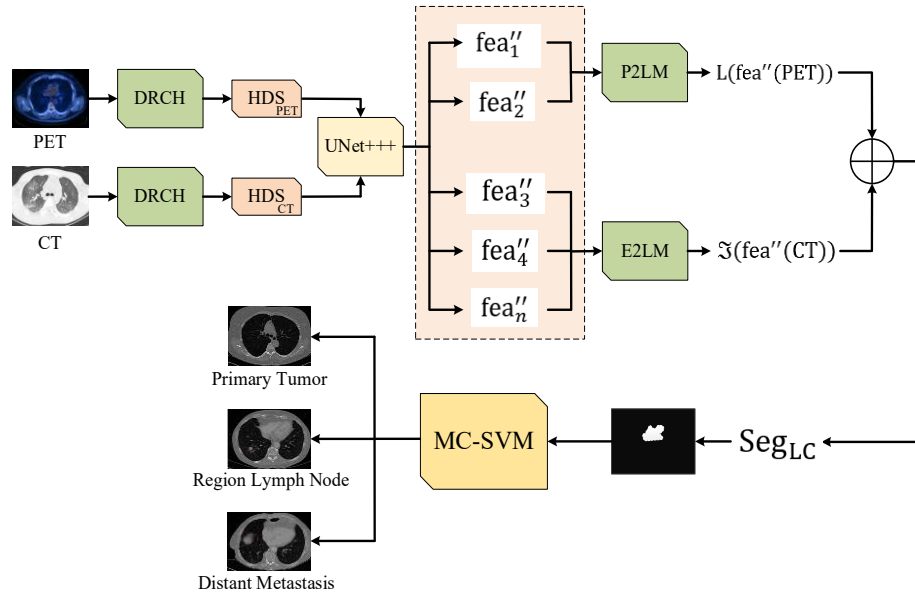


Figure 1. Overall architecture of the proposed REDEM-NET model

#### 4.1. Parallel feature extractors

The pre-processed PET and CT image  $Im'''_{PET}$ , and  $Im'''_{CT}$  is provided as an input to the parallel feature extractors named DRCN which provides output as associated pixel-superpixels  $\forall$  and high dimensional spatial features  $\aleph_{ds}$ . Note that, pixels and super-pixels in the medical images can perform transformation among each other  $\forall \in \mathbb{R}^{M \times N}$ ,  $Im \in \mathbb{R}^{M \times C}$ , and  $SP \in \mathbb{R}^{N \times C}$  in which pixels and super-pixels are represented by M and N respectively. Furthermore, the color scaling and positional factors are denoted by  $\delta^{posi}$ , and  $\delta^{CSC}$ . The formulation of  $\delta^{posi}$  is computed in (1):

$$\delta^{posi} = \beth \left( \frac{N_{\omega}, N_h}{M_{\omega}, M_h} \right) \quad (1)$$

From (1),  $N_{\omega}$ ,  $N_h$ ,  $M_{\omega}$ , and  $M_h$  denotes the number of pixels with height and width respectively. With the color scaling and positional factors, the spatial compactness and color similarities can overlook the tumor boundaries. The designed DRCN enhances the learning capability and efficacy of the feature extraction which transforms the given  $Im'''_{PET}$  or  $Im'''_{CT}$  into two level upsampled and downsampled feature maps respectively. The mathematical (1) of the dense convolutional blocks and stacked residual blocks are computed as (2) to (5):

$$IF_{DC} = PS + MPLR \uparrow_4 \quad (2)$$

$$IF_{res}^1 = Fea_{DC} \quad (3)$$

$$IF_{res}^2 = Fea_{DC} \oplus Fea_{res}^1 \quad (4)$$

$$IF_{res}^3 = (Fea_{DC} \oplus Fea_{res}^1) \oplus Fea_{res}^2 \quad (5)$$

From (2)-(5)  $IF_{DC}$  and  $Fea_{DC}$  denotes the input and output feature maps of the dense convolutional blocks respectively;  $IF_{res}^j$  and  $Fea_{res}^j$  denotes the input and output features maps in the j-th residual blocks; and  $\uparrow_4$  denotes the up-sampling operation with  $\oplus$  as addition operation of every feature element.

