

UNET Model for Lung and Infection Segmentation

UMESH VANGAPALLY - Z2008841

SRI VARSHA MULAKALA - Z1988135

MANEESH REDDY THUDI - Z2000485

ABSTRACT

This work proposes a deep learning-based methodology for lung and infection segmentation in chest CT scans using the U-Net model. The project focuses on two key tasks: identifying lung boundaries and detecting infection regions. Advanced preprocessing techniques, such as contrast enhancement with CLAHE, cropping, and resizing, prepare the data, while data augmentation methods like flipping and scaling improve model performance.

The U-Net model is trained separately for lung and infection segmentation, achieving high accuracy based on metrics like Dice Coefficient and IoU. These models are then combined into a unified system capable of segmenting both lungs and infections simultaneously. Additionally, a user-friendly interface is developed for testing and visualizing segmentation results on new CT images.

This integrated solution offers an efficient and precise tool for automated chest CT analysis, aiding healthcare professionals in diagnosing and assessing pulmonary conditions. Future efforts will focus on extending the model's applications and enhancing its accuracy.

INTRODUCTION

Lung and infection segmentation from chest CT scans is critical in medical imaging for diagnosing and managing respiratory diseases such as COVID-19. Automated segmentation accelerates diagnostic workflows while reducing reliance on manual annotations, which are time-intensive and prone to variability. This study leverages deep learning, particularly the U-Net architecture, to develop an efficient pipeline for accurate segmentation of lungs and infections. The process begins with extensive preprocessing, including contrast enhancement using CLAHE, contour-based cropping of lung regions, and resizing images to standard dimensions. Non-informative slices unrelated to lung anatomy are excluded to optimize the training dataset. Separate U-Net models are trained for lung and infection segmentation using labeled CT scans, lung masks, and infection masks. These models are then integrated into a unified framework that allows simultaneous segmentation, streamlining the diagnostic process. Data augmentation techniques like flipping, rotation, and scaling enhance model generalization, making the system robust to variations in CT scans.

The pipeline's effectiveness is validated using metrics such as Dice Coefficient, IoU, precision, recall, and F1-score, confirming the reliability and accuracy of the segmentation. Additionally, a user-friendly interface has been designed to enable real-time testing and visualization of segmentation results, providing practical utility for clinicians. By combining lung and infection segmentation into a single, scalable framework, the project simplifies the analysis of CT scans, improving diagnostic efficiency and supporting better clinical decision-making. This deep learning-based approach underscores the transformative potential of automated medical imaging,

offering a robust tool to enhance healthcare workflows and facilitate effective treatment planning.

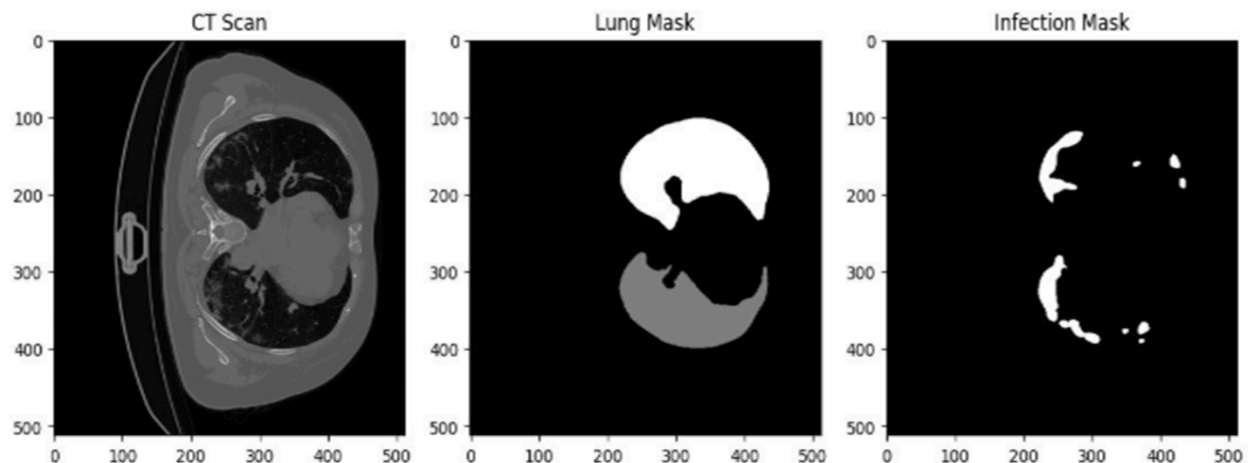
METHODOLOGY

1. Dataset Overview

The dataset includes 20 cases of chest CT scans, stored in the NIfTI (.nii) format, a widely used standard in medical imaging. Each case comprises the following components:

- CT Scans: Three-dimensional volumetric images capturing the chest anatomy.
- Lung Masks: Binary masks that outline the regions of the lungs.
- Infection Masks: Binary masks specifically marking infection regions within the lungs.

All files are neatly organized into directories, with a metadata file ([metadata.csv](#)) included to provide quick and easy reference. This metadata file contains essential details like file paths for each mask and the CT scans, ensuring seamless access to the dataset.



2. Metadata Information

The **metadata.csv** file serves as a comprehensive guide to the dataset, with the following key columns:

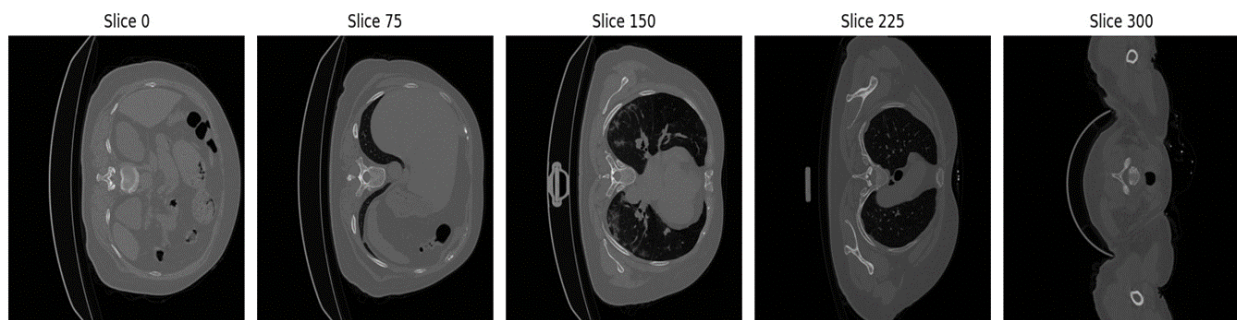
- **ct_scan**: File path for the CT scan images.
- **lung_mask**: File path for the binary lung mask files.
- **infection_mask**: File path for the binary infection mask files.
- **lung_and_infection_mask**: File path for the combined mask files integrating lung and infection segmentations.

