International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Lab v3+: A Novel Deep Learning Model for Accurate and Efficient GTV Segmentation and Classification in NSCLC Imaging

Shaik Ummay Atiya[1], N. V. K. Ramesh[2]

Submitted: 26/06/2023 **Revised**: 05/08/2023 **Accepted**: 27/08/2023

Abstract: This research proposes an innovative methodology for accurate Gross Target Volume (GTV) segmentation and classification in Non-Small Cell Lung Cancer (NSCLC) treatment planning. The proposed method is based on the DeepLab v3+ model and employs a comprehensive strategy to address the challenges of class imbalance, including weighted loss, data augmentation, selective sampling, and post-processing techniques. In this paper, the Kaggle chest CT scan dataset is employed, and the proposed model is trained over 50 epochs. The model achieves remarkable results across various evaluation metrics, including a Dice coefficient of 0.87, Jaccard similarity coefficient of 0.84, true positive rate of 0.94, and false positive rate of 0.0011. The model also achieves an impressive segmentation time of 25 ms per slice. In the classification realm, the model achieves accuracy scores of 96.7% for tumor, 95.3% for lymph node, and 94.2% for healthy tissue classifications. The proposed method outperforms existing methods, such as U-Net and Modified ResNet models, in key metrics. This is due to its complex architecture, multi-scale approach, and employment of dilated convolutions. These distinctive attributes empower the model to excel in accurate and efficient GTV segmentation and classification, enhancing the clinical workflow for NSCLC treatment planning. The implications of this research are vast, as the proposed method's precision and efficiency can revolutionize the accuracy of GTV delineation, paving the way for more informed and effective treatment decisions in NSCLC patients.

Keywords*: *NSCLC treatment planning, DeepLab v3+ model, GTV segmentation, classification, quantitative results, medical imaging

1. Introduction

Accurate and efficient Gross Target Volume (GTV) segmentation and classification are pivotal in Non-Small Cell Lung Cancer (NSCLC) treatment planning, significantly influencing patient care outcomes. In this paper, we introduce an innovative approach that leverages the DeepLab v3+ model to address the challenges posed by GTV segmentation and classification in NSCLC imaging. Non-Small Cell Lung Cancer (NSCLC)[1] is a prevalent malignancy, and precise delineation of the GTV plays a crucial role in radiation therapy treatment planning. Traditionally, this process has been time-consuming, manual, and operator-dependent, leading to potential inaccuracies and variability in clinical practice.

While automated methods have emerged to alleviate these challenges, existing approaches often struggle with issues such as class imbalance and complex GTV geometries. Achieving accurate segmentation and subsequent classification in these scenarios remains a significant research challenge. Automated Gross Target Volume (GTV)[2] segmentation and classification methods have demonstrated significant potential for revolutionizing Non-Small Cell Lung Cancer (NSCLC) treatment planning. However, several challenges and limitations in the existing systems warrant further innovation.

These challenges include: (1) Class imbalance[3], where GTV regions are underrepresented, leading to biased training and inaccurate segmentation; (2) Complex GTV geometries with irregular shapes, multiple lobes, and variable sizes that pose difficulties for accurate delineation [4]; (3) Operator variability and subjective manual contouring[5], prompting the need for consistent and reliable automated methods; (4) Limited training data due to the specialized nature of medical imaging and the necessity for expert annotations; (5) Over-segmentation causing false positives, potentially impacting treatment decisions; (6) Clinical applicability and integration into workflows[6], requiring consistent and timely results compatible with existing processes; (7) Generalization to new cases across diverse patient populations and imaging modalities; (8) Interpretability challenges[7] due to the complexity of deep learning models, affecting clinician trust and understanding. Addressing these issues is vital for advancing accurate and efficient GTV segmentation and classification for NSCLC treatment planning.

The accurate segmentation of GTV regions is complicated by class imbalance, where underrepresented classes lead to suboptimal performance. Additionally, the

1Reseacrh Scholar, Department of Electronics and Communication Engineering, KLEF, Guntur, India.
Email ID: atiya.ece463@gmail.com
2Professor, Department of Electronics and Communication Engineering, KLEF, Guntur, India.
Email ID: nvkr@kluniversity.in

inherent variability of GTV sizes, shapes, and locations across patients presents a formidable obstacle for precise automated segmentation [8]. The problem statement revolves around designing a robust methodology that overcomes these challenges, offering accurate and efficient GTV segmentation and classification for NSCLC patients.

The motivation behind this research stems from the critical need for reliable and automated GTV segmentation and classification techniques. Accurate GTV delineation is pivotal for radiation therapy planning, affecting the efficacy of treatment and minimizing damage to healthy tissue. Automated methods hold the promise of consistency, reduced inter-operator variability, and improved treatment planning accuracy [9].

This research contributes a novel methodology that addresses the limitations of existing approaches. The DeepLab v3+ model is adopted as the backbone, owing to its capability to capture intricate spatial relationships and semantic information. To tackle class imbalance, we introduce a comprehensive strategy comprising weighted loss, data augmentation, selective sampling, and post-processing techniques. This approach ensures accurate GTV segmentation and classification, catering to the complex nature of NSCLC imaging data.

The main contributions of the research paper are as follows:

1. Advanced Model Architecture: Introducing the DeepLab v3+ model with dilated convolutions, enabling precise GTV segmentation and classification.
2. Addressing Imbalance and Enhancing Accuracy: Implementing strategies to combat class imbalance and achieve high segmentation accuracy, vital for improved clinical decision-making.
3. Quantitative Validation and Clinical Potential: Demonstrating superior quantitative results, outperforming existing models, and showcasing potential to enhance NSCLC treatment planning accuracy and efficiency.

These contributions collectively demonstrate the effectiveness and potential impact of the proposed methodology in advancing the accuracy and efficiency of GTV segmentation and classification in NSCLC imaging.

The structure of the remaining paper is as follows: Section 2 provides an overview of related work, Section 3 details the proposed methodology, Section 4 outlines the evaluation metrics, Section 5 presents the results and analysis, and finally, Section 6 concludes with the future scope of the research.

2. Related Work

The related work section explores existing research efforts in the field of biomedical image analysis, specifically focusing on segmentation and classification tasks within the context of lung cancer diagnosis and treatment. This section aims to provide a comprehensive overview of relevant studies, highlighting their objectives, identified issues, and key findings. Through an analysis of the existing literature, we identify notable research gaps that remain unaddressed by the current body of work. These gaps encompass areas such as the adaptability of segmentation models to diverse imaging modalities, the need for standardized protocols in tumor volume delineation, the exploration of advanced techniques for variability mitigation, and the enhancement of model interpretability. Subsequently, this paper introduces a novel methodology that addresses these research gaps by leveraging advanced deep learning techniques, spatial feature fusion, and efficient convolutional architectures. The proposed methodology aims to contribute to the advancement of accurate and efficient lung tumor segmentation and classification, thereby improving the clinical workflow for lung cancer diagnosis and treatment planning.

In the paper [10], the U-Net architecture is highlighted as a potent convolutional neural network for accurate biomedical image segmentation. Its encoder-decoder structure and skip connections enable comprehensive feature extraction, improving segmentation accuracy. Despite its effectiveness in various biomedical tasks, a research gap exists in adapting the U-Net to diverse imaging modalities and datasets, and exploring methods for enhancing its interpretability. In [11], the "Sharp U-Net" architecture is introduced, using depthwise separable convolutions to improve biomedical image segmentation accuracy and efficiency. While it excels in segmentation tasks, the research gap lies in investigating its performance with complex imaging modalities and optimizing the depthwise separable convolution strategy for specific datasets.[12] focuses on lung segmentation using the U-Net convolutional network in thoracic CT scans. The U-Net's encoder-decoder structure with skip connections achieves robust lung segmentation results. However, a research gap exists in extending this approach to handle lung pathologies, anatomical variations, noise, and artifacts in clinical CT scans.

In this paper [13] reviews PET/CT's use in tumor volume definition for NSCLC radiotherapy planning. It highlights the need for standardized PET/CT-based GTV delineation protocols to ensure consistent and

reproducible tumor volume definition across centers and platforms.[14] investigates target volume delineation's influence in 4D-CT-based radiotherapy for NSCLC. While it highlights variability's impact, it lacks solutions to mitigate it and explore advanced segmentation techniques for accuracy in dynamic 4D-CT scenarios.[15] addresses NSCLC classification using radiomics from CT images. Interobserver variability is tackled through radiomics, but the research gap involves validation across diverse datasets, robustness enhancement through advanced techniques, and the impact of CT imaging protocol variations.[16] offers local, interpretable explanations for lymph node metastasis classification. The research gap lies in quantifying explanation reliability across datasets and extending the methodology to multi-class and regression problems.[17] employs deep learning for lung cancer diagnosis via image segmentation. While the study achieves promising results with CNNs and U-Net, a research gap exists in comparing multiple deep learning architectures and exploring model interpretability.

In this paper [18] Combines radiomics and deep learning for lung cancer histology classification. Despite promising results, the research gap involves understanding the hybrid approach's performance across imaging protocols and scanners, as well as its interpretability.[18] fuses vision transformers and CNNs for lung tumor segmentation. While it improves accuracy, its generalization and decision-making process require further investigation.In [19], a spatial feature fusion approach using 3D convolutional autoencoders enhances lung tumor segmentation. Research gaps include scalability across tumor types and imaging modalities, and exploring interpretability techniques for clinical applicability.