By entrenching dense convolutional layers, element wise addition, and stacked residual blocks sums the feature maps of every depth (i.e.64 feature maps). The deeper layer better will be spatial information in the feature maps. Totally, we have utilized 200 feature maps in triple levels and are amalgamated using the  $2 \times 2$  convolutional layer operation. To be clearer, the output residual information of the fused with the Upsample residual feature map to produce multi scale and multi branch feature maps with enriched spatial information. The outputs of the  $\text{Im}_{\text{PET}}'''$  or  $\text{Im}_{\text{CT}}'''$  are high dimensional spatial features as  $\text{HDS}^{\text{PET}}$ , and  $\text{HDS}^{\text{CT}}$  respectively.

#### 4.2. Multi-level decoded feature extraction

The output from the DRCN  $\text{HDS}^{\text{PET}}$ , and  $\text{HDS}^{\text{CT}}$  are provided as an input to the  $\text{UNet}^{+++}$  for multi-level decoded feature extraction. Those features are provided to every encoder stage  $\text{HDS}_{\text{Enc}}^1, \text{HDS}_{\text{Enc}}^2, \text{HDS}_{\text{Enc}}^3$ , and  $\text{HDS}_{\text{Enc}}^4$  and are amalgamated and passed to the CBAS in the decoder block  $\text{HDS}_{\text{Dec}}^1, \text{HDS}_{\text{Dec}}^2, \text{HDS}_{\text{Dec}}^3$ , and  $\text{HDS}_{\text{Dec}}^4$  by performing  $2 \times 2$  convolutional and batch normalization operation respectively. The major advantage of utilizing  $\text{UNet}^{+++}$  over conventional UNet is that the designed model utilizes only lesser parameters for multi scale decoded feature extraction and fusion respectively.

The skip connections in  $\text{UNet}^{+++}$  enhance multi-scale feature extraction but can cause redundancy. To address this, features from  $\text{HDS}_{\text{Dec}}^3$  are processed through  $\text{HDS}_{\text{Enc}}^1$  and  $\text{HDS}_{\text{Enc}}^2$  from different max-pooling layers.  $\text{HDS}_{\text{Enc}}^3, \text{HDS}_{\text{Dec}}^4$ , and  $\text{HDS}_{\text{Dec}}^5$  are refined using  $2 \times 2$  convolutions with sigmoid activation and bi-linear upsampling, enabling efficient semantic feature learning with reduced computational complexity.

The formulation of skip connected multiscale  $\text{UNet}^{+++}$  is provided as follows; assume that  $j$  be the present encoder and decoder layer respectively with  $M$  number if overall layers. The feature maps arranged in stacked format are denoted by  $\text{HDS}_{\text{Dec}}^j$  that can be calculated as (6):

$$\text{HDS}_{\text{Dec}}^j = \begin{cases} B \left( \text{CBN} \left[ \left( \text{BNR} \left( \text{HDS}_{\text{Enc}}^j \right), \text{UP} \left( \text{HDS}_{\text{Dec}}^k \right)_{k=j+1} \right) \right] \right)^{M-2}, j = 1 \\ B \left( \text{CBN} \left[ \left( \text{BNR} \left( \text{HDS}_{\text{Enc}}^j \right), \text{UP} \left( \text{HDS}_{\text{Dec}}^k \right)_{k=j+1}, \text{DS} \left( \text{HDS}_{\text{Dec}}^k \right)_{i=j-2} \right) \right] \right), j > 1, i > 0 \end{cases} \quad (6)$$

From (6), the CBAS is denoted by  $B(\cdot)$  which is succeeded by the ReLU activation function. The  $\text{CBN}(\cdot)$ , and  $\text{BNR}(\cdot)$  is denoted by convolution batch normalization set and ReLU batch normalization set respectively,  $\text{UP}(\cdot)$ , and  $\text{DS}(\cdot)$  denotes the up sampling and down sampling operation respectively. The adoption of CBAS with both the spatial and channel attention enhances the contextual information in multi scale fashion. In our design, decoder layer attains feature map  $\text{MSF} \in \mathbb{R}^{W \times H \times C}$  in which it is provided as an input to the CBAS. Aftermath, the channel entrenched feature map  $N_c \in \mathbb{R}^{C \times 1 \times 1}$  and spatial entrenched feature map  $N_s \in \mathbb{R}^{1 \times H \times W}$ . The final feature map output  $\text{fea}''$  is computed as (7) and (8):

$$\text{fea}' = N_c(\text{fea}) \otimes \text{fea} \quad (7)$$

$$\text{fea}'' = N_s(\text{fea}') \otimes \text{fea}' \quad (8)$$

The refined multiscale feature map  $\text{fea}''$  holds the multi scale semantic information of the lung tumors.