This structured organization and detailed metadata make the dataset user-friendly and well-suited for training advanced models in medical image segmentation.

3. Shape and Size

The CT scans and masks in this dataset are represented as 3D arrays, with dimensions varying across samples:

- **Spatial Resolution:** Each slice ranges from 512×512 pixels to 630×630 pixels.
- **Number of Slices:** Each scan includes between 39 and 418 slices.
- **Total Slices:** The dataset contains a total of 3,520 slices, offering a substantial volume of data for training and validation purposes.
- Chest CT scans consist of multiple cross-sectional slices, capturing detailed anatomical structures at different depths. The progression of slices provides critical information for segmentation tasks. Below is an example of slices from a single CT scan, visualized at intervals (Slice 0, Slice 75, Slice 150, Slice 225, and Slice 300):



4. NIfTI Format

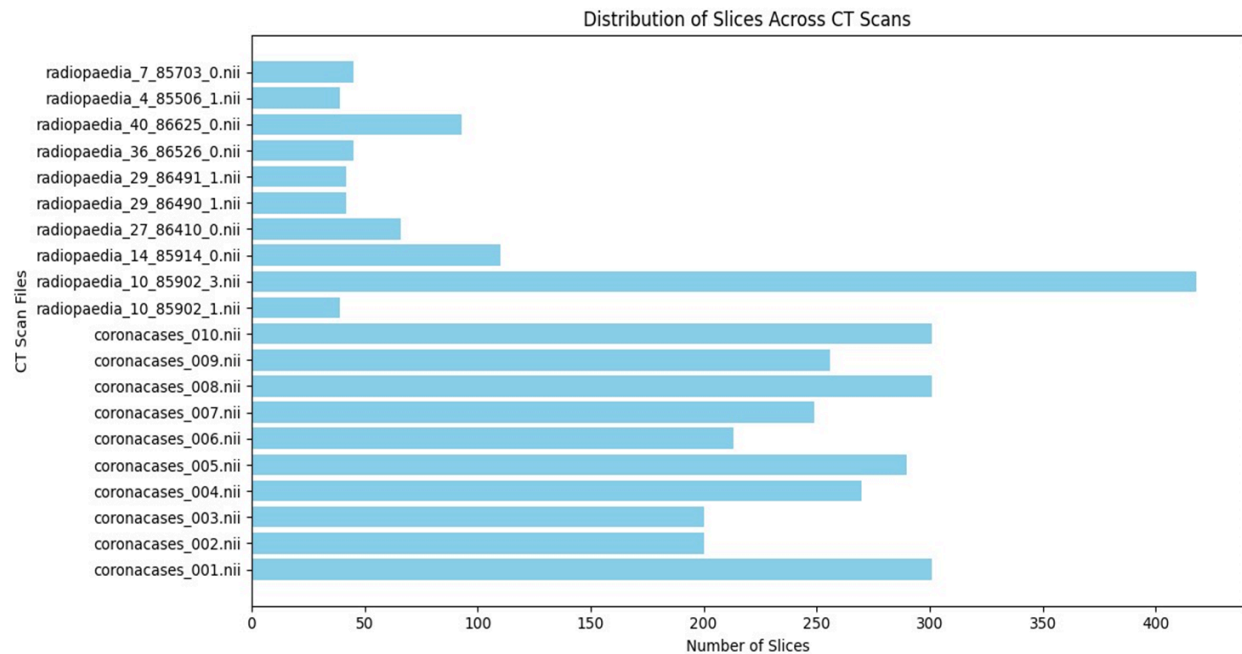
The dataset uses the NIfTI (.nii) format, a widely adopted standard in medical imaging. These files store both 3D volumetric data and accompanying metadata, such as pixel spacing and orientation, making them well-suited for segmentation tasks. Libraries like **nibabel** allow for easy loading and processing of these files, enabling efficient data manipulation.

5. Dataset Statistics

- **Intensity Ranges:** The intensity values in the CT scans exhibit significant variability, with examples ranging from (-1023, 8575) to normalized ranges like (0, 255).
- **Class Balance:** Infection regions account for just 10.39% of the total segmented area, reflecting a notable class imbalance. This underscores the importance of data augmentation to ensure model robustness during training.

6. Slice Distribution

The distribution of slices varies across scans, with each case contributing a different number of slices. This variability introduces diversity in the dataset, which is beneficial for training models to handle a wide range of inputs effectively.