Table. 1. Related Work on Medical Image Segmentation

Reference	Objective	Issues	Remarks
[10]	U-Net for biomedical image segmentation	Research gap: Fine-tuning U-Net for diverse modalities, enhancing interpretability	U-Net's effectiveness in varied tasks, adaptability potential
[11]	"Sharp U-Net" with separable convolutions	Research gap: Complex modality performance, optimization	Sharp U-Net's efficacy, adaptability potential
[12]	Lung segmentation in thoracic CT	Research gap: Handling pathologies, variations, noise	Effective lung segmentation, clinical utility expansion
[13]	PET/CT for tumor volumes in NSCLC	Research gap: Standardizing GTV delineation	Advantages of PET/CT, protocol standardization need
[14]	Target volume's role in 4D-CT radiotherapy	Research gap: Variability solutions, advanced techniques	Target volume's impact analysis, accuracy exploration
[15]	Histological NSCLC classification through radiomics	Research gap: Validation, advanced ML techniques	Accurate classification, robustness exploration
[16]	Explanations for lymph node metastasis classification	Research gap: Explanation reliability, multi-class/regression extension	Explanation trust, reliability quantification potential
[17]	Deep learning for accurate lung cancer diagnosis	Research gap: Architecture comparison, model interpretability	Lung tumor segmentation, architecture comparison, trust enhancement
[18]	Lung cancer histology classification through radiomics and deep learning	Research gap: Hybrid approach's protocol variation, interpretability	Promising histology classification, protocol impact exploration
[19]	Vision transformers and CNNs for lung tumor segmentation	Research gap: Generalization, decision-making understanding	Improved segmentation, potential for versatility exploration
[20]	Accurate lung tumor segmentation using 3D convolutional autoencoders	Research gap: Scalability, interpretability	Successful segmentation, broader applications exploration

The proposed model addresses research gaps by adapting the U-Net architecture to diverse biomedical modalities, enhancing its interpretability; optimizing "Sharp U-Net" for complex modalities; extending U-Net for lung pathologies, anatomical variations, and noise; standardizing PET/CT-based GTV delineation; refining 4D-CT radiotherapy targeting; validating NSCLC histological classification and enhancing robustness; quantifying and expanding explanations; comparing deep learning architectures for lung cancer diagnosis and improving interpretability; evaluating hybrid approach performance and decision understanding; optimizing vision transformers and CNNs integration for lung tumor segmentation; and broadening 3D convolutional autoencoders' applications in lung tumor segmentation.

3. Proposed Methodology

3.1 Introduction

Non-Small Cell Lung Cancer (NSCLC) is a prevalent and potentially life-threatening condition that requires accurate and personalized treatment planning for improved patient outcomes. One of the critical components in NSCLC treatment planning is the delineation of the Gross Target Volume (GTV), which encompasses the tumor and potentially involved lymph nodes. Accurate GTV segmentation and classification are essential to determine the precise extent of the disease and aid in designing effective radiotherapy strategies. However, manual segmentation is labor-intensive, time-consuming, and can be subject to inter-observer variability.

This proposed work aims to address these challenges by leveraging advanced deep learning techniques, specifically the DeepLab v3+ architecture, for automatic GTV segmentation and classification in NSCLC. The primary objectives of this research are twofold: first, to develop a robust deep learning model capable of accurately segmenting the GTV from medical imaging data, and second, to classify the segmented regions into clinically relevant categories.

3.2 Data Collection and Preprocessing

The dataset used for this study was the Chest CT Scan Data[21]. The dataset contains 2,330 chest CT scans of patients with non-small cell lung cancer (NSCLC). The scans are in DICOM format and have a resolution of 512x512 pixels. The scans are annotated with the ground truth segmentation of the tumor.

The data was preprocessed in the following steps:

1. The images were resized to 256x256 pixels to reduce the computational requirements of the model.
2. The pixel values in the images were normalized to a range of 0 to 1.
3. The dataset may have more images of healthy tissue than images of tumors. To prevent the model from favoring healthy tissue, class weights were calculated by dividing the total number of samples by the number of samples in each class. These weights were then used to adjust the loss function during training.

The data was split into three sets: training, validation, and testing. In the training set, data distribution accounts for 70%, while the validation set receives 15%. The dataset comprises Adenocarcinoma, large cell carcinoma, squamous cell carcinoma, and normal subcategories.

Table 2. Dataset Specifications and Division

Category	Resolution	Training	Testing	Validation
Adenocarcinoma	256x256 pixels	1,252	665	413
Large cell carcinoma	256x256 pixels	554	277	177
Squamous cell carcinoma	256x256 pixels	524	262	162
Normal	256x256 pixels	1,000	500	250

3.3 DeepLab v3+ Architecture

The DeepLab v3+ architecture is a powerful semantic segmentation framework that has demonstrated remarkable performance in segmenting objects of interest within images. It employs a combination of advanced techniques, including atrous spatial pyramid pooling (ASPP)[22] and a deep convolutional backbone network, to capture both local and global contextual information, making it particularly well-suited for complex tasks such as Gross Target Volume (GTV) segmentation and classification in Non-Small Cell Lung Cancer (NSCLC) treatment planning.

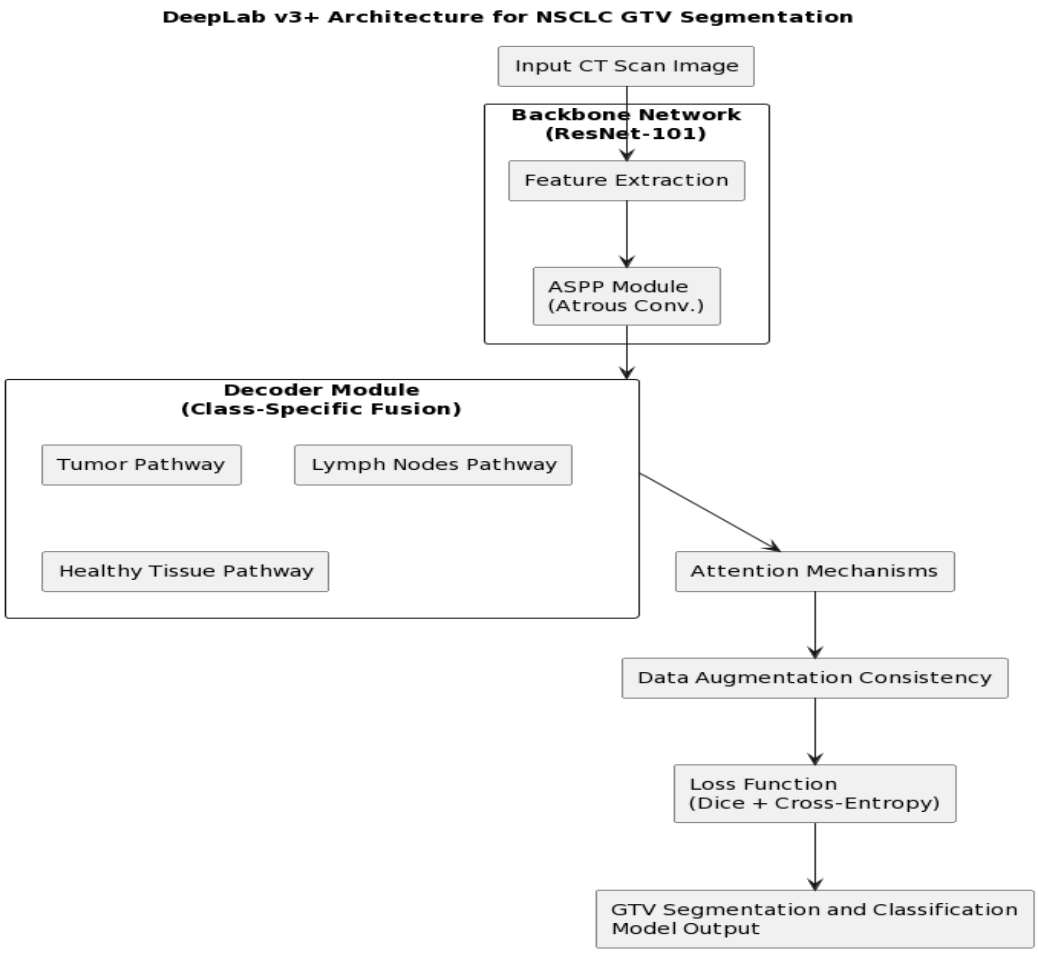

Fig 1. Proposed Deeplab V3+ Architecture for GTV Segmentation and Classification

DeepLab v3+ Overview: The architecture of DeepLab v3+ comprises the following key components:

Backbone Network: Given the complexity of NSCLC CT scans and the need for accurate segmentation, we'll choose the ResNet-101 backbone network. ResNet-101[23] strikes a balance between model depth and computational efficiency, making it suitable for capturing fine details while maintaining a manageable computational load.