#### 4.3. Pixel-edge level learning

The extracted  $\text{fea}''$  is then passed simultaneously to the P2LM and E2LM for highly précised lung tumor segmentation. The P2LM is composed of series of convolutional layers which outputs the pixel wise classification map. The  $\text{fea}''(\text{PET})$  is passed to the series of  $3 \times 3$  convolution layers in which the final convolutional layer  $\text{conv}^{\text{op}}$  can be formulated as (9):

$$L = \text{conv}^{\text{op}}(\text{fea}''(\text{PET})) \quad (9)$$

From (9),  $L \in \mathbb{R}^{H \times W \times C}$  denotes the logit map here  $C$  denotes the number of classes. We apply SoftMax function to feature map channel dimension to attain the probability map  $\text{Pr}$ . The formulation of  $\text{Pr}$  is denoted as (10):

$$\text{Pr}_{j,i,c} = \frac{e^{L_{j,i,c}}}{\sum_{c'=1}^C e^{L_{j,i,c'}}} \quad (10)$$

The loss function for pixel wise learning utilized is cross entropy loss and can be authorized as (11):

$$Lo = \frac{1}{H'W'} \sum_{j=1}^{H'} \sum_{i=1}^{W'} \sum_{c=1}^{C'} Gr_{j,i,c} \log (Pr_{j,i,c}) \quad (11)$$

From (11),  $Gr \in R^{H' \times W' \times C'}$  denotes the ground truth pixel level probability map.

The edge features can be extracted from the  $fea''(CT)$  feature map using the general edge detection operator  $Y$  as shown in (12):

$$Ed = Y(fea''(CT)) \quad (12)$$

From (12),  $Ed \in R^{H \times W}$  from the  $fea''(CT)$  input. Aftermath, feature extraction can be performed based on the contextual information around the image edges. Assume that convolutional layers are adopted for feature map as (13):

$$Fea_k = \text{con}_k(Fea_{k-1}), \quad k \in \{1, 2, \dots, K\} \quad (13)$$

From (13),  $Fea_0 = fea''(CT)$  as the input image, and  $Fea_k \in R^{H_k \times W_k \times D_k}$  is the  $k$ -th layer feature map. Adopted feature map is then provided on convolutional layer for edge level classification as (14):

$$\mathfrak{S} = \text{con}_{\text{edg}}(fea'_i) \quad (14)$$

From (14),  $\mathfrak{S}$  denotes logit edge map with  $c'$  as the edge classes. In similar manner to (9), softmax operation also performed for generating edge level feature map. By combining the P2LM and E2LM, we obtain the segmented output and can be formulated as (15):

$$\text{Seg}_{LC} = L(fea''(PET)) \oplus \mathfrak{S}(fea''(CT)) \quad (15)$$

#### 4.4. Machine learning based lung cancer stage classification

From the  $\text{Seg}_{LC}$  we derive  $w$ -classes with  $w$ -binary classifiers are trained in which every binary classifiers are trained to detached the  $j$ -th class from other classes. For enabling MC-SVM, this research utilized one Vs rest (OVR) approach which splits the multiple binary classifiers. To be more distinctive, the training process with every  $w$ -th binary classifier with  $j$ -th class can be formulated as (16):

$$\text{mini}_{w_{ej}, b_{ij}, \mu_j} \frac{1}{2} \|w_{ej}\|^2 + RP \sum_{j=1}^n \mu_{ji} \quad (16)$$

Subject to the constraint in (17):

$$d_{ji}(w_{ej} \cdot \delta_i + b_{ij}) \geq 1 - \mu_{ji}, \quad \forall i = 1, 2, \dots, n \quad (17)$$

Where the label function is defined in (18):

$$d_{ji} = \begin{cases} -1 & \text{if } d_i = j \\ 1 & \text{if } d_i \neq j \end{cases} \quad (18)$$

Based on the formulation, the MC-SVM classifies the  $\text{Seg}_{LC}$  tumor into three stages as primary tumor, regional lymph node, and distant metastasis.

## 5. EXPERIMENTAL EVALUATION

### 5.1. Implementation details

We utilized the python TensorFlow and Keras packages to employ our approach. We also implemented an NVIDIA Tesla T4 GPU which Google Colab offers for training and controlling the classifier. We utilized mini batches of Size 16 to train the network for 100 epochs with the past stopping conditions set to 10 epochs to avoid over fitting. We commonly adjust the amount of the existing work and begin the training process from scratch. We also employed the Adam optimizer to train the network above to below. Especially, we place the initial learning rate to 0.0001 and the optimizer's parameters to 0.9 and 0.999, if there is no validation loss enhancement is noticed for five consecutive epochs. We also fine-tuned the recent model for five epochs at a learning rate to adversely damage the weights of the model.