PREPROCESSING

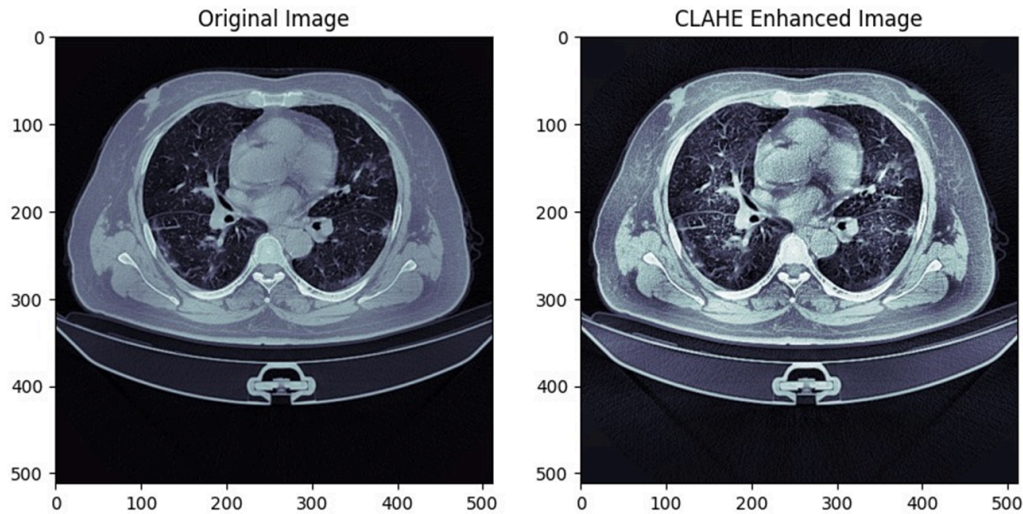
Overview

Preprocessing is a vital step in preparing the dataset for training deep learning models. It ensures consistency, improves image quality, and isolates relevant regions of interest, enhancing segmentation performance. For this dataset, the following preprocessing steps were applied to the CT scans and segmentation masks:

Steps in Preprocessing

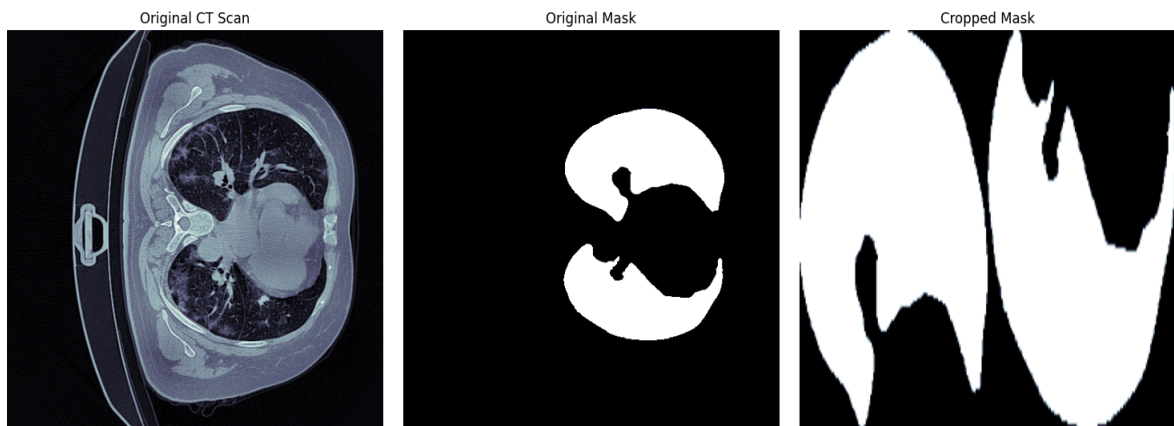
1. Contrast Enhancement (CLAHE):

- Contrast Limited Adaptive Histogram Equalization (CLAHE) was used to enhance local contrast in the CT scans.



2. Lung Region Cropping:

- Using lung masks, the scans were cropped to focus solely on the lung regions.
- Contour-based cropping identified the two largest connected regions in the binary lung mask, effectively isolating the lungs while excluding irrelevant areas like the chest wall and air .



3. Normalization:

- Pixel intensity values in the CT scans were normalized to a range of $[0, 1]$, ensuring uniform intensity scaling across all images.

4. Resizing to Uniform Dimensions:

- Cropped CT scans and masks were resized to a fixed resolution of 224×224 pixels, maintaining the lung region aspect ratio.

5. Non-Informative Slice Removal:

- Slices that lacked meaningful structures or lung regions were excluded from the dataset.
- This step reduced noise and optimized the dataset for effective training.

6. Data Augmentation:

- To address class imbalance and enhance model generalization, several augmentation techniques were applied:
 - **Horizontal flipping** for symmetry considerations.
 - **Rotation** by random angles within a range of -20° to $+20^\circ$.
 - **Scaling** between 90% and 110% to handle size variability.
- These augmentations introduced diversity and improved the model's robustness to imaging inconsistencies and patient positioning variations.

The final preprocessed dataset consists of **2112 CT scans** and **2112 lung masks**, each resized to dimensions **224×224×1** for compatibility with the U-Net model.

Splitting of the Train-Test Set:

The preprocessed dataset was then divided into 70% for training and 30% for validation using the `train_test_split` function to have a good balance between training the model well and assessing its performance.

This gave the following:

Training Data (X): 1478 CT scans, Training Masks (Y): 1478 masks, dimension 224×224×1;
Validation Data (X): 634 CT scans, Validation Masks (Y): 634 masks, dimension 224×224×1.