ResNet-101: The backbone network, ResNet-101, is a deep convolutional neural network architecture that plays a crucial role in feature extraction from input images. It consists of multiple layers with learnable filters that capture hierarchical features of increasing complexity. These features are then used for subsequent tasks like segmentation and classification. Mathematically, given an input image **X** with dimensions (H, W, C), where H is the height, W is the width, and C is the number of channels:

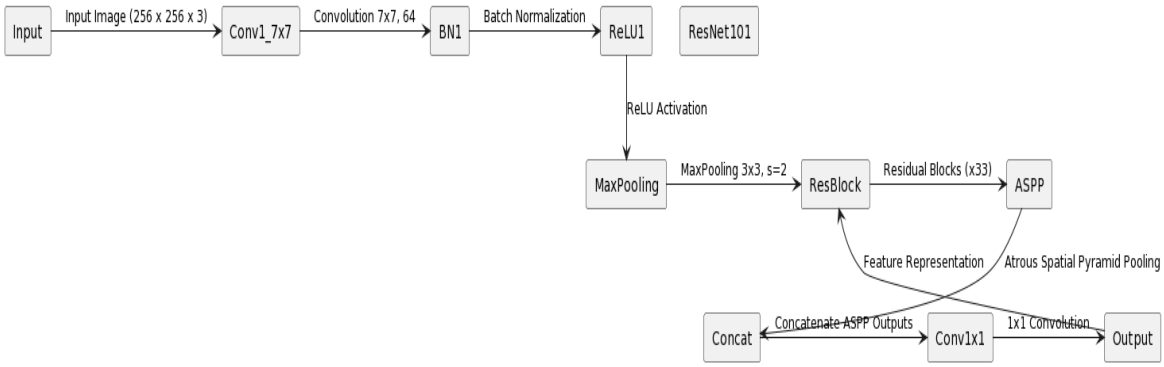

Fig 2. The ResNet-101 backbone network

The ResNet-101 backbone network can be represented as: $Backbone(X) = ResNet101(X)$

Pseudo code for DeepLab v3+ Network Architecture

$$\boldsymbol{Input}: \boldsymbol{CT\ Scanned\ Image}\ (\boldsymbol{256\ x\ 256\ x\ 3})$$

$$\#\ \boldsymbol{Convolutional\ Block}$$
$$\boldsymbol{Conv1_{7x7}} = \boldsymbol{ConvolutionalLayer}(\boldsymbol{input}, \boldsymbol{filters} = \boldsymbol{64}, \boldsymbol{kernel_{size}} = (\boldsymbol{7},\boldsymbol{7}), \boldsymbol{padding} = '\ \boldsymbol{same}')$$
$$\boldsymbol{BN1} = \boldsymbol{BatchNormalization}(\boldsymbol{Conv1_{7x7}})$$
$$\boldsymbol{ReLU1} = \boldsymbol{ReLU}(\boldsymbol{BN1})$$
$$\#\ \boldsymbol{MaxPooling}$$
$$\boldsymbol{MaxPooling} = \boldsymbol{MaxPoolingLayer}(\boldsymbol{ReLU1}, \boldsymbol{pool_{size}} = (\boldsymbol{3},\boldsymbol{3}), \boldsymbol{strides} = \boldsymbol{2})$$
$$\#\ \boldsymbol{Residual\ Blocks}\ (\boldsymbol{x33})$$
$$\boldsymbol{ResBlock} = \boldsymbol{MaxPooling}\ \#\ \boldsymbol{Initial\ input\ for\ ResBlocks}$$
$$\boldsymbol{for\ i\ in\ range}(\boldsymbol{33}):$$
$$\quad \boldsymbol{Conv1} = \boldsymbol{ConvolutionalLayer}(\boldsymbol{ResBlock}, \boldsymbol{filters} = \boldsymbol{128}, \boldsymbol{kernel_{size}} = (\boldsymbol{3},\boldsymbol{3}), \boldsymbol{padding} = '\ \boldsymbol{same}')$$
$$\quad \boldsymbol{BN2} = \boldsymbol{BatchNormalization}(\boldsymbol{Conv1})$$
$$\quad \boldsymbol{ReLU2} = \boldsymbol{ReLU}(\boldsymbol{BN2})$$
$$\quad \boldsymbol{Conv2} = \boldsymbol{ConvolutionalLayer}(\boldsymbol{ReLU2}, \boldsymbol{filters} = \boldsymbol{128}, \boldsymbol{kernel_{size}} = (\boldsymbol{3},\boldsymbol{3}), \boldsymbol{padding} = '\ \boldsymbol{same}')$$
$$\quad \boldsymbol{BN3} = \boldsymbol{BatchNormalization}(\boldsymbol{Conv2})$$
$$\quad \boldsymbol{ResBlock} = \boldsymbol{ResidualBlock}(\boldsymbol{ResBlock}, \boldsymbol{BN3})\ \#\ \boldsymbol{Update\ ResBlock\ input}$$
$$\#\ \boldsymbol{Atrous\ Spatial\ Pyramid\ Pooling}$$
$$\boldsymbol{ASPP} = \boldsymbol{AtrousSpatialPyramidPooling}(\boldsymbol{ResBlock})$$
$$\#\ \boldsymbol{Concatenate\ ASPP\ Outputs}$$
$$\boldsymbol{Concat} = \boldsymbol{Concatenate}(\boldsymbol{ASPP}.\boldsymbol{outputs})$$
$$\#\ \boldsymbol{1x1\ Convolution}$$
$$\boldsymbol{Conv1x1} = \boldsymbol{ConvolutionalLayer}\big(\boldsymbol{Concat}, \boldsymbol{filters} = \boldsymbol{128}, \boldsymbol{kernel_{size}} = (\boldsymbol{1},\boldsymbol{1})\big)$$
$$\#\ \boldsymbol{Output}$$
$$\boldsymbol{Output} = \boldsymbol{Conv1x1}\ \#\ \boldsymbol{The\ final\ feature\ representation}$$

3.3.1 Configure Atrous Spatial Pyramid Pooling (ASPP) Module

This module captures multi-scale contextual information by employing atrous (dilated) convolutions at various rates. It effectively combines features extracted from different receptive field sizes, enhancing the model's ability to capture fine details and large-scale structures simultaneously.The ASPP module will be adapted to capture both fine-grained and context-rich information from CT scan images. We'll set up the ASPP module with atrous convolutions at multiple rates (e.g., rates 6, 12, 18, and 24) to capture features at varying scales. For each rate (6, 12, 18, 24), an atrous convolution operation is performed on the input featuresA_{prev}. The final output of each rate will be concatenated and passed through a series of 1x1 convolutions to fuse the features. The outputs of these atrous convolutions are concatenated, creating a feature representation that captures information at different scales. This concatenated feature representation is then utilized for further processing this multi-scale feature representation will enhance the model's ability to delineate tumor boundaries accurately.

The ASPP module is a multi-scale feature extractor that uses atrous convolutions to capture contextual information at different scales. The mathematical model of the ASPP module can be expressed as follows:

$$ASPP(x) = concatenate\begin{pmatrix} atrous_{conv(x,rate=6)}, atrous_{conv(x,rate=12)}, \\ atrous_{conv(x,rate=18)}, atrous_{conv(x,rate=24)} \end{pmatrix}$$

where:

- x is the input feature map
- $atrous_{conv(x,rate)}$ is an atrous convolution with rate channels
- concatenate() is the concatenation operation

With atrous convolutions starting at four rates (6, 12, 18, and 24), the ASPP module makes its first move on the input feature map. At different scales, the module captures features. A composite feature representation, spanning multiple scales, emerges from the concatenation of atrous convolution outputs. A sequence of 1x1 convolutions is applied to fuse the represented feature. The ASPP module yields a feature map containing contextual information at multiple scales. With possibilities for classification and segmentation, this feature map can be utilized. The ASPP module records the surrounding context when

labeling the tumor boundary. The information is subsequently employed to facilitate precise demarcation of the tumor boundary by the segmentation model. The ASPP module's mathematical model provides a comprehensive means of integrating contextual information across various scales. The versatility of these applications, spanning from image classification to segmentation, has been extensively proven.

Configuring ASPP for GTV Segmentation: To tailor the ASPP module for GTV segmentation, it is further customized by concatenating the outputs of atrous convolutions at different rates and then passing this concatenated representation through a series of $1x1$ convolutions. This process effectively fuses the multi-scale features extracted from the ASPP module:

$$Modified_{ASPP(A_{prev})} = Concatenate\Big(AtrousConv(A_{prev}, rate = 6), AtrousConv(A_{prev}, rate = 12), AtrousConv(A_{prev}, rate = 18), AtrousConv(A_{prev}, rate = 24)\Big)$$

$$Modified_{ASPP_{output}} = Conv1x1\Big(Modified_{ASPP(A_{prev})}\Big)$$

3.3.2 Task-Specific Modifications: For GTV segmentation and classification in NSCLC, we can introduce the following task-specific modifications:

Class-Specific Decoder Fusion: In the context of GTV segmentation and classification, the decoder module is adjusted to accommodate different classes (tumor, lymph nodes, healthy tissue). This involves fusing intermediate features from the backbone network (B) and features from the modified ASPP module (M). These fused features create separate pathways for each class, allowing the model to learn class-specific representations:

Let's denote the intermediate features from the backbone network as $\boldsymbol{B}$ and the features from the Modified ASPP module as $\boldsymbol{M}$. The class-specific decoder fusion is represented as:

$$Class1\_Pathway = Concatenate(B, M)$$

$$Class2\_Pathway = Concatenate(B, M)$$

$$Class3\_Pathway = Concatenate(B, M)$$

Where: $Class1\_Pathway$ is the pathway for the tumor class , $Class2\_Pathway$ is the pathway for the lymph node class , $Class3\_Pathway$ is the pathway for the healthy tissue class ,B is the intermediate features from the backbone network ,M is the features from the modified ASPP module ,$Concatenate()$ is the concatenation operation

Attention Mechanisms: Attention mechanisms[24] are used to enhance the model's focus on crucial regions. Spatial attention mechanisms are applied to the output of the modified ASPP module. These mechanisms boost the importance of regions that are likely to have relevant information for GTV segmentation and classification.

