JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AUTOMATED LUNG TUMOR SEGMENTATION USING U-NET & U-NET+

[1]Mr. A. Rama Krishna Prasad, [2]G. K. Pavan Teja, [3]A. Bhargavi, [4]V. Narayana Reddy, [5]S. Charan Teja

[1]Assistant Professor, [2,3,4,5]UG Scholars
[1]Department of Electronics and Communication Engineering
[1]N.B.K.R Institute of Science and Technology, Vidyanagar, Andhra Pradesh

Abstract— Lung tumor segmentation plays a critical role in the early detection and treatment of lung cancer. This project focuses on automating the segmentation process of lung tumors using deep learning techniques, specifically leveraging the power of U-Net and U-Net+ architectures. The dataset used for this study is sourced from Kaggle, containing pre-processed CT scan images of lung tumors. The segmentation model aims to accurately identify and localize the tumor regions within the lung, which is essential for diagnosis and treatment planning. U-Net, a widely used convolutional neural network architecture, is designed to handle medical image segmentation tasks by capturing both local and global context through its encoder-decoder structure. To further improve performance, we utilize U-Net+, an advanced version of U-Net, which incorporates additional modifications to enhance segmentation accuracy and deal with challenges like small tumor regions and unclear boundaries. The model training is performed on the pre-processed dataset, and the results are evaluated based on various performance metrics, such as Intersection over Union (IoU), Dice Coefficient, and accuracy. The outcomes of this study aim to provide a robust tool for radiologists to assist in tumor localization and enhance the efficiency of lung cancer diagnosis.

Index Terms— Lung tumor segmentation, U-Net, U-Net+, deep learning, medical image processing, CT scan, tumor localization, image dataset, convolutional neural network, early detection, diagnosis.

I. INTRODUCTION

Lung cancer is one of the leading causes of cancer-related mortality worldwide, accounting for a significant number of deaths each year. Early detection and accurate localization of lung tumors are crucial for improving patient survival rates and enabling effective treatment planning. Computed Tomography (CT) imaging is widely used for diagnosing lung abnormalities; however, manual analysis of CT scans by radiologists is time-consuming, subjective, and prone to inter-observer variability, particularly when tumors are small or exhibit irregular boundaries.

With the advancement of Artificial Intelligence (AI) and Deep Learning (DL), automated medical image analysis has gained significant attention. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in image segmentation tasks by automatically learning hierarchical features from medical images. Among various architectures, U-Net has emerged as a powerful model for biomedical image segmentation due to its encoder–decoder structure and skip connections, which preserve spatial information and improve localization accuracy.

Despite its effectiveness, traditional U-Net faces challenges in handling complex tumor structures and reducing the semantic gap between encoder and decoder features. To overcome these limitations, advanced architectures such as U-Net++ have been introduced, incorporating nested skip connections and deep supervision to enhance segmentation performance. These improvements enable better detection of small and poorly defined tumor regions.

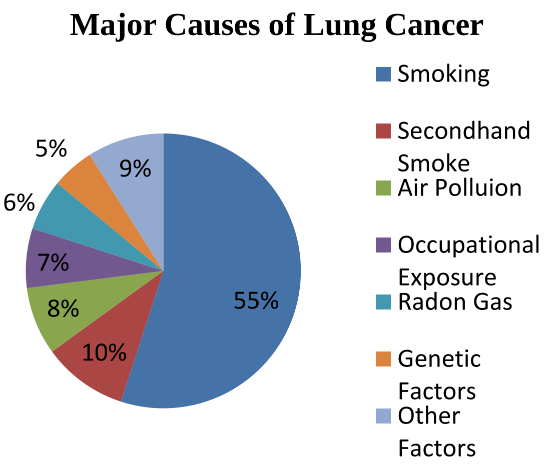


Figure 1: Major Causes of Lung Cancer

The pie chart illustrates the major causes of lung cancer, where smoking accounts for the largest proportion (55%), making it the primary risk factor. Secondhand smoke contributes 10%, followed by air pollution (8%) and occupational exposure (7%). Radon gas (6%) and genetic factors (5%) also play significant roles, while other factors collectively account for 9%. This distribution highlights that most lung cancer cases are linked to preventable environmental and lifestyle factors.