### 5.2. Evaluation metrics

The employed DL techniques are evaluated in the provisions of five metrics for separate plane such as accuracy, precision, recall, F1-score, intersection over union (IoU), and area under curve (AUC).

- Accuracy: accuracy is defined as the ratio of properly detected the various pixels and the easiest metric on the five metrics.
- Precision: precision is specified as the measure of the properly detected malignant pixels to the different pixels categorized as malignant in the plane and thus it displays the better model is while the results are positive.
- Recall: recall is the piece of the malignant pixels in the bottom truth which were properly detected and classified as to all malignant pixels in the data set. A less recall signifies a huge number of fake alarms, yet a less precision value specifies a huge number of fake alarms.
- F1-score: F1-score is defined as the consonant mean of the precision and recall.
- IoU: IoU is defined as the intersection measure of the ground truth union and detected places and is the essential metric on several object classification and segmentation issues.
- AUC: AUC is specified as the place below the ROC curve. ROC curve has the recall of the design on the y-axis and false positive percentage + false positive examples (FPE). FPE true negative examples (TNE) on the x-axis. AUC is a percentage of the model's effectiveness in an increased unbalanced outline and will be evaluated as the vital below the ROC curve. The mentioned metrics are evaluated for given plane as in (19) to (22):

$$\text{Accuracy} = \frac{\text{TPE} + \text{TNE}}{\text{TPE} + \text{TNE} + \text{FPE} + \text{FNE}} \quad (19)$$

$$\text{Precision} = \frac{\text{TPE}}{\text{TPE} + \text{FPE}} \quad (20)$$

$$\text{Recall} = \frac{\text{TPE}}{\text{TPE} + \text{FNE}} \quad (21)$$

$$\text{F1 - score} = \frac{\text{TPE}}{\text{TPE} + \frac{1}{2}(\text{FPE} + \text{FNE})} \quad (22)$$

Here true positive examples (TPE), TNE, FPE, and false negative examples (FNE) signifies the amount of true positive, true negative, false positive, and false negative on the detected binary segmentation mask of a taken image. Here we have estimated the evaluation metrics above for each of the individual plane of the testing dataset and we record the common values of the mentioned metrics through all images of the testing set.

### 5.3. Comparative analysis

PET or CT scan specialists designed pixel level segmentation masks for individual 35 patients to estimate our model which represent the test set. Ground truth from the real bounding boxes was implemented to train the model. However, the evaluation metrics were evaluated among the model detection and the pixel level segmentation makes to assure the evaluation's important effectiveness. Table 3 shows the performance of the combination approach for individual plane with a 95% confidence level.

From Table 3, the proposed REDEM-Net model consistently outperforms alternative models across five folds in fusing PET or CT planes. In fold 1, it achieved 98.55% accuracy and 71.23 F1-score, surpassing modality-specific segmentation network (MSSN-Net) significantly. In fold 2, it outperformed two-stage segmentation network (TDNN-Net) with 97.32% accuracy and 80.69 F1-score. Fold 3 results showed superior performance over context-aware convolutional network (CAC-Net), while fold 4 also favored REDEM-Net over TDNN-Net. In fold 5, our model reached 96.68% accuracy and 85.69 F1-score. Overall, REDEM-Net improves accuracy by 20.81%, precision by 32.64%, F1 by 15.36%, and IoU by 20.36% on average. In Figure 2 REDEM-Net model only utilizes one modality PET or CT, generally collapse short of exactly identifying the site of the malignant tumor. This can be clarified by the truth which not the functional (PET) or the anatomical (CT) data that are important for diagnosis and carried into account. PET and CT modalities fusion crucially enhances the diagnostic abilities of the model. Moreover, the pixel wise segmentation masks are enhanced, and they can exactly define the tumor location and size after the online few shot retraining process will be implemented. We utilized the nonparametric Wilcoxon signed rank test that connects the two paired groups to explore whether the F1 and IoU development is crucial as an outcome of the PET and CT modality fusion and the presented few shots retraining approach.