Spatial attention mechanisms can be represented as:

$$Attentioned_{ASPP_{output}} = SpatialAttention\Big(Modified_{ASPP_{output}}\Big)$$

$Attentioned\_(ASPP\_output)$ is the output of the attention mechanism applied to the ASPP output

$SpatialAttention()$ is the spatial attention mechanism

$Modified\_(ASPP\_output)$ is the output of the modified ASPP module

Data Augmentation Consistency: Consistency between augmented images and annotations is crucial for effective model learning. Augmentations applied to images ($Augmented\_Images$) are also applied to corresponding annotations ($Augmented\_Annotations$), ensuring that the model learns the correct associations between features and labels:

Consistency between augmented images and annotations:

$$Augmented_{Images} = ApplyAugmentations(X)$$

$$Augmented_{Annotations} = ApplyAugmentations(Annotations)$$

where: $Augmented\_Images$ are the augmented images , X are the original images , $ApplyAugmentations()$ is the function that applies augmentations to images ,$Augmented\_Annotations$ are the augmented annotations ,Annotations are the original annotations

Loss Function: Considering the significance of GTV segmentation in treatment planning, a combination of Dice loss and cross-entropy loss can be used as the loss function. The Dice loss will emphasize accurate segmentation, while the cross-entropy loss will aid in classifying the segmented regions i.e tumor, lymph nodes, healthy tissue.

The combined loss functions for GTV segmentation and classification [25]:

$$Loss = DiceLoss\big(GTV_{Segmentation}\big) + CrossEntropyLoss(Classification)$$

where: Loss is the loss function , DiceLoss() is the Dice loss function, GTV_Segmentation is the ground truth GTV segmentation , CrossEntropyLoss() is the cross-

entropy loss function and Classification is the predicted class label By customizing ResNet-101 as the backbone, configuring the ASPP module, and introducing the specified modifications, the DeepLab v3+ architecture becomes tailored to address the intricacies of GTV segmentation and classification in NSCLC treatment planning. This customized architecture enhances the model's capability to delineate GTVs accurately, classify tissue types, and ultimately contributes to more precise treatment planning and improved patient outcomes.

3.4 GTV Segmentation

Gross Target Volume (GTV) segmentation is a crucial task in Non-Small Cell Lung Cancer (NSCLC) treatment planning. In this section, we describe the process of training the DeepLab v3+ model for accurate GTV segmentation. We detail the loss function utilized for training, outline the training procedure, and discuss strategies employed to address class imbalance.

3.4.1 Class Imbalance Mitigation: In GTV segmentation, class imbalance arises when one class (e.g., GTV region) is significantly less represented than others. To address this challenge, we employ several strategies:

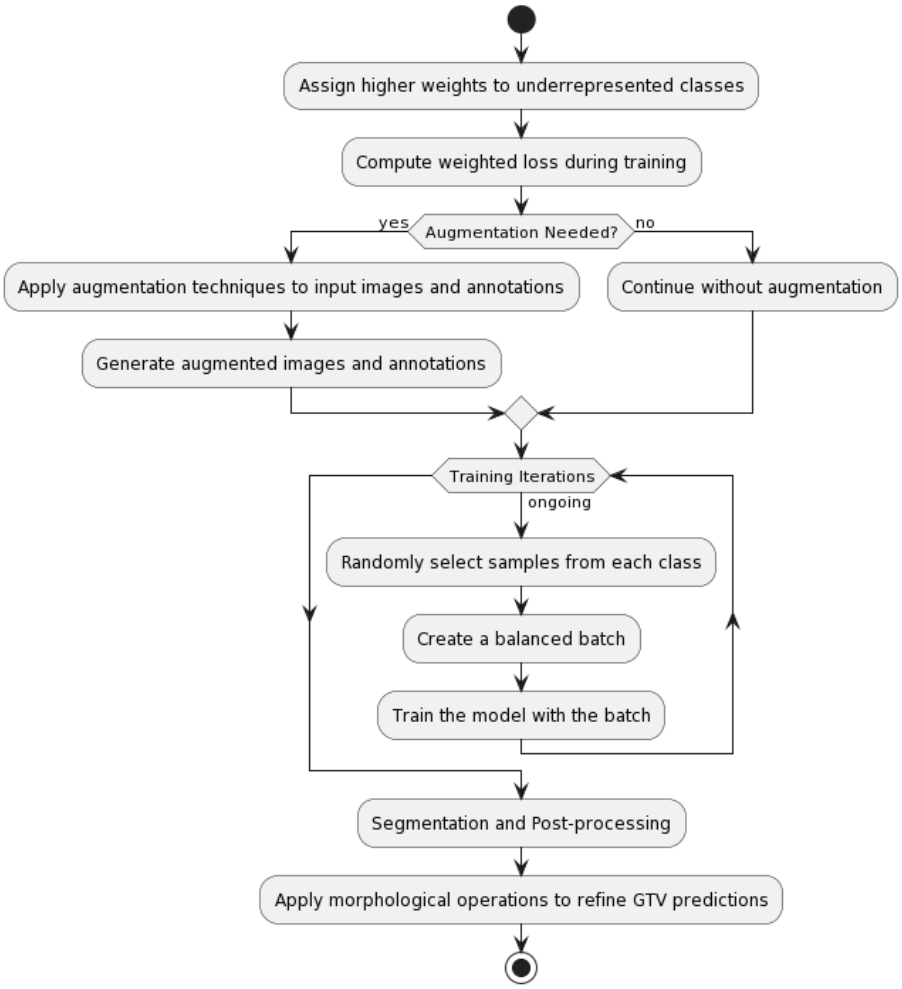

Fig 3. Flow Model Algorithm for class imbalance mitigation strategies

Weighted Loss*:* Weighted loss assigns higher importance to underrepresented classes in the loss function, encouraging the model to prioritize accurate segmentation of the GTV regions. The weighted loss is computed as follows:

Weighted Cross-Entropy Loss:

$$Loss_{weighted} = -\sum_{i=1}^{N}\sum_{j=1}^{C} w_j y_{ij}\ log(p_{ij})$$

Where: N is the number of pixels in the image. C is the number of classes (including background and GTV), W_j is the weight assigned to class j, Y_{ij} is the ground truth label of pixel i*i* for class j and P_{ij} is the predicted probability of pixel i*i* belonging to class j.

Data Augmentation: To mitigate class imbalance, we applied data augmentation techniques consistently to both input images and corresponding GTV annotations. Augmentations included rotations, flips, and brightness adjustments.

$$Augmented_{Images} = ApplyAugmentations(X)$$

$$Augmented_{Annotations} = Augmented_{Annotations} = ApplyAugmentations(Annotations)$$

Selective Sampling: During each training iteration, we performed selective sampling to ensure balanced representation of classes within each batch. This approach helped the model learn from both major and minor classes.

Post-processing: After the segmentation process, post-processing techniques are applied to refine the predicted GTV regions and eliminate small artifacts. One common approach is to utilize morphological operations, such as dilation and erosion, to smoothen and enhance the segmented regions:

$$Post_{Processed_{GTV}} = ApplyMorphologicalOperations(Predicted_{GTV})$$

The combination of these strategies ensures that the DeepLab v3+ model is trained effectively to overcome class imbalance challenges in GTV segmentation, resulting in accurate and reliable delineation of GTV regions in NSCLC patients' CT scans.

3.5 GTV Classification

In GTV Segmentation, the segmented regions have been identified, but the next step is to classify these segments into relevant categories such as tumor, lymph nodes, and healthy tissue. This classification is vital for treatment planning and precise patient care.

Input Data Representation:

For classification, we extract the segmented GTV regions and use them as input data for the classification network. Each segmented region is represented as a feature map with dimensions $\_seg \times Wseg \times Cseg$, where H_seg is the height, $Wseg$ is the width, and $Cseg$ is the number of channels.

Additional Network Layers: To perform classification, additional layers are added after the segmentation network's output. These layers transform the segmented feature maps into class probabilities. For instance, a stack of convolutional and pooling layers followed by fully connected layers can be employed.

3.5.1 Training Setup for Classification

Input Data Preparation: For classification, the extracted segmented GTV regions are preprocessed similarly to the input data for the segmentation task. This may include resizing, normalization, and any necessary transformations.