This paper proposes an automated lung tumor segmentation system using U-Net and U-Net++ architectures. The system processes CT scan images to accurately segment tumor regions and assist radiologists in diagnosis. The performance of the proposed models is evaluated using standard metrics such as Dice Coefficient, Intersection over Union (IoU), and accuracy, demonstrating improved segmentation efficiency and reliability.

II. LITERATURE REVIEW

The rapid advancement of deep learning techniques has significantly transformed the field of medical image analysis, particularly in lung tumor detection and segmentation. Accurate segmentation of pulmonary nodules from CT images is a challenging task due to variations in tumor size, shape, intensity, and unclear boundaries. Numerous research efforts have focused on developing robust models to overcome these challenges and improve diagnostic accuracy.

An extensive overview of deep neural network-based pulmonary nodule segmentation techniques is presented in [1]. The study systematically analyzes various architectures, including Convolutional Neural Networks (CNNs), Fully Convolutional Networks (FCNs), and U-Net-based models. It highlights that deep learning approaches outperform traditional image processing methods by automatically extracting hierarchical features from CT images. However, the paper also identifies challenges such as the need for large annotated datasets and difficulties in segmenting nodules with low contrast or irregular shapes.

To address these challenges, a cascaded multi-stage framework is proposed in [2] for the automatic detection and segmentation of pulmonary nodules. The framework integrates multiple deep learning models in sequential stages, allowing the system to refine predictions progressively. This approach enhances detection accuracy and reduces false positives, particularly in complex clinical scenarios. Moreover, the study emphasizes the importance of designing cost-effective and computationally efficient solutions for deployment in resource-limited healthcare environments.

In [3], the authors introduce an advanced architecture called AWEU-Net (Attention-Aware Weight Excitation U-Net), which incorporates attention mechanisms into the traditional U-Net model. The attention modules enable the network to focus on the most relevant regions of the image, thereby improving segmentation performance for small and ambiguous tumor regions. The study demonstrates that integrating attention mechanisms significantly enhances feature representation and leads to improved accuracy compared to standard U-Net architectures.

Further improvements in segmentation performance are explored in [4], where a multiscale dense residual neural network is proposed. This model leverages multiscale feature extraction to capture both fine-grained details and global contextual information. The integration of dense connections and residual learning helps in mitigating the vanishing gradient problem and improves information flow across the network. As a result, the model achieves higher robustness and accuracy, especially in cases involving low-contrast CT images and complex tumor structures.

A comprehensive systematic review is presented in [5], which examines the role of deep learning in pulmonary nodule detection and segmentation. The review covers a wide range of architectures, including CNNs, U-Net, and hybrid models, and evaluates their performance across different datasets and clinical conditions. The findings indicate that deep learning models have achieved remarkable accuracy, often surpassing traditional machine learning approaches. However, the study also highlights limitations such as overfitting, lack of interpretability, and dependency on large-scale annotated datasets, suggesting the need for more generalized and explainable models.

In addition to segmentation-focused approaches, object detection-based frameworks have also been explored. In [6], a Mask R-CNN-based model is proposed for the detection and segmentation of pulmonary nodules, along with 3D visualization capabilities. This approach combines object detection and instance segmentation, allowing for precise localization and detailed analysis of tumor structures. The inclusion of 3D visualization enhances clinical applicability, particularly in surgical planning and disease monitoring. Despite its advantages, the model requires high computational resources and large datasets for optimal performance.

Overall, the literature indicates that while significant progress has been made in lung tumor segmentation using deep learning techniques, several challenges remain. These include handling small and irregular tumor regions, reducing false positives, and improving generalization across diverse datasets. Advanced architectures such as U-Net++ aim to address these limitations by enhancing feature fusion and reducing the semantic gap between encoder and decoder layers. Therefore, the proposed work focuses on leveraging U-Net and U-Net++ to achieve improved segmentation accuracy and reliability in lung tumor detection.