MODEL ARCHITECTURE

U-Net Architecture

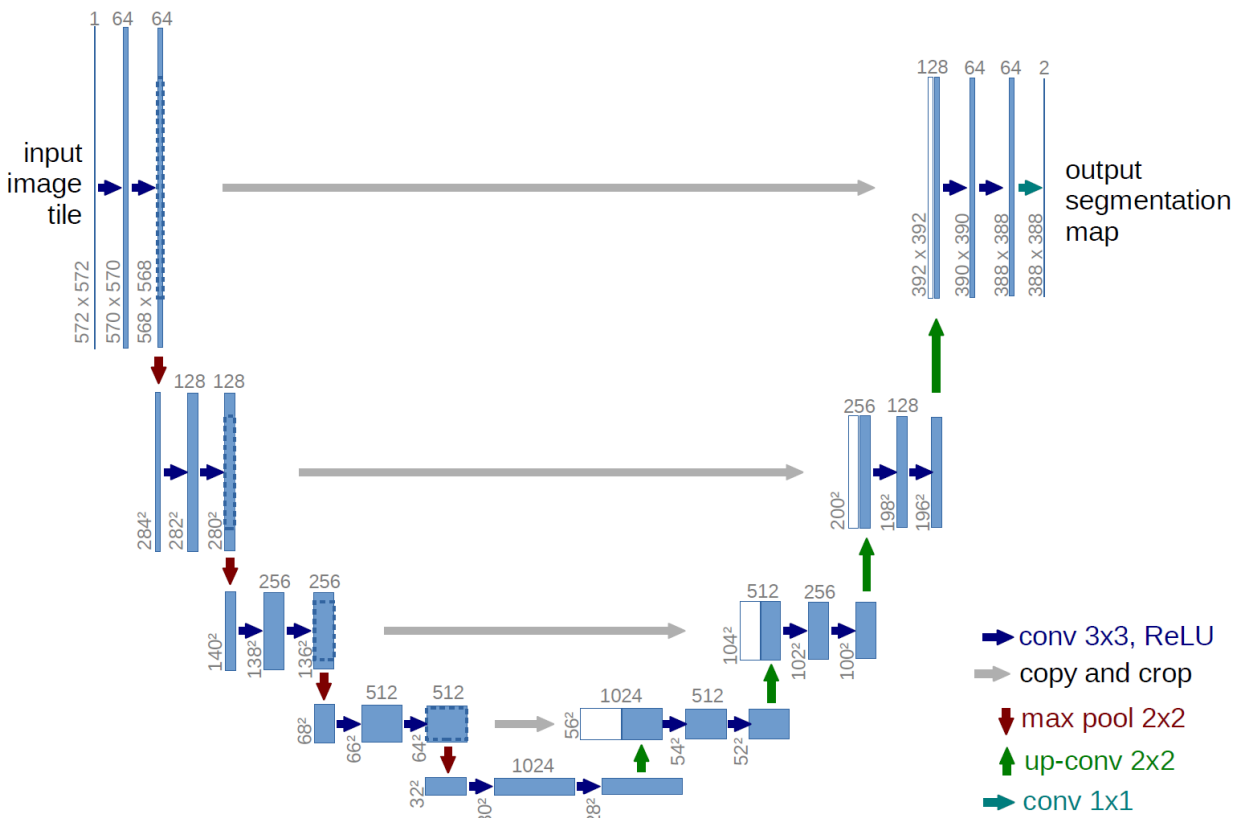
U-Net is a convolutional neural network (CNN) architecture, proposed by Ronneberger et al., explicitly designed for biomedical image segmentation tasks. Developed for use in medical imaging, U-Net is characterized by its symmetric architecture comprising an encoder-decoder structure with skip connections that bridge corresponding layers of the encoder and decoder. These skip connections allow the model to retain spatial information while capturing hierarchical feature representations.

- **Encoder (Contracting Path):** The encoder path captures contextual information through successive layers of:
 - **Convolutions:** Two 3×3 convolutional layers with ReLU activations.
 - **Batch Normalization:** Applied to stabilize training and improve convergence.
 - **Pooling:** A 2×2 max-pooling layer down-samples the feature maps and increases their depth at each stage.
- **Bottleneck:** The bottleneck serves as a bridge between the encoder and decoder, capturing high-level abstract representations while maintaining the spatial dimensions essential for reconstruction.

- **Decoder (Expanding Path):** The decoder reconstructs the segmentation map through:
 - **Upsampling:** Transposed convolutions (2×2) for spatial up-sampling.
 - **Skip Connections:** Features from the encoder are concatenated with corresponding up-sampled features in the decoder, ensuring the preservation of spatial details.
 - **Final Output:** A convolutional layer with softmax activation generates probability maps for each class.

U-Net is highly effective for segmentation tasks due to its ability to handle small datasets, leverage spatial information, and efficiently segment complex structures.

The below is the block diagram for U-Net architecture :



DeepLabV3+ Architecture

The **DeepLabV3+** architecture is a state-of-the-art model for semantic segmentation. It combines a **ResNet50 backbone** for feature extraction with an **Atrous Spatial Pyramid Pooling (ASPP)** module to capture multi-scale context. The key components include:

- **Encoder:** ResNet50 extracts low-level spatial features (`conv2_block3_out`) and high-level semantic features (`conv4_block6_out`).

- **ASPP:** Multi-scale dilated convolutions capture context at different receptive fields, complemented by global average pooling.
- **Decoder:** Combines ASPP features with low-level features via upsampling and convolutions to refine boundary details.
- **Output Layer:** A 1x1 convolution with a sigmoid activation generates the binary segmentation mask.

This architecture excels in precise segmentation by leveraging multi-scale context and fine spatial details, making it ideal for medical image segmentation tasks like lung and infection segmentation.