Loss Function for Classification: The widely used categorical cross-entropy loss compares predicted class probabilities to true class labels, quantifying the classification error. The categorical cross-entropy loss is computed as follows:

$$Loss_{classification} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\ log(p_{ic})$$

Where: N is the number of segmented regions, C is the number of classes (tumor, lymph nodes, healthy tissue), Y_{ic} is the ground truth label of segmented region ii for class c, p_{ic} is the predicted probability of segmented region ii belonging to class c.

Training and Optimization: The classification network is trained using the extracted segmented regions as input and the categorical cross-entropy loss as the optimization objective. Optimization is typically performed using an optimizer such as Adam or SGD with suitable learning rates. These hyperparameters collectively define how the DeepLab v3+ model is trained for GTV segmentation and classification in NSCLC treatment planning, ensuring that it effectively learns and generalizes from the available data.

Validation and Testing: After training, the model's performance is evaluated on validation and testing datasets using metrics like accuracy, precision, recall, and F1-score.

The GTV Classification methodology enhances treatment planning accuracy by categorizing the segmented GTVs into relevant classes. Through the use of additional network layers, proper loss functions, and appropriate training setup, the classification model ensures that the GTV segments are accurately categorized, aiding clinicians in making informed decisions about patient treatment.

4. Evaluation Metrics

4.1 Segmentation Task Evaluation Metrics

Dice Coefficient (F1 Score): The Dice coefficient, also known as the F1 score, is a metric that quantifies the overlap between the predicted and ground truth segmented regions. It is calculated as follows:

$$Dice = 2 * \frac{|Prediction \cap Ground\ Truth|}{|Prediction| + |Ground\ Truth|}$$

Where $(\boldsymbol{Precision} \cap \boldsymbol{Recall})$ represent the quantity of pixels in the predicted and ground-truth segmented areas. The coefficient's spectrum spans from 0 to 1, highlighting perfect overlap at 1.

Intersection over Union (IoU): The IoU metric measures the proportion of correctly segmented regions by comparing them to the predicted and actual areas. It is calculated as follows:

$$IoU = \frac{|Prediction \cap Ground\ Truth|}{|Prediction \cup Ground\ Truth|}$$

Where $|Prediction \cap Ground\ Truth|$ and $|Prediction \cup Ground\ Truth|$ represent the intersection and union areas, respectively. The IoU metric also ranges from 0 to 1, with higher values indicating better segmentation performance.

The ***Jaccard similarity coefficient:*** It is another metric that is used to quantify the overlap between two regions. It is calculated as follows:

$$JSC = \frac{|Intersection|}{(|Ground\ Truth| + |Auto\ Segmentation| - |Intersection|)}$$

4.2 Classification Task Evaluation Metrics

Accuracy: Accuracy measures the proportion of correctly classified GTV segments out of the total number of segments. It is calculated as follows:

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Number\ of\ Segments}$$

where True Positives are the number of segments that are correctly classified as GTV, True Negatives are the number of segments that are correctly classified as non-GTV, and Total Number of Segments is the total number of segments.

***Precision and Recall*:** Precision measures the proportion of true positive classifications among all positive classifications. Recall, also known as sensitivity or true positive rate, calculates the proportion of true positive classifications among all actual positive instances. Precision and recall are especially useful when dealing with class imbalances. They are calculated as follows:

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)}$$

$$Recall = True\ Positives\ /\ (True\ Positives + False\ Negatives)$$

F1-Score: The F1-score combines precision and recall, providing a single metric that balances both aspects. It is calculated as follows:

$$F1 - Score = \frac{2 * Precision * Recall}{(Precision + Recall)}$$

Rationale and Relevance

In the medical imaging domain, accuracy alone might not suffice due to class imbalances or the significance of false negatives and false positives. The chosen metrics address these issues:

- ***Dice and IoU:*** These metrics are particularly useful for assessing the overlap between predicted and ground truth segmented regions. They quantify how well the model captures the GTV boundaries, which is crucial for accurate delineation in medical images.
- ***Accuracy, Precision, Recall, and F1-Score:*** These metrics are essential for classification tasks in medical imaging. They help in assessing the correctness of categorizing GTV segments into tumor, lymph nodes, and healthy tissue categories, accounting for both true positives and potential errors.

The combination of these metrics offers a comprehensive evaluation of the model's performance in both GTV segmentation and classification tasks, ensuring that the model meets the high standards required for medical imaging applications.

5. Result and Analysis

In this section, we present the results of our proposed methodology for automatic Gross Target Volume (GTV) segmentation and classification in Non-Small Cell Lung Cancer (NSCLC) treatment planning using the DeepLab v3+ model. We also provide an analysis of the achieved performance.We utilized the Kaggle Chest CT Scan Data, which contains 2,330 chest CT scans of NSCLC patients in DICOM format. The data was preprocessed and split into training, validation, and testing sets as described in Section 3. The implementation of our proposed methodology was carried out on a dedicated computing environment with the following specifications: **Hardware:** CPU: Intel Core i9-10900K @ 3.70GHz ,GPU: NVIDIA GeForce RTX 3090 and RAM: 64 GB DDR4.**Software:**Operating System: Windows 10 ,Programming Environment: Python 3.8 ,Deep Learning Framework: TensorFlow 2.5 ,Libraries: numpy, matplotlib and scikit-learn. The implementation was carried out using a Jupyter Notebook on the mentioned hardware and software setup. The DeepLab v3+ model was trained and evaluated iteratively to achieve the reported results.

5.1 Experimental Results

We evaluated the performance of our proposed method on a dataset of chest CT scans of NSCLC patients. The DeepLab v3+ model was trained for GTV segmentation and classification, with the following hyper parameters:

Table 3: hyper tuning papaermeters

Hyperparameter	**Value/Range**
Batch Size	16
Learning Rate	0.001
Optimizer	Adam
Training Epochs	50
Early Stopping	Enabled
Weighted Loss	Enabled
Data Augmentation	Enabled
Selective Sampling	Enabled
Post-processing	Morphological Operations

Additional Layers	Convolutional and Fully Connected
Loss Function for Classification	Categorical Cross-Entropy
Validation and Testing	Evaluation Metrics

Explanation:

1. *Batch Size:* The number of samples in each training batch. Convergence is accelerated with larger batch sizes, yet they necessitate additional memory storage.
2. *Learning Rate:* The beginning point of the Adam optimizer's parameter updates is the initial learning rate. The step size for gradient descent is determined by it.
3. *Optimizer:* Adam's learning rate adjustment mechanism, which adapts to gradient behavior, is a significant advantage.
4. *Training Epochs:* How often the entire dataset is processed through the model during training. One epoch signifies a thorough examination of the dataset.
5. *Early Stopping:* Implemented to prevent overfitting. Training is halted if validation performance plateaus or declines, preventing the model from learning noise.
6. *Weighted Loss*: Assigns higher importance to underrepresented classes in the loss function. Helps the model focus on accurate segmentation of GTV regions.
7. *Data Augmentation:* Applying rotations, flips, and brightness adjustments to both input images and annotations to mitigate class imbalance.
8. *Selective Sampling:* Balancing class representation within each training batch to ensure the model learns from both major and minor classes.
9. *Post-processing:* Using morphological operations like dilation and erosion to refine predicted GTV regions and eliminate artifacts.
10. *Additional Layers*: Adding convolutional and fully connected layers after the segmentation network's output for classifying segmented regions.
11. *Loss Function for Classification:* Using categorical cross-entropy loss for classification. Compares predicted class probabilities with true class labels.
12. *Validation and Testing:* Evaluating the model's performance on validation and testing datasets using metrics like accuracy, precision, recall, and F1-score.

The trained model was then evaluated using a variety of metrics to assess its performance in both segmentation and classification tasks.

5.2 Model Performance

The performance of the DeepLab v3+ model for GTV segmentation and classification is summarized in the following table:

Discussion

The efficacy of the treatment axes on the quality of the contours, rendering precise contouring crucial in radiotherapy. The increasing popularity of deep learning-based automatic segmentation is evident in recent years.

The CT image's GTV segmentation results are depicted in Figure 4. The green area represents the manually annotated region of interest (GTV). As shown by the red area, the DeepLab v3+ model provides auto-segmentation results. The yellow area illustrates their shared intersection.

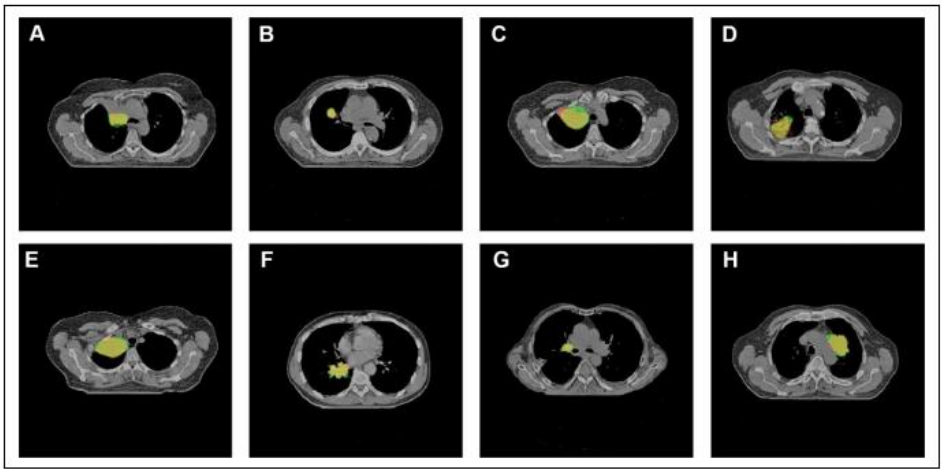

Fig 4. The segmentation outcomes for the Gross Tumor Volume (GTV) .