III. PROPOSED SYSTEM

The proposed method focuses on developing an automated system for lung tumor segmentation from CT scan images using deep learning techniques, specifically U-Net and U-Net++ architectures. The aim is to improve segmentation accuracy while reducing the dependency on manual analysis by radiologists.

The overall workflow of the proposed system consists of several stages, including data acquisition, preprocessing, model training, segmentation, and evaluation.

Initially, lung CT scan images are collected from publicly available datasets. These images are then subjected to preprocessing steps such as resizing, normalization, and noise reduction to enhance image quality and ensure uniformity. The preprocessed images are then fed into the U-Net model, which serves as the baseline architecture. U-Net follows an encoder–decoder structure, where the encoder extracts important features from the input image, and the decoder reconstructs the segmented output. Skip connections are used to preserve spatial information and improve localization accuracy.

To further enhance segmentation performance, the U-Net++ architecture is employed. U-Net++ introduces nested and dense skip connections that reduce the semantic gap between encoder and decoder feature maps. This allows the model to capture fine details and improves segmentation accuracy, especially for small and irregular tumor regions.

During the training phase, both models are trained using annotated CT images. The training process utilizes optimization techniques such as the Adam optimizer and loss functions like Dice Loss and Binary Cross-Entropy to handle class imbalance and improve segmentation quality.

After training, the models generate segmentation masks that highlight tumor regions in the CT images. Post-processing techniques are applied to remove noise and refine the segmented boundaries, resulting in more accurate outputs.

Finally, the performance of the models is evaluated using metrics such as Dice Coefficient, Intersection over Union (IoU), and accuracy. The results demonstrate that U-Net++

provides better segmentation performance compared to the standard U-Net model.

IV. METHODOLOGY

The proposed methodology focuses on automated lung tumor segmentation using deep learning architectures, specifically U-Net and U-Net++. These models are designed to perform pixel-wise classification of CT scan images, enabling accurate identification of tumor regions.

1. U-Net

U-Net is a convolutional neural network (CNN) architecture specifically designed for biomedical image segmentation. It is a fully convolutional network that performs pixel-level predictions, making it highly effective for segmenting lung tumors from CT scan images.

A. Architecture Overview

The U-Net architecture consists of two primary components: the contracting path (encoder) and the expansive path (decoder). The encoder captures contextual information from the input image, while the decoder ensures precise localization for accurate segmentation.

B. Contracting Path (Encoder):

The encoder is composed of repeated applications of convolutional layers followed by ReLU activation and max-pooling operations. This process progressively reduces the spatial dimensions of the image while increasing the depth of feature maps. As a result, the network learns both low-level features (such as edges and textures) and high-level features (such as shapes and boundaries), which are essential for tumor detection.

C. Bottleneck:

At the deepest level of the network lies the bottleneck layer, where the feature maps have minimal spatial dimensions but maximum depth. This layer captures the most abstract representation of the input image and serves as a bridge between the encoder and decoder.

D. Expansive Path (Decoder):

The decoder reconstructs the segmentation map by progressively increasing the spatial resolution using upsampling or transposed convolution operations. At each stage, the upsampled feature maps are concatenated with corresponding feature maps from the encoder via skip connections, enabling the recovery of fine-grained spatial details.

E. Skip Connections:

Skip connections are a key feature of U-Net, linking corresponding encoder and decoder layers. These connections allow the transfer of low-level spatial information directly to the decoder, which significantly improves localization accuracy and helps in detecting small or unclear tumor boundaries.

F. Final Layer:

The final layer applies a 1×1 convolution to generate the segmentation map with the same spatial dimensions as the input image. Each pixel is classified as either tumor or non-tumor (background), resulting in a binary segmentation output.

G. Training and Loss Function:

U-Net is trained using loss functions such as Binary Cross-Entropy Loss and Dice Loss. Dice Loss is particularly effective for medical imaging tasks, as it addresses class imbalance by focusing on the overlap between predicted and ground truth regions.

Overall, U-Net effectively combines local and global contextual information, making it a powerful baseline model for lung tumor segmentation.