MODEL TRAINING

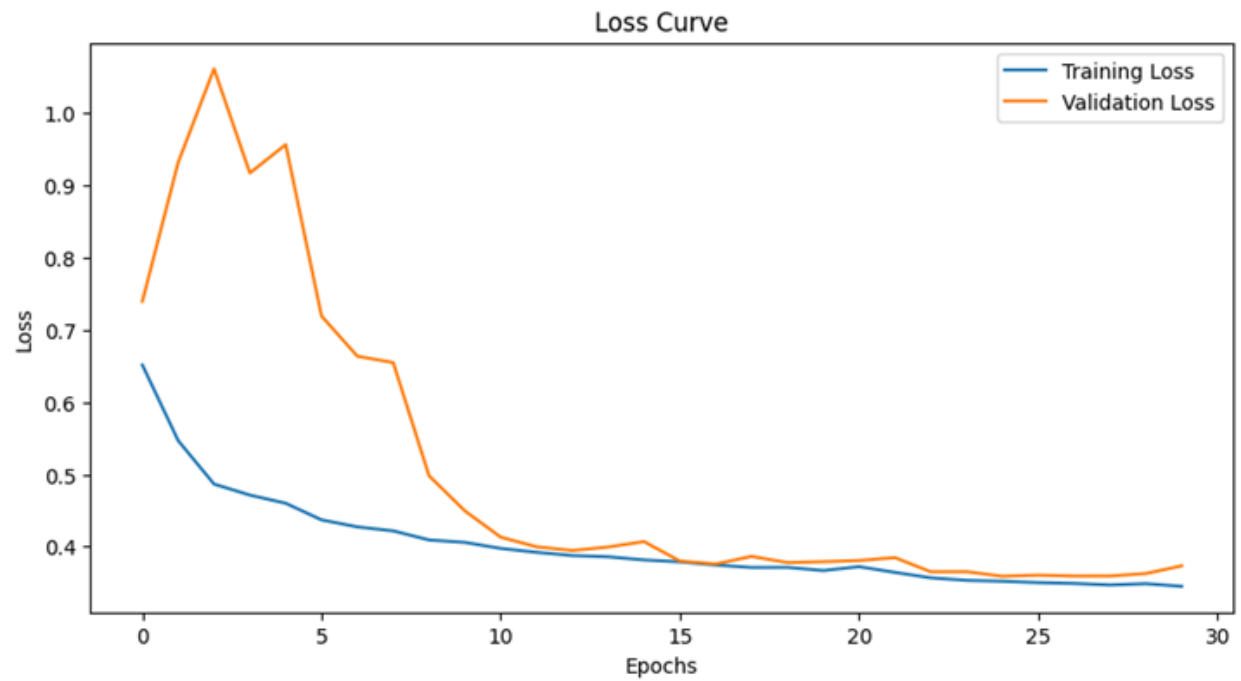
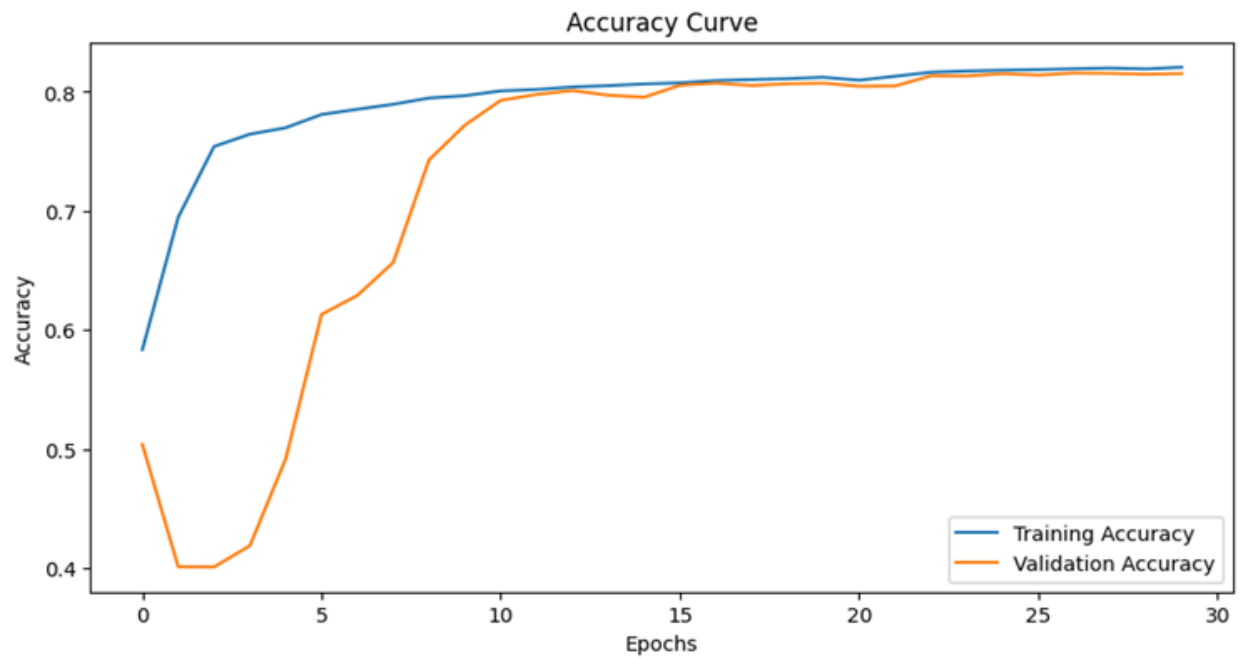
The preprocessed dataset was used for training the U-Net model in segmenting lungs and infections from CT scans. The training was done with much attention paid to its optimization using the binary cross-entropy loss and Adam optimizer. Several hyperparameters were chosen to ensure that learning is effective, and different callbacks such as ModelCheckpoint and ReduceLROnPlateau were used to improve the performance of the model and avoid overfitting. The model was trained for a total of 30 epochs, saving its best performance based on validation loss for further evaluation.