Majorly it will be applied in the location of the paired student's test without acquiring the recognized data with normal distribution. We utilized the Wilcoxon test to reliably access that our approach has advanced F1 and IoU score while compared to other approaches due to the results are connected and differ from a normal distribution. We have achieved the value less than 0.02 for the F1 and IoU scores by

utilizing the Wilcoxon test. We will avoid the hypothesis with a confidence level of 99.98% while we compare the suggested REDM-Net model with different approaches.

Figure 3 presents the ROC curves and AUC values of five models evaluated for semantic segmentation performance. An AUC of 0.5 indicates random classification, while 1.0 denotes perfect accuracy. Our proposed REDEM-Net achieved the highest AUC of 0.998, outperforming CAC-Net (0.856), TDNN-Net (0.756), MSSN-Net (0.625), and MSAM-Net (0.925), demonstrating its superior diagnostic capability.

Table 3. Five-fold comparison of proposed vs existing model

Fold	Model	Accuracy ± 95% CI	Precision ± 95% CI	Recall ± 95% CI	F1 ± 95% CI	IOU ± 95% CI
Fold 1	CAC-Net [8]	96.21±0.04	65.32±2.62	35.64±2.19	43.21 ± 1.80	33.26 ± 0.06
	TDNN-Net [6]	97.25±0.06	69.75±2.35	54.32±2.15	49.35 ± 1.65	29.65 ± 0.35
	MSSN-Net [13]	96.54±0.07	71.25±2.45	49.68±2.69	56.32 ± 2.35	19.32 ± 1.25
	MSAM-Net [14]	94.32±0.05	70.65±2.05	50.36±2.25	60.65 ± 2.65	25.62 ± 2.36
	REDEM-NET	98.55±0.05	74.69±2.95	65.87±2.36	71.23 ± 3.30	40.67 ± 2.96
Fold 2	CAC-Net [8]	95.36±0.02	70.25±2.65	40.32±2.65	70.32±1.25	15.32±0.35
	TDNN-Net [6]	94.32±0.36	72.05±2.36	45.02±3.06	65.32±2.36	19.65±0.36
	MSSN-Net [13]	95.68±0.28	74.02±2.96	39.65±3.01	49.25±3.65	22.35±2.30
	MSAM-Net [14]	94.35±0.02	71.32±1.02	42.35±2.65	55.62±3.85	30.89±2.58
	REDEM-NET	97.32±0.02	77.65±2.65	55.32±2.78	80.69±2.65	35.64±3.32
Fold 3	CAC-Net [8]	96.32±0.25	69.35±2.87	45.32±2.15	69.32±2.05	20.36±0.75
	TDNN-Net [6]	92.65±0.39	60.55±2.95	39.65±2.03	70.52±3.87	25.37±3.25
	MSSN-Net [13]	93.45±0.87	71.25±2.62	42.35±2.65	71.35±1.02	30.65±1.02
	MSAM-Net [14]	94.35±0.36	72.35±2.39	40.33±2.96	72.65±1.01	24.97±2.36
	REDEM-NET	98.65±0.15	79.65±1.03	50.65±2.75	80.65±2.36	39.65±3.69
Fold 4	CAC-Net [8]	93.65±0.25	78.65±2.03	38.33±2.85	72.65±3.95	15.36±2.58
	TDNN-Net [6]	90.78±1.20	70.65±2.05	30.01±3.25	71.25±3.75	19.37±2.75
	MSSN-Net [13]	92.64±1.26	71.55±2.85	32.65±3.01	46.35±3.05	25.64±2.14
	MSAM-Net [14]	90.65±0.31	73.64±2.96	40.56±3.05	65.39±2.65	15.36±2.65
	REDEM-NET	95.66±0.20	80.65±1.26	45.64±2.65	78.36±2.75	35.61±0.58
Fold 5	CAC-Net [8]	92.35±0.12	79.36±2.75	35.67±2.65	80.56±2.89	25.68±0.31
	TDNN-Net [6]	90.36±1.32	75.36±2.58	42.97±2.05	72.56±3.07	20.31±2.35
	MSSN-Net [13]	92.87±2.36	74.37±2.48	39.67±2.09	80.23±2.65	31.48±3.25
	MSAM-Net [14]	94.36±1.35	72.89±2.60	40.23±2.33	75.62±3.98	35.67±3.69
	REDEM-NET	96.68±2.01	82.36±2.75	50.67±2.85	85.69±3.99	40.65±3.97