2. U-Net+

U-Net++ is an advanced extension of the U-Net architecture designed to improve segmentation accuracy, particularly for complex and small tumor regions. It enhances the original architecture by introducing dense connectivity and improved feature fusion mechanisms.

A. Attention Mechanism:

One of the main improvements in U-Net+ is the introduction of attention mechanisms. Attention modules help the network focus on the most relevant parts of the input image by assigning higher weights to important regions (such as tumor boundaries) and suppressing less important ones. This helps in segmenting small or weakly defined tumors that might be overlooked in traditional U-Net.

B. Nested Skip Pathways:

Unlike U-Net, which uses direct skip connections, U-Net++ introduces nested and dense skip pathways between encoder and decoder layers. These connections help reduce the semantic gap between feature maps, enabling better feature alignment and improving segmentation performance.

C. Deep Supervision:

U-Net+ also incorporates deep supervision, which provides auxiliary supervision signals at intermediate layers during training. By applying loss functions at different depths of the network, the model can learn better feature representations at various stages of the encoding process, leading to improved performance in segmentation tasks. This multi-level supervision helps the model generalize better and avoids overfitting.

D. Dense Connections:

Dense connections between different layers of the encoder and decoder help improve information flow throughout the network. These connections allow the network to efficiently utilize feature maps from previous layers, enabling better feature reuse and strengthening the overall representation of the image. This is particularly useful in cases where the tumor has complex or irregular shapes.

E. Improved Decoder:

The decoder in U-Net++ is more sophisticated compared to U-Net, as it combines multi-scale features from different levels of the network. This allows for more accurate reconstruction of the segmentation map and better delineation of tumor boundaries.

F. Training and Loss Function:

Similar to U-Net, U-Net++ is trained using Dice Loss and Binary Cross-Entropy Loss. However, due to deep supervision, multiple loss functions may be applied at different layers, leading to improved optimization and segmentation accuracy.

V. SYSTEM ARCHITECTURE

The system architecture represents the overall workflow of the proposed Automated Lung Tumor segmentation using U-Net & U-Net+

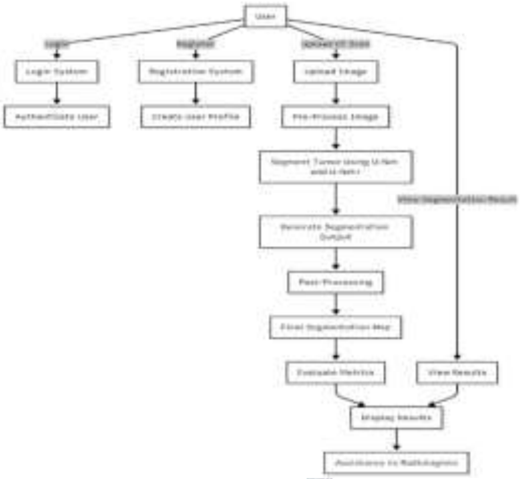


Figure 2: Workflow for CT Scan Tumor Segmentation Using U-Net & U-Net+ Models

A. User Actions

The user enters credentials, which go through the Login System to Authenticate User. New users create an account via the Registration System, which stores their User Profile. The main workflow begins when a user uploads a CT scan image.

B. Image Handling:

The CT scan is received by the system. Enhancements like noise reduction, normalization, or resizing are applied to prepare the scan for segmentation.

C. Tumor Segmentation:

Deep learning models (U-Net and its improved variant U-Net+) are applied to detect and outline tumor regions. The raw segmentation mask is produced, showing detected tumor boundaries.

D. Post-Processing:

Refinements such as smoothing, removing false positives, or morphological operations are applied. A clean, clinically usable tumor map is generated.

E. Evaluation & Results:

Performance is assessed using metrics like Dice coefficient, sensitivity, specificity, etc. The user can check the segmentation outcome. Results are presented visually (overlay on CT scan, charts, or tables). The final output supports radiologists in diagnosis and treatment planning.

The below architecture illustrates the integration of user interaction, web technologies, databases, and deep learning models for automated tumor segmentation using U-Net and U-Net+.