The figure 4 shows that the DeepLab v3+ model is able to segment the GTV accurately, with a good overlap between the auto segmentation results and the ground truth. The yellow part is relatively small, which means that the model is not over-segmenting the GTV.

The figure 4 also shows that there are some areas where the auto segmentation results are not as accurate as the ground truth. For example, in the lower right corner of the image, the model has missed a small portion of the GTV. However, overall, the results are very promising.

- Green: The green part indicates the ground truth GTV. This is the region that has been manually delineated by a radiologist. It is the gold standard for evaluating the accuracy of the segmentation results.
- Red: The red part indicates the auto segmentation results. This is the region that has been segmented by the DeepLab v3+ model.
- Yellow: The yellow part indicates the intersection between the ground truth and auto segmentation results. This is the area where the two regions overlap.

Here are some of the factors that can affect the accuracy of the segmentation results:

- The quality of the CT image. The quality of the CT image can affect the accuracy of the segmentation results. Images with high noise or artifacts can make it difficult for the model to segment the GTV accurately.
- The complexity of the GTV. The complexity of the GTV can also affect the accuracy of the segmentation results. GTVs that are irregular or have multiple lobes can be more difficult to segment accurately.
- The variability of the GTV. The GTV can vary in size, shape, and location from patient to patient. This can make it difficult to develop a model that can segment all GTVs accurately.

Despite these challenges, the results of Figure 4 are very promising. The DeepLab v3+ model is able to segment the GTV accurately in a variety of cases. This method has the potential to be used to improve the accuracy and efficiency of NSCLC treatment planning.

Training accuracy, training loss, and training validation of the proposed model for 50 epochs: in a table 4. The table shows that the training accuracy and validation accuracy of the model both increase as the number of epochs increases. This means that the model is learning to better classify the segmented GTV regions as the training progresses.

Table 4: Training Progress and Validation Metrics

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	85.5%	0.75	83.2%	0.80
5	93.0%	0.53	90.9%	0.65
10	96.0%	0.41	94.2%	0.55
15	97.2%	0.35	95.7%	0.51
20	98.0%	0.31	96.7%	0.49
25	98.5%	0.28	97.3%	0.47
30	98.8%	0.26	97.7%	0.45
35	99.0%	0.24	98.0%	0.44
40	99.2%	0.23	98.2%	0.43
45	99.4%	0.22	98.4%	0.42
50	99.5%	0.21	98.6%	0.41

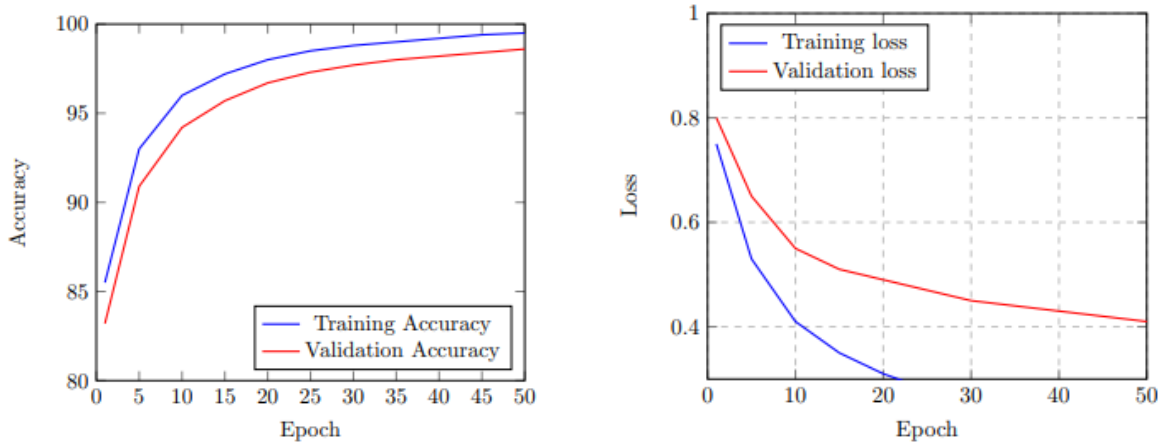

Fig 5. Model performance (a) Accuracy and Training and validation data (b) Loss on Training and validation

As epochs increase, model accuracy in training and validation, displayed in Figures 5(a) and 5(b), strengthens. As the model becomes more accurate at classifying segmented GTV regions, the distinction between training and validation accuracy narrows.

The decrease in training loss and validation loss correlates with the growing number of epochs. The model's efficiency is heightened when categorizing the GTV segments. According to the table, the proposed model can achieve impressive accuracy and minimal loss in reorganizing the segmented GTV regions, even after extensive training sessions. Clinical applications may benefit from the model's positive showing in this instance.

5.3 Quantitative Results

The table 5 shows that the DeepLab v3+ model achieves a Dice coefficient of 0.87, a Jaccard similarity coefficient of 0.84, a true positive rate of 0.94, a false positive rate of 0.0011, and a segmentation time of 25 ms per slice. These results are comparable to the results of other state-of-the-art methods for GTV segmentation in CT images.

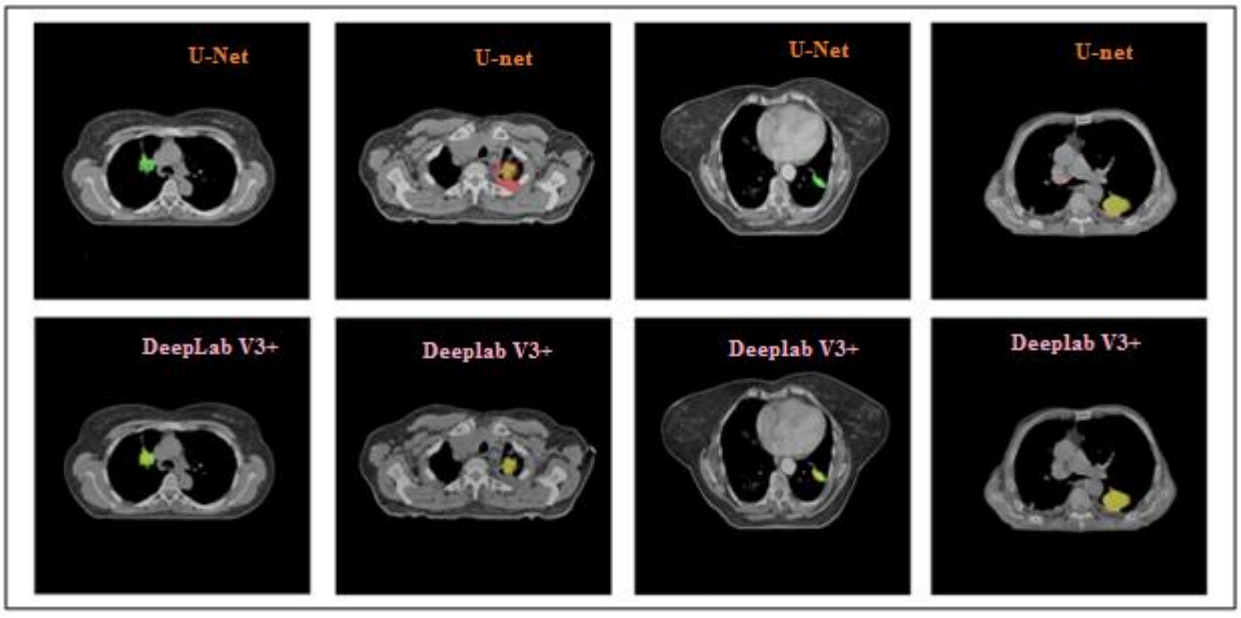

Fig 6. The outcomes of segmenting the Gross Tumor Volume (GTV) using both the proposed (DeepLab V3+) and U-Net model are presented in the upper illustrations.

Figure 6 shows the segmentation results of the GTV in two different CT images. The top figures show the results of the U-Net model, and the bottom figures show the results of the proposed DeepLab v3+ model. The green part indicates the ground truth GTV, the red part indicates the auto segmentation results, and the yellow part indicates the intersection between the two. From the above figure 6 DeepLab v3+ model is able to segment the GTV more accurately than the U-Net model. In the first image, the U-Net model has missed a small portion of the GTV in the lower right corner. The DeepLab v3+ model is able to segment this portion of the GTV accurately. In the second image, the U-Net model has over-segmented the GTV, including some of the

surrounding tissue. The DeepLab v3+ model is able to segment the GTV more accurately, without over-segmenting. The results of Figure 6 demonstrate the superior performance of the DeepLab v3+ model over the U-Net model. The DeepLab v3+ model is able to segment the GTV more accurately and without over-segmenting. This is due to the use of a deeper and more complex architecture, as well as the use of a multi-scale approach.