Figure 2. Segmentation result comparison of proposed vs existing works

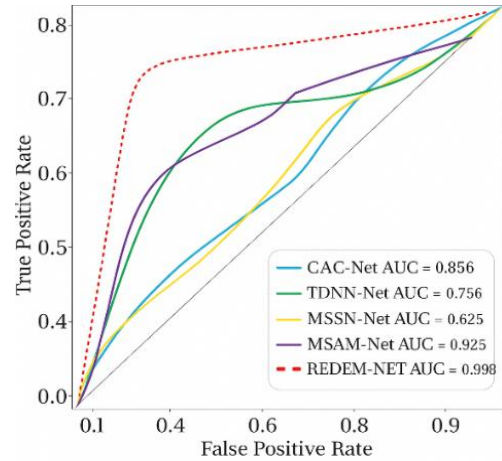


Figure 3. ROC-AUC illustration of proposed vs existing works

#### 5.4. ML based stage classification analysis

In staging, there were 28 patients in the first stage, 45 patients in second stage and 98 patients in third stage. With a common development on 41.05 months in the range of 1.8-85.1 months, the median OS was 14.3 months in the range of 1.2-51 months. At last, 75 patients have died (50%). Classification outcomes of individual ML methods and their unity following the two techniques are expressed in Table 4. Concerning the Median OS endpoint in the training set, the better model constructed by DT with 20 features that have



achieved an accuracy 85% [80% CI 55-71] in the training set. In the testing set, DT has achieved 67 % [95% CI 52-61] accuracy.

Table 4. Stage classification comparison among ML models

Classification	Accuracy (Median OS %)			Balanced accuracy (OS<6months %)		
	Training	Testing	No of patients properly classified	Training	Testing	No of patients properly classified
DT	85 [80 CI 55-71]	67 [95 CI 52-61]	30	59 [95 CI 65-78]	51 [95 CI 50-55]	26
RF	88 [95 CI 65–81]	66 [95 CI 60–70]	33	85 [95 CI 52-61]	59 [95 CI 60-70]	36
XGBoost	90 [95 CI 90–95]	65 [95 CI 50–65]	32	72 [95 CI 61-75]	55 [95 CI 75-80]	40
LR	87 [95 CI 62-75]	62 [95 CI 50–70]	35	80 [95 CI 51-71]	65 [95 CI 56-71]	42
MC-SVM	95 [95 CI 95-95]	75 [95 CI 65-76]	37	92 [95 CI 75-86]	90 [95 CI 65-76]	45

In addition, the next model RF has combined with 25 features and achieved an accuracy of 88% [95% CI 65–81] in training set. In the testing set, RF achieved 66% [95% CI 60–70] accuracy. Moreover, XG Boost combined with 30 features and attained an accuracy of 90% [95% CI 90–95] in the training set. In the testing set, XG boost achieved 65% [95% CI 50–65] of accuracy. LR combined with 32 features and achieved an accuracy of 87% [95% CI 62-75] in training set.

In the testing set, LR achieved 62% [95% CI 50–70] of accuracy. At last, our suggested MC-SVM has combined with 40 features and achieved an accuracy of 95% [95% CI 95-95] in training test. In the testing set, MC-SVM has attained 75% [95% CI 65-76] of accuracy. This means that the individual algorithms separately properly detected the patient’s survival above or below value for the mentioned algorithms. By comparing these approaches, our suggested approach has reached higher accuracy in the training set and the testing set. Concerning the second endpoint on OS below 6 months while compared to other approaches our suggested model MC-SVM has achieved higher accuracy 92% [95% CI 75-86] in training set. In the testing set, MC-SVM has attained 90% [95% CI 65-76] of accuracy.

## 6. CONCLUSION

We develop a REDEM-NET technique for accurate segment and classify lung tumor stage by processing PET and CT images. Initially, DRCN to collect input data and extract high-dimensional features simultaneously. Then the extracted features of both images are then passed into UNet++ to acquire multi-level decoded features. The decoded feature is further processed through two specialized modules, the PELM and E2LM in order to create accurate tumor segmentation. Finally, the outputs of these modules are merged to produce a precise segmentation. Then the segmented tumor is classified into stages by utilizing a MC-SVM to identifies primary tumor, region lymph node and distant metastasis to classify lung tumor stages.

## ACKNOWLEDGMENTS

The authors acknowledge SRM Institute of Science and Technology (SRMIST) for providing the necessary facilities to carry out this research and thank the reviewers for their constructive suggestions.

## FUNDING INFORMATION

Authors state no funding involved.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Prabakaran Jayaraman	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Pandiaraj Selvaraj		✓				✓		✓	✓	✓	✓	✓	✓	✓
Ashwini Elango	✓			✓		✓	✓			✓	✓	✓	✓	✓

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

## CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

## DATA AVAILABILITY

The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.