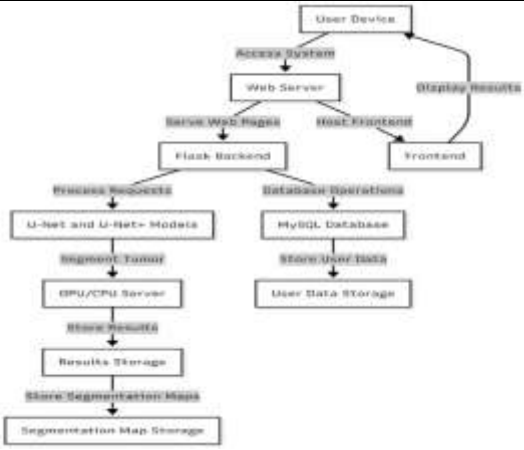


Figure 3: System Architecture for Automated Tumor Segmentation Using U-Net and U-Net++ Models

A. User Device:

The user device serves as the entry point to the system, enabling radiologists or medical staff to interact with the segmentation platform. Through a web interface, users can log in, upload CT scans, and view results. This block emphasizes accessibility, ensuring that the system can be operated remotely without requiring specialized hardware on the user's side.

B. **Web Server:**

The web server acts as the communication hub between the user device and the backend. It hosts the frontend interface, delivers web pages, and manages requests. By serving as the gateway, it ensures smooth data transfer and provides a secure environment for handling sensitive medical information.

C. Frontend:

The frontend is the user-facing component of the system. It displays segmentation results, evaluation metrics, and visual overlays of tumor regions on CT scans. Designed for clarity and usability, the frontend ensures that radiologists can easily interpret outputs and integrate them into clinical decision-making.

D. Backend (Flask Framework):

The backend, implemented using the Flask framework, processes incoming requests from the frontend. It coordinates tasks such as invoking U-Net and U-Net+ models for segmentation, handling database operations, and managing storage systems. This block represents the core logic of the system, bridging user interaction with computational processing.

E. Model Servers (U-Net & U-Net+)

The model servers host the deep learning models—U-Net and U-Net+—on GPU/CPU infrastructure. These models perform the actual tumor segmentation by analyzing CT scan images and generating segmentation masks. Their deployment on dedicated servers ensures computational efficiency and scalability for clinical workloads.

F. Results Storage:

Results storage maintains the outputs generated by the segmentation models. This includes raw segmentation masks, processed tumor maps, and evaluation metrics. By storing results systematically, the system allows users to retrieve past analyses and supports longitudinal studies of patient data.

G. Segmentation Map Storage:

Segmentation map storage is dedicated to preserving the final, post-processed tumor maps. These maps are clinically validated outputs that radiologists can directly use for diagnosis and treatment planning. This block ensures that refined, reliable data is available for medical decision support.

H. User Data Storage (MySQL Database):

The MySQL database manages user-related information such as login credentials, registration details, and profile metadata. It ensures secure authentication and provides a structured repository for tracking user activity. This block highlights the importance of data integrity and confidentiality in medical applications.

I. Frontend Display:

The frontend display presents the final segmentation maps and evaluation metrics to the user in a clear, visual format. By overlaying tumor boundaries on CT scans and providing quantitative measures, it supports radiologists in making accurate, evidence-based decisions. This block closes the workflow loop, transforming computational outputs into actionable clinical insights.

VI. RESULTS

A. Homepage Overview:

NeuroCare AI offers advanced lung segmentation technology powered by AI to enhance precision in diagnosing lung conditions. Our platform enables rapid, accurate analysis of CT scans for informed clinical decision-making.

A. Service Overview:

NeuroCare AI offers advanced lung segmentation and analysis using deep learning models to improve clinical decision-making. Our platform automates lung segmentation, provides quantitative reporting, and integrates seamlessly into healthcare workflows.