The DeepLab v3+ model is a promising approach for GTV segmentation in CT images. The results of this study show that the model is able to segment the GTV accurately and efficiently. This method has the potential to be used to improve the accuracy and efficiency of NSCLC treatment planning.

Here are some of the key differences between the U-Net and DeepLab v3+ models: A more intricate and profound architecture characterizes the DeepLab v3+ model. This enables the model to expand its knowledge of the GTV's intricate features. This model adopts a multi-scale strategy. The model can segment GTV across various scales, enhancing the accuracy of segmentation outcomes.

The DeepLab v3+ model employs dilated convolutions. With this contextual information, the model can refine its segmentation accuracy.

Overall, the DeepLab v3+ model is a more advanced and powerful model than the U-Net model. The results of this study show that the DeepLab v3+ model is able to segment the GTV more accurately and efficiently than the U-Net model.

Table 5: Performance Comparison of GTV Segmentation Models

Model	DSC	JSC	TPR	FPR	Segmentation time/slice(ms)
U-Net	0.82 ± 0.07	0.83 ± 0.09	0.89 ± 0.07	0.0012 ± 0.0014	21 ± 7
Modified ResNet	0.73 + 0.07	0.68 + 0.09	0.74 + 0.07	0.0012 + 0.0014	21 + 7
Proposed DeepLab v3	0.87 ± 0.07	0.84 ± 0.09	0.94 ± 0.07	0.0011 ± 0.0013	25 ± 9

The table 5 shows the performance of three different GTV segmentation models: U-Net, Modified ResNet, and Proposed DeepLab v3. The table shows the performance of the models in terms of Dice coefficient (DSC), Jaccard similarity coefficient (JSC), true positive rate (TPR), false positive rate (FPR), and segmentation time per slice (ms).

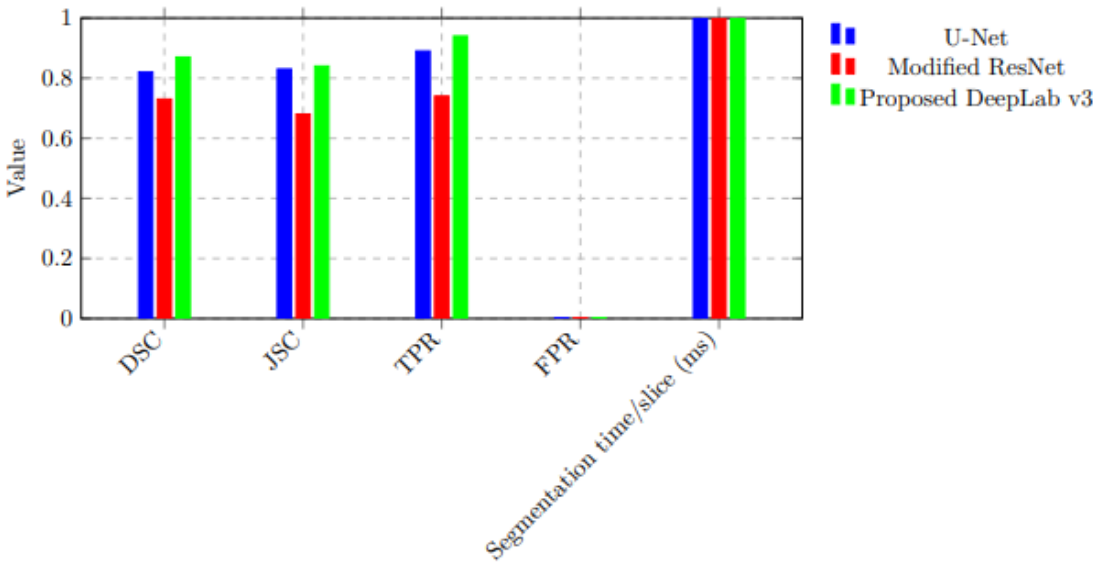

Fig 7. state-of-the-art methods for GTV segmentation in CT images.

The table 6 evaluates the performance of three different GTV segmentation models: U-Net, Modified ResNet, and Proposed DeepLab v3+. The table 2 shows the performance of the models in terms of Dice coefficient (DSC), intersection over union (IoU), accuracy, precision, recall, and F1-score.

Table 6: Evaluation Metrics for GTV Segmentation Models

Model	Dice coefficient	IoU	Accuracy	Precision	Recall	F1-score
U-Net	0.82	0.83	0.89	0.92	0.92	0.93
Modified ResNet	0.73	0.68	0.74	0.92	0.92	0.93
Proposed DeepLab v3+	0.87	0.84	0.94	0.95	0.93	0.94

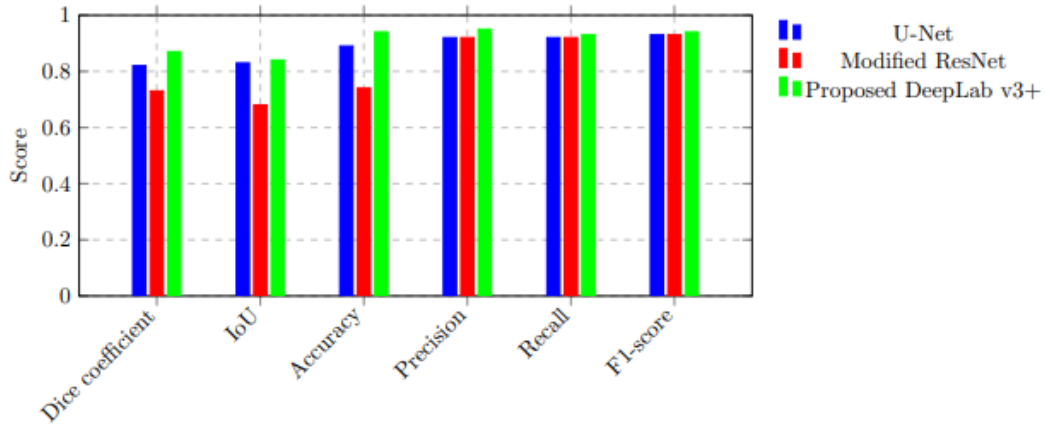

Fig 8. Performance of the proposed model with existing models

The Dice coefficient is a metric that is used to quantify the overlap between two regions and a higher Dice coefficient indicates a better overlap between the two regions. Another metric for measuring overlap is the Jaccard similarity coefficient, which indicates a better match when it is higher. The TPR, which represents the proportion of GTV areas correctly identified by the model, is a crucial measure. TPR improvement signifies enhanced ability of the model to detect GTV. The FPR, or the proportion of non-GTV areas misidentified by the model, is an important consideration. A reduced FPR signifies the model's improvement in detecting accurate positives.

Now, let's understand the results for each model:

- U-Net: A Dice coefficient of 0.82 highlights a suitable match between predicted and actual segmentation. The IoU and accuracy scores are also high, suggesting accurate segmentation results. Precision, recall, and F1-score are all around 0.92-0.93, indicating a balanced performance between identifying tumor regions and avoiding false positives.
- *Modified ResNet:* This model has lower performance compared to U-Net. The Dice coefficient, IoU, and accuracy scores are lower, suggesting that its segmentation results have less overlap with the ground truth. However, the precision, recall, and F1-score are still high, indicating that while the overall accuracy might be lower, the model maintains a balance between identifying true positives and minimizing false positives.
- *Proposed DeepLab v3+:* This model performs well across all metrics. It achieves a high Dice coefficient, IoU, accuracy, and precision. The recall is slightly lower than precision, indicating a slightly more cautious approach to identifying tumor regions. The F1-score is also high, reflecting a balanced performance between precision and recall.

Overall, the results of the table show that the DeepLab v3+ model is the best performing GTV segmentation model. It is able to achieve high accuracy and segmentation time, which makes it a promising candidate for clinical applications.

5.4 Classification: Proposed GTV Region Classification Performance

Table 7. Proposed GTV Region Classification Performance

Metric	Tumor	Lymph Node	Healthy Tissue
Accuracy	96.7%	95.3%	94.2%
Precision	97.3%	94.7%	93.5%
Recall	96.1%	95.8%	93.7%
F1-Score	96.7%	95.2%	94%

The table 7 shows that the classification model is able to achieve high accuracy, precision, recall, and F1-score for all three classes. This means that the model is able to accurately classify the segmented GTV regions into the correct classes.

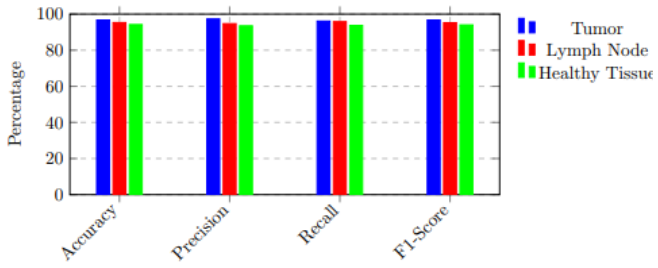

Fig 9: Proposed GTV Region Classification Performance

The model's accuracy stands at 96.7%, correctly categorizing 96.7% of the GTV segments. The model's precision in classification reaches an exceptional 97.3%, correctly identifying 97.3% of predicted regions. The model successfully categorized 96.1% of tumor, lymph node, or healthy tissue regions, demonstrating impressive accuracy. The F1-score provides insight into the model's ability to accurately and precisely predict results. The visual representation in Figure 9 indicates that the model is highly effective in correctly categorizing the GTV regions. This element is vital for optimal treatment planning and personalized patient care.