## REFERENCES

- [1] R. Fujikawa *et al.*, "Clinicopathologic and genotypic features of lung adenocarcinoma characterized by the international association for the study of lung cancer grading system," *Journal of Thoracic Oncology*, vol. 17, no. 5, pp. 700–707, 2022, doi: 10.1016/j.jtho.2022.02.005.
- [2] R. Rami-Porta, "Future perspectives on the tnm staging for lung cancer," *Cancers*, vol. 13, no. 8, 2021, doi: 10.3390/cancers13081940.
- [3] R. U. Osarogiabon *et al.*, "The international association for the study of lung cancer lung cancer staging project: overview of challenges and opportunities in revising the nodal classification of lung cancer," *Journal of Thoracic Oncology*, vol. 18, no. 4, pp. 410–418, 2023, doi: 10.1016/j.jtho.2022.12.009.
- [4] Y. Ohno *et al.*, "Small cell lung cancer staging: prospective comparison of conventional staging tests, FDG PET/CT, whole-body MRI, and coregistered FDG PET/MRI," *American Journal of Roentgenology*, vol. 218, no. 5, pp. 899–908, 2022, doi: 10.2214/AJR.21.26868.
- [5] N. M. Batouty *et al.*, "State of the art: lung cancer staging using updated imaging modalities," *Bioengineering*, vol. 9, no. 10, 2022, doi: 10.3390/bioengineering9100493.
- [6] J. Park *et al.*, "Automatic lung cancer segmentation in [18F]FDG PET/CT using a two-stage deep learning approach," *Nuclear Medicine and Molecular Imaging*, vol. 57, no. 2, pp. 86–93, 2023, doi: 10.1007/s13139-022-00745-7.
- [7] H. Shen *et al.*, "A subregion-based positron emission tomography/computed tomography (PET/CT) radiomics model for the classification of non-small cell lung cancer histopathological subtypes," *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 7, pp. 2918–2932, 2021, doi: 10.21037/qims-20-1182.
- [8] D. El Hamdi, I. Elouedi, and I. Slim, "Computer-aided classification of cell lung cancer via PET/CT images using convolutional neural network," *International Journal of Image and Graphics*, vol. 24, no. 4, 2024, doi: 10.1142/S0219467824500402.
- [9] G. Kasinathan and S. Jayakumar, "Cloud-based lung tumor detection and stage classification using deep learning techniques," *BioMed Research International*, vol. 2022, 2022, doi: 10.1155/2022/4185835.
- [10] K. Barbouchi, D. El Hamdi, I. Elouedi, T. Ben Aïcha, A. K. Echi, and I. Slim, "A transformer-based deep neural network for detection and classification of lung cancer via PET/CT images," *International Journal of Imaging Systems and Technology*, vol. 33, no. 4, pp. 1383–1395, 2023, doi: 10.1002/ima.22858.
- [11] D. Q. Wang *et al.*, "Assessing dynamic metabolic heterogeneity in non-small cell lung cancer patients via ultra-high sensitivity total-body [18F]FDG PET/CT imaging: quantitative analysis of [18F]FDG uptake in primary tumors and metastatic lymph nodes," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 49, no. 13, pp. 4692–4704, 2022, doi: 10.1007/s00259-022-05904-8.
- [12] J. D. Rose, K. Jaspin, and K. Vijayakumar, "Lung cancer diagnosis based on image fusion and prediction using CT and PET image," in *Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems*, Singapore: Springer, 2021, pp. 67–86, doi: 10.1007/978-981-15-6141-2\_4.
- [13] D. Xiang, B. Zhang, Y. Lu, and S. Deng, "Modality-specific segmentation network for lung tumor segmentation in PET-CT images," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1237–1248, 2023, doi: 10.1109/JBHI.2022.3186275.
- [14] X. Fu, L. Bi, A. Kumar, M. Fulham, and J. Kim, "Multimodal spatial attention module for targeting multimodal PET-CT lung tumor segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3507–3516, 2021, doi: 10.1109/JBHI.2021.3059453.
- [15] Y. Xie *et al.*, "A PET/CT nomogram incorporating SUVmax and CT radiomics for preoperative nodal staging in non-small cell lung cancer," *European Radiology*, vol. 31, no. 8, pp. 6030–6038, 2021, doi: 10.1007/s00330-020-07624-9.
- [16] X. Xia and R. Zhang, "A novel lung nodule accurate segmentation of PET-CT images based on convolutional neural network and graph model," *IEEE Access*, vol. 11, pp. 34015–34031, 2023, doi: 10.1109/ACCESS.2023.3262729.
- [17] N. Nawreen, U. Hany, and T. Islam, "Lung cancer detection and classification using CT scan image processing," in *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, 2021, pp. 1–6, doi: 10.1109/ACMI53878.2021.9528297.
- [18] A. Rehman, M. Kashif, I. Abunadi, and N. Ayesha, "Lung cancer detection and classification from chest CT scans using machine learning techniques," in *2021 1st International Conference on Artificial Intelligence and Data Analytics, CAIDA 2021*, 2021, pp. 101–104, doi: 10.1109/CAIDA51941.2021.9425269.
- [19] P. Yadav, N. Menon, V. Ravi, and S. Vishvanathan, "Lung-GANs: Unsupervised representation learning for lung disease classification using chest CT and X-ray images," *IEEE Transactions on Engineering Management*, vol. 70, no. 8, pp. 2774–2786, 2023, doi: 10.1109/TEM.2021.3103334.
- [20] R. Raza *et al.*, "Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images," *Engineering Applications of Artificial Intelligence*, vol. 126, 2023, doi: 10.1016/j.engappai.2023.106902.