B. Registration Overview:

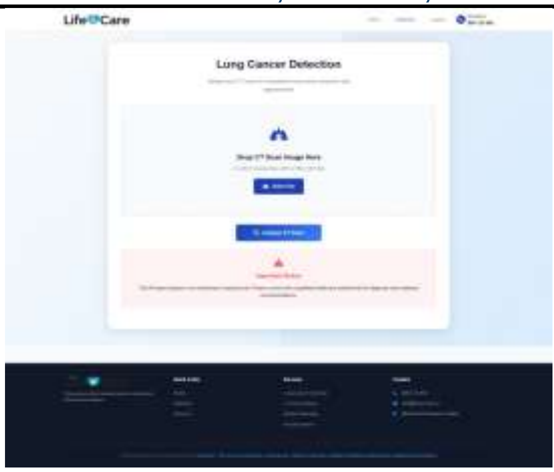
The registration page allows new users to create an account for the Lung Segmentation platform by providing details such as name, email, password, age, gender, and mobile number. Once registered, users can access the platform's features for lung segmentation analysis.

C. Login Overview:

The login page for the Lung Segmentation platform allows users to securely sign in with their email and password. New users can register for an account directly from this page.

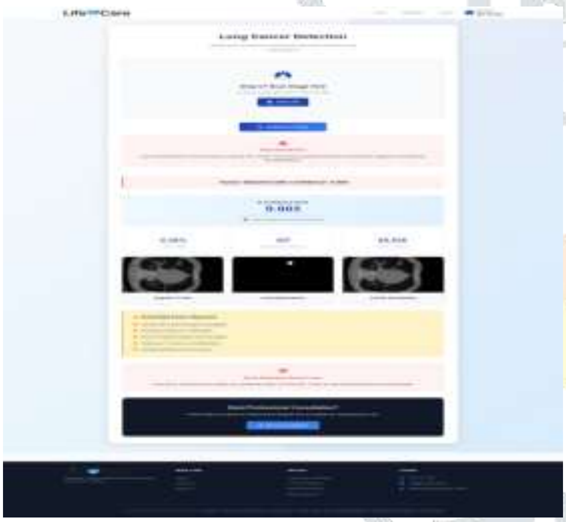
D. Upload Overview:

The Lung Cancer Detection page allows users to upload CT scan images for AI-powered analysis and segmentation. A disclaimer notifies users that the analysis is for preliminary screening only and advises consulting a healthcare professional for further recommendations.



E. Detection Result Overview:

The Lung Cancer Detection page displays the AI confidence score and detailed lung scan analysis after uploading a CT scan image. It provides a visual segmentation of the lung and tumor, along with immediate action recommendations and consultation options.



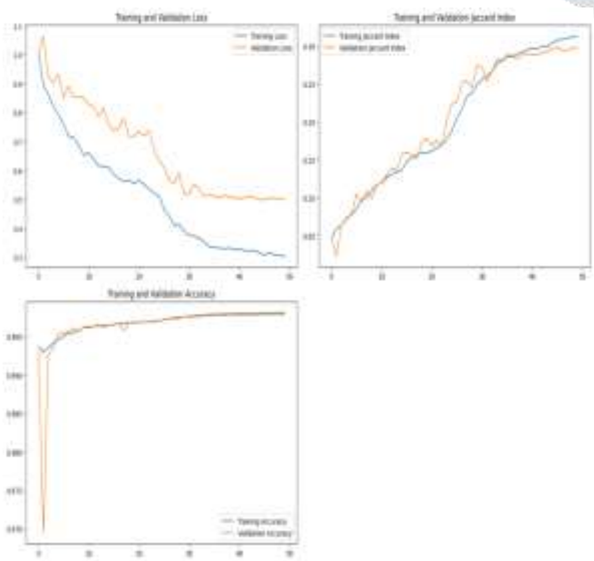


Fig: U-Net Model

The UNet model results show strong performance in lung segmentation tasks. After 50 epochs, the training accuracy is exceptionally high at 99.81%, with a training loss of 0.3032 and a Jaccard index of 0.3133, indicating good model learning. However, the validation accuracy is slightly lower at 99.79%, with a validation loss of 0.5002 and a Jaccard index of 0.2972. These results suggest the model has a slight overfitting issue, as the validation metrics are slightly worse than the training metrics. Despite this, the model performs well for lung segmentation, maintaining high accuracy and reasonable loss values on both training and validation sets.