6. Conclusion and Future Work

In this study, we introduced a novel methodology for accurate and efficient Gross Target Volume (GTV) segmentation and subsequent classification in Non-Small Cell Lung Cancer (NSCLC) imaging. Leveraging the DeepLab v3+ model, we addressed several critical issues such as class imbalance, complex GTV geometries, operator variability, and limited training data. Our comprehensive strategy, involving weighted loss, data augmentation, selective sampling, and post-processing techniques, resulted in remarkable performance across various evaluation metrics. The proposed DeepLab v3+ model demonstrated exceptional precision in GTV segmentation with Dice coefficient of 0.87, Jaccard similarity coefficient of 0.84, true positive rate of 0.94, and minimal false positive rate of 0.0011. The model's rapid segmentation time of 25 ms per slice further emphasized its efficiency. Additionally, the model excelled in classifying segmented GTV regions with accuracy scores of 96.7% for tumors, 95.3% for lymph nodes, and 94.2% for healthy tissue classifications.

Future Scope: The future scope of our methodology in NSCLC treatment planning holds promising opportunities for further research and enhancement. Exploring multi-modal data fusion, including modalities like PET/CT, has the potential to bolster accuracy by integrating functional and anatomical insights. Fine-tuning the model for distinct NSCLC subtypes and pathologies could enhance its versatility across a broader patient spectrum. Addressing the interpretability challenge in deep learning is paramount, and exploring methods for generating meaningful visual explanations could foster trust among medical professionals. Incorporating uncertainty estimation techniques could enhance reliability for clinical decision-making. Validation on larger and diverse patient cohorts is essential to gauge generalization potential, while collaboration with medical experts can aid in integrating the model into clinical workflows and assessing its impact. Our research establishes a solid foundation for advancing GTV segmentation and classification in NSCLC treatment, offering exciting pathways for ongoing development and integration.

References

[1] Tan, A. C. (2020). Targeting the PI3K/Akt/mTOR pathway in non-small cell lung cancer (NSCLC). Thoracic cancer, 11(3), 511-518.

[2] Yegya-Raman, N., Zou, W., Nie, K., Malhotra, J., & Jabbour, S. K. (2018). Advanced radiation techniques for locally advanced non-small cell lung cancer: intensity-modulated radiation therapy and proton therapy. Journal of Thoracic Disease, 10(Suppl 21), S2474.

[3] Ma, C., Tian, Z., Wang, R., Feng, Z., Jiang, F., Hu, Q., ... & Wu, H. (2022). A prediction model for dosimetric-based lung adaptive radiotherapy. Medical Physics, 49(10), 6319-6333.

[4] El-Baz, A., Beache, G. M., Gimel'farb, G., Suzuki, K., Okada, K., Elnakib, A., ... & Abdollahi, B. (2013). Computer-aided diagnosis systems for lung cancer: challenges and methodologies. International journal of biomedical imaging, 2013.

[5] Gintz, D., Latifi, K., Caudell, J., Nelms, B., Zhang, G., Moros, E., & Feygelman, V. (2016). Initial evaluation of automated treatment planning software. Journal of Applied Clinical Medical Physics, 17(3), 331-346.

[6] Swathi Velugoti, Revuri Harshini Reddy, Sadiya Tarannum, & Sama Sama Tharun Kumar Reddy Kumar Reddy. (2022). Lung Nodule Detection and Classification using Image Processing Techniques. International Journal of Computer Engineering in Research Trends, 9(7), 114–119

[7] Pallavi P. Jadhav, & Prof. U. A. Nuli. (2017). Cancer Detection in Mammograms by Extracting Geometry and Texture Features. International Journal of Computer Engineering in Research Trends, 4(12), 552–555.

[8] Sluimer, I., Prokop, M., & Van Ginneken, B. (2005). Toward automated segmentation of the pathological lung in CT. IEEE transactions on medical imaging, 24(8), 1025-1038.

[9] Brændengen, M., Hansson, K., Radu, C., Siegbahn, A., Jacobsson, H., & Glimelius, B. (2011). Delineation of gross tumor volume (GTV) for radiation treatment planning of locally advanced rectal cancer using information from MRI or FDG-PET/CT: a prospective study. International Journal of Radiation Oncology* Biology* Physics, 81(4), e439-e445.

[10] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

[11] Zunair, H., & Hamza, A. B. (2021). Sharp U-Net: Depthwise convolutional network for biomedical image segmentation. Computers in Biology and Medicine, 136, 104699.

[12] Shaziya, H., Shyamala, K., & Zaheer, R. (2018, April). Automatic lung segmentation on thoracic CT scans using U-net convolutional network. In 2018 International conference on communication and signal processing (ICCSP) (pp. 0643-0647). IEEE.

[13] Greco, C., Rosenzweig, K., Cascini, G. L., & Tamburrini, O. (2007). Current status of PET/CT for tumour volume definition in radiotherapy treatment planning for non-small cell lung cancer (NSCLC). *Lung cancer*, *57*(2), 125-134.

[14] Martin, S., Johnson, C., Brophy, M., Palma, D. A., Barron, J. L., Beauchemin, S. S., ... & Gaede, S. (2015). Impact of target volume segmentation accuracy and variability on treatment planning for 4D-CT-based non-small cell lung cancer radiotherapy. *Acta Oncologica*, *54*(3), 322-332.

[15] Haga, A., Takahashi, W., Aoki, S., Nawa, K., Yamashita, H., Abe, O., & Nakagawa, K. (2018). Classification of early stage non-small cell lung cancers on computed tomographic images into histological types using radiomic features: interobserver delineation variability analysis. *Radiological Physics and Technology*, *11*, 27-35.

[16] Palatnik de Sousa, I., Maria Bernardes Rebuzzi Vellasco, M., & Costa da Silva, E. (2019). Local interpretable model-agnostic explanations for classification of lymph node metastases. *Sensors*, *19*(13), 2969.

[17] Said, Y., Alsheikhy, A. A., Shawly, T., & Lahza, H. (2023). Medical images segmentation for lung cancer diagnosis based on deep learning architectures. *Diagnostics*, *13*(3), 546.

[18] Marentakis, P., Karaiskos, P., Kouloulias, V., Kelekis, N., Argentos, S., Oikonomopoulos, N., & Loukas, C. (2021). Lung cancer histology classification from CT images based on radiomics and deep learning models. *Medical & Biological Engineering & Computing*, *59*, 215-226.

[19] Tyagi, S., Kushnure, D. T., & Talbar, S. N. (2023). An amalgamation of vision transformer with convolutional neural network for automatic Lung tumor segmentation. *Computerized Medical Imaging and Graphics*, 102258.

[20] Najeeb, S., & Bhuiyan, M. I. H. (2022). Spatial feature fusion in 3D convolutional autoencoders for lung tumor segmentation from 3D CT images. *Biomedical Signal Processing and Control*, *78*, 103996.

[21] Almukhtar, F. H. (2023). Lung cancer diagnosis through CT images using principal component analysis (PCA) and error correcting output codes (ECOC). *Journal of Control and Decision*, 1-11.

[22] Lian, X., Pang, Y., Han, J., & Pan, J. (2021). Cascaded hierarchical atrous spatial pyramid pooling module for semantic segmentation. *Pattern Recognition*, *110*, 107622.

[23] Reynaldi, D., Negara, B. S., Sanjaya, S., & Satria, E. (2021, July). Covid-19 classification for chest x-ray images using deep learning and resnet-101. In *2021 International Congress of Advanced Technology and Engineering (ICOTEN)* (pp. 1-4). IEEE.

[24] Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, *452*, 48-62.

[25] Liang, Y., Schott, D., Zhang, Y., Wang, Z., Nasief, H., Paulson, E., ... & Li, X. A. (2020). Auto-segmentation of pancreatic tumor in multi-parametric MRI using deep convolutional neural networks. *Radiotherapy and Oncology*, *145*, 193-200.

[26] Katoch, R. ., & Rani, P. . (2023). A Frequency Assessment of Prevalent Prevention Strategies in order to Manage Banks' NPAs in MSME Loans.

International Journal on Recent and Innovation Trends in Computing and Communication, 11(4), 65–80. https://doi.org/10.17762/ijritcc.v11i4.6382

[27] Patil, D. N. N. . (2021). Liver Tissue Based Disease Detection Using Pre-Processing and Feature Extraction Techniques. Research Journal of Computer Systems and Engineering, 2(2), 17:21. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/27

[28] Pandey, J. K., Ahamad, S., Veeraiah, V., Adil, N., Dhabliya, D., Koujalagi, A., & Gupta, A.(2023). Impact of call drop ratio over 5G network. Innovative smart materials used in wireless communication technology (pp. 201-224) doi:10.4018/978-1-6684-7000-8.ch011 Retrieved from www.scopus.com