- [21] C. Venkatesh, K. Ramana, S. Y. Lakkisetty, S. S. Band, S. Agarwal, and A. Mosavi, "A neural network and optimization-based lung cancer detection system in CT images," *Frontiers in Public Health*, vol. 10, 2022, doi: 10.3389/fpubh.2022.769692.
- [22] A. R. Bushara and R. S. V. Kumar, "Deep learning-based lung cancer classification of CT images using augmented convolutional neural networks," *Electronic Letters on Computer Vision and Image Analysis*, vol. 21, no. 1, pp. 130–142, 2022, doi: 10.5565/REV/ELCVIA.1490.
- [23] N. Faruqui, M. A. Yousuf, M. Whaiduzzaman, A. K. M. Azad, A. Barros, and M. A. Moni, "LungNet: A hybrid deep-CNN model for lung cancer diagnosis using CT and wearable sensor-based medical IoT data," *Computers in Biology and Medicine*, vol. 139, 2021, doi: 10.1016/j.combiomed.2021.104961.
- [24] I. Nazir, I. U. Haq, M. M. Khan, M. B. Qureshi, H. Ullah, and S. Butt, "Efficient pre-processing and segmentation for lung cancer detection using fused CT images," *Electronics*, vol. 11, no. 1, 2022, doi: 10.3390/electronics11010034.
- [25] S. U. Aswathy, P. P. F. Rajeena, A. Abraham, and D. Stephen, "Deep learning-based BoVW–CRNN model for lung tumor detection in nano-segmented CT images," *Electronics*, vol. 12, no. 1, 2023, doi: 10.3390/electronics12010014.
- [26] L. J. Crasta, R. Neema, and A. R. Pais, "A novel deep learning architecture for lung cancer detection and diagnosis from computed tomography image analysis," *Healthcare Analytics*, vol. 5, 2024, doi: 10.1016/j.health.2024.100316.

## BIOGRAPHIES OF AUTHOR







**Prabakaran Jayaraman**     is currently working as an Assistant Professor in the Department of Networking and communications, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, India. He has been working in SRMIST for the past 7 years. He published 7 indexed journals in the fields of ML, deep learning, and computer vision. He is constantly involved in research related to deep learning and transformer models in medical applications. He can be contacted at email: [prabakaj@srmist.edu.in](mailto:prabakaj@srmist.edu.in).



**Pandiaraj Selvaraj**     is currently working as an Associate Professor in the Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamil Nadu, India. He has been working in SRMIST for the past 18 years. He published 37 indexed journals in the fields of IoT, bigdata, computer vision, and intelligent networks. He is constantly involved in research related to artificial intelligence with IoT. He is currently guiding 5 research scholars. He can be contacted at email: [selvarap@srmist.edu.in](mailto:selvarap@srmist.edu.in).



**Ashwini Elango**     is doing her postgraduate studies in the Department of Anaesthesia at Velammal Medical College Hospital and Research Institute. She is involved in research related to the application of modern technologies in the medical field. She is interested in research regarding pharmacological drugs and their therapeutic effects, as well as their adverse effects. She can be contacted at email: [cashwini2k@gmail.com](mailto:cashwini2k@gmail.com).