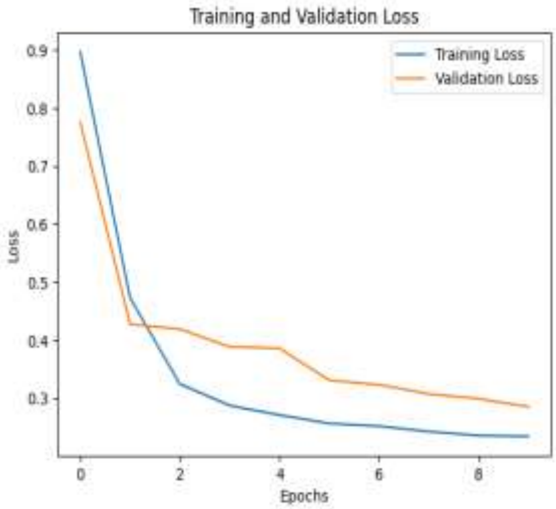


Fig: U-Net+ Model

The UNet++ model shows good performance for the task, with a training loss of 0.2343 and a validation loss of 0.2851, indicating that the model is learning effectively. The training Dice coefficient (0.7657) and Intersection over Union (IoU, 0.6027) suggest solid segmentation accuracy on the training set, while the validation metrics are slightly lower, with a Dice coefficient of 0.7129 and an IoU of 0.5879. This gap suggests some overfitting, as the model performs better on the training data. Precision is high (0.8567 for training, 0.7705 for validation), and recall is also strong (0.7877 for training, 0.8626 for validation), indicating balanced performance between detecting true positives and minimizing false negatives.

B. CONCLUSION

This paper presented an automated lung tumor segmentation framework using U-Net and U-Net++ architectures applied to CT scan images. The proposed method integrates preprocessing techniques such as normalization and noise reduction with deep learning models to improve segmentation accuracy and reliability. Both architectures effectively capture spatial and contextual features, enabling precise delineation of tumor regions.

Experimental results indicate that U-Net provides strong baseline performance, while U-Net++ achieves superior results due to enhanced feature fusion and optimized skip connections. The models demonstrate consistent performance across validation data, with the ability to handle complex tumor structures.

The proposed system reduces the need for manual intervention and supports radiologists in efficient diagnosis and treatment planning. These findings emphasize the potential of deep learning-based segmentation methods for accurate and scalable lung cancer analysis in clinical applications.

C. REFERENCES

[1] L. Zhi, W. Jiang, S. Zhang, and T. Zhou, "An overview of deep neural network pulmonary nodule segmentation methods for CT images," Comput. Biol. Med., vol. XX, 2023.

[2] L. Chen, Y. Gou, Y. Tan, and J. Wu, "A cascaded multi-stage framework for automatic detection and segmentation of pulmonary nodules in developing countries," IEEE J. Biomed. Health Inf., vol. 26, pp. 5619–5630, 2022, doi: 10.1109/JBHI.2022.3198509.

[3] S. F. Banu, M. M. K. Sarker, M. Abdel-Nasser et al., "AWEU-Net: An Attention-Aware Weight Excitation U-Net for Lung Nodule Segmentation," arXiv, Oct. 2021.

[4] X. Zhang, S. Kong, Y. Han, B. Xie, and C. Liu, "Lung Nodule CT Image Segmentation Model Based on Multiscale Dense Residual Neural Network," Mathematics, vol. 11, no. 6, Art. 1363, 2023.

[5] L. Gao, L. Wu, W. Wu, Y. Huang, and M. Xu, "Deep learning in pulmonary nodule detection and segmentation: Systematic review," J. Med. Imaging, 2024.

[6] L. Cai, T. Long, Y. Dai, and Y. Huang, "Mask R-CNN-based detection and segmentation for pulmonary nodule 3D visualization," IEEE Access, vol. 8, pp.44400–44409, 2020.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *in* Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 234–241.