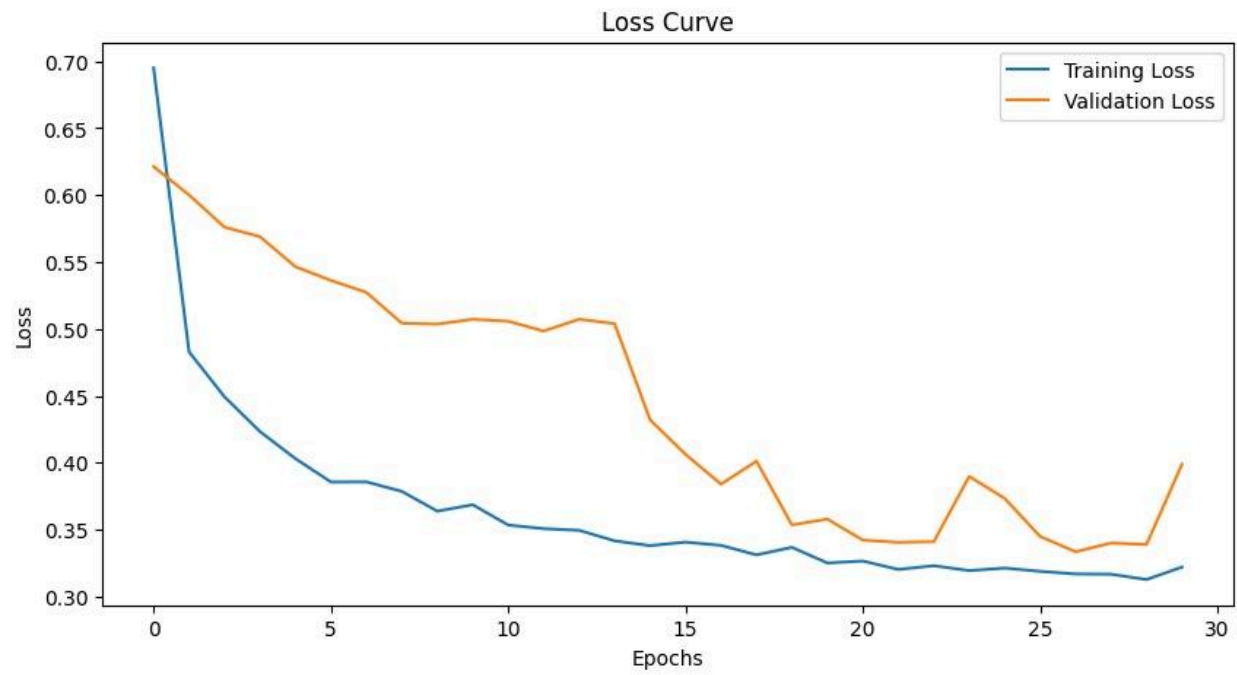
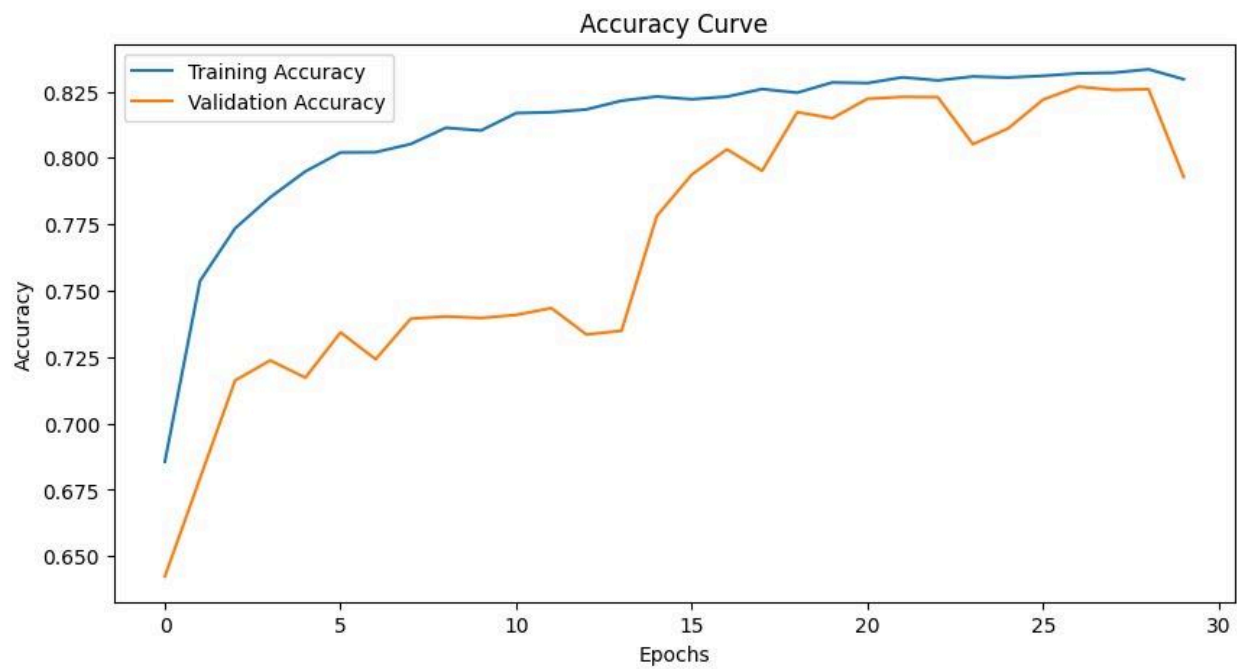
Hyperparameters Used:

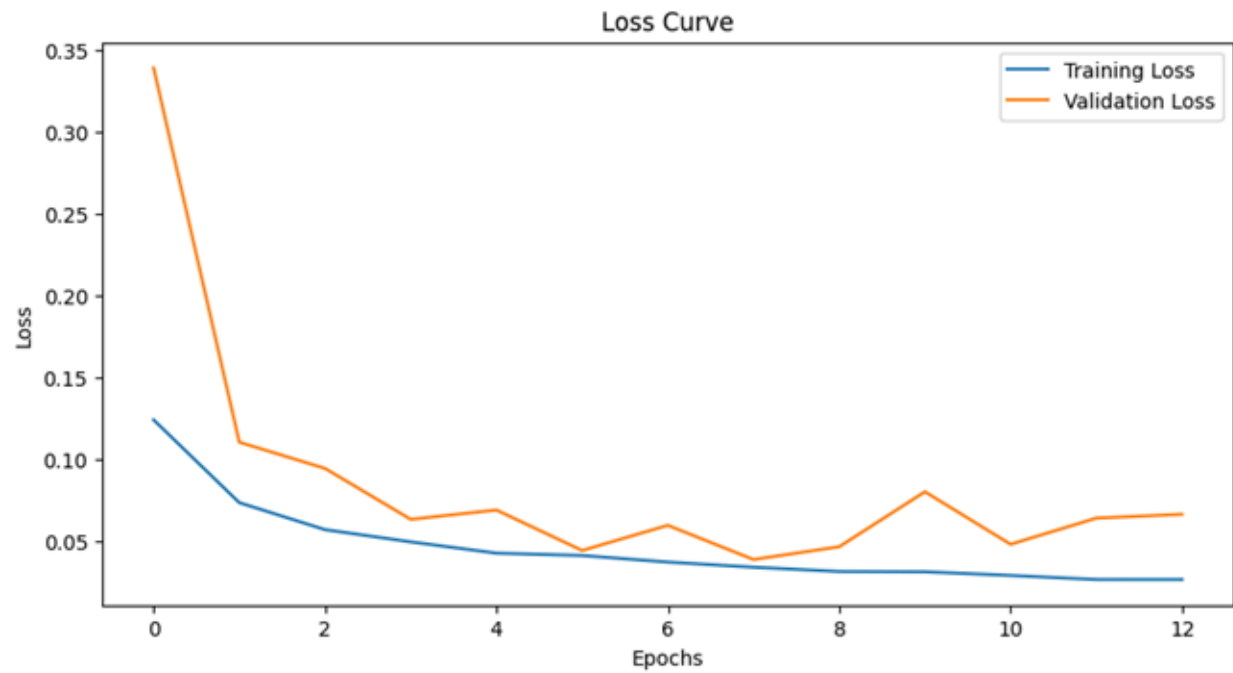
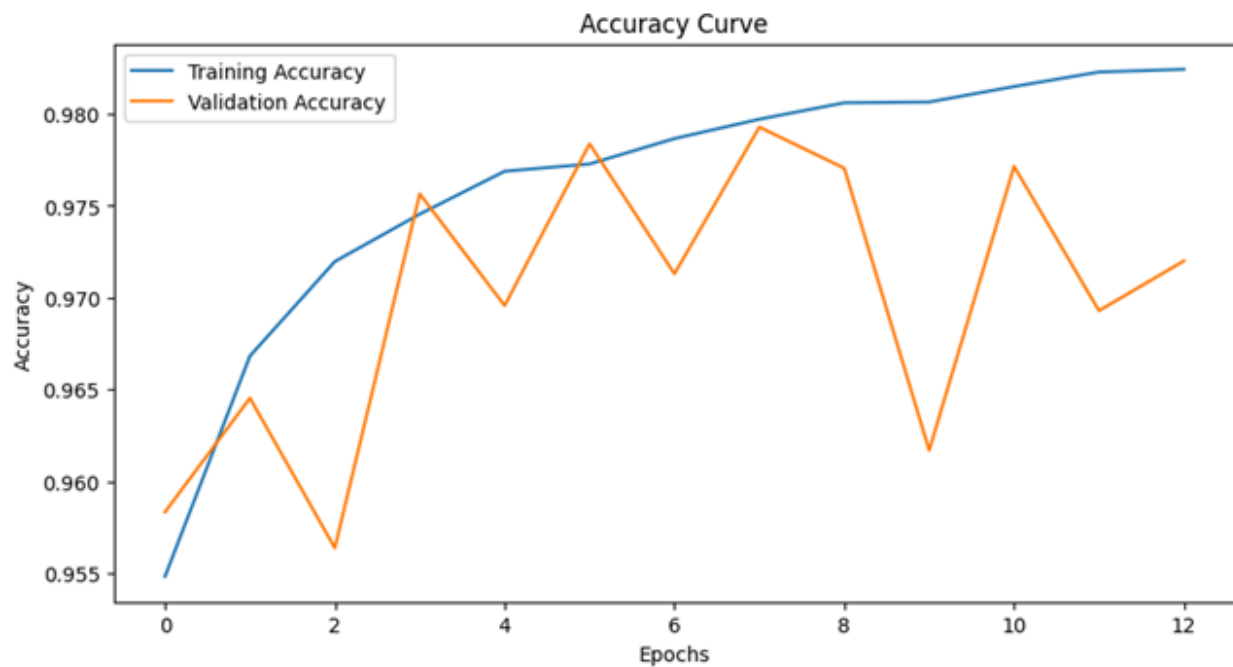
- Model: U-Net architecture and DeepLabV3+
- Loss Function: Binary Cross-Entropy.
- Optimizer: Adam with a learning rate of 0.0005.
- Batch Size: 32.
- Number of Epochs: 30.
- Callbacks:
 - ModelCheckpoint: Saved the best model based on validation loss.
 - ReduceLROnPlateau: Dynamically reduced the learning rate when validation loss plateaued.

RESULTS

The training process was evaluated using accuracy and loss metrics for both the training and validation datasets. The accuracy curve indicates the model's ability to improve performance over epochs, while the loss curve demonstrates the minimization of errors during training.

Lung Segmentation Loss Curve_U-NET:**Lung Segmentation Accuracy Curve_U-NET:**

Lung Segmentation Loss Curve_DeepLabV3+ :**Lung Segmentation Accuracy Curve_DeepLabV3+ :**

Infection Segmentation Loss Curve_U-NET:**Infection Segmentation Accuracy Curve_U-NET:****EVALUATION METRICS**

To evaluate the performance of the trained U-Net models for lung and infection segmentation, various metrics were used, including the Dice Coefficient, Intersection over Union (IoU), Precision, Recall, and F1-Score. These metrics provide a robust measure of segmentation quality by assessing overlap, accuracy, and boundary detection.

Lung Segmentation Metrics_U-NET

- Mean Dice Coefficient: 0.8513
- Mean IoU Score: 0.7652

Lung Segmentation Metrics_DEEPLABV3+

- Mean Dice Coefficient: 0.8623
- Mean IoU Score: 0.7831

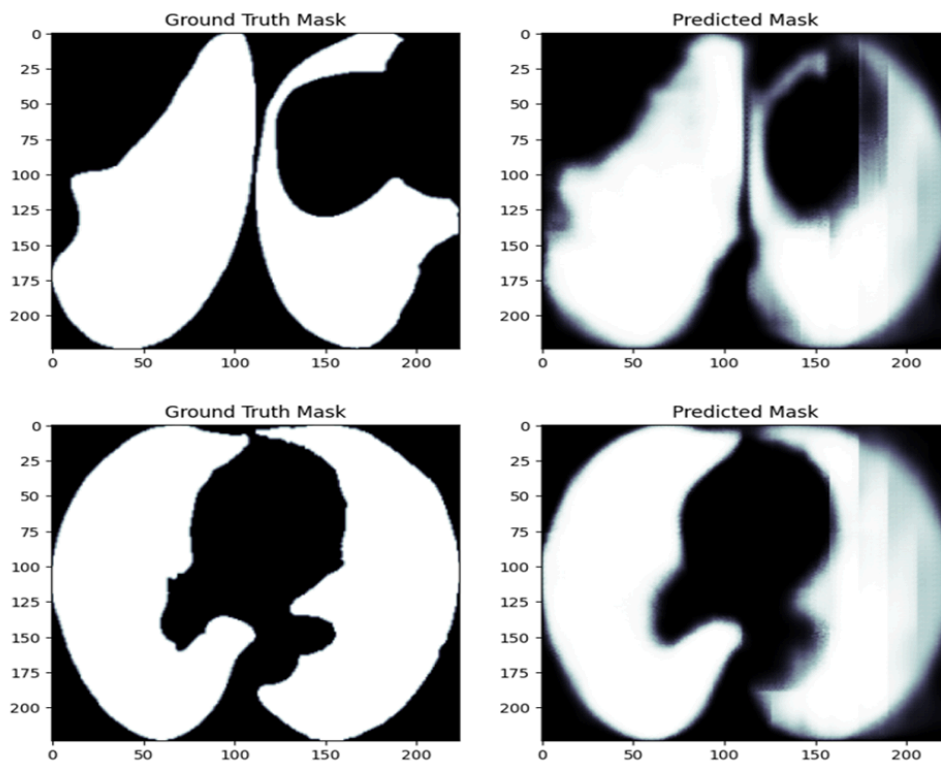
Infection Segmentation Metrics

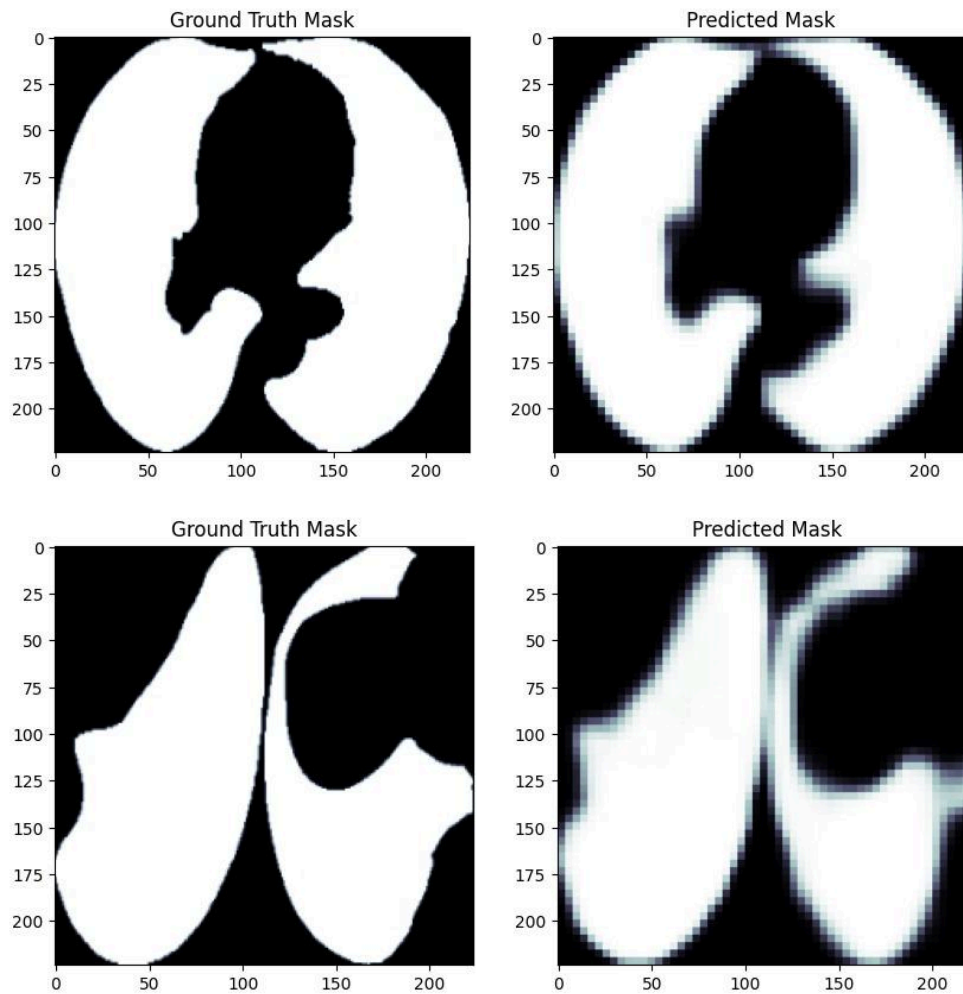
- Mean Dice Coefficient: 0.7210
- Mean IoU Score: 0.6450

OUTPUT

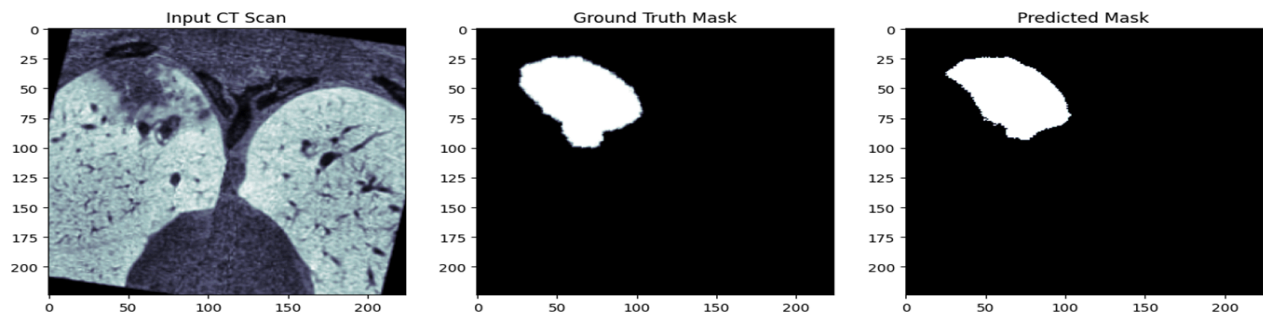
Lung Segmentation_U-NET

The top image shows the ground truth and predicted masks for lung segmentation. The model effectively captures lung contours, with minor inconsistencies observed at the edges.



Lung Segmentation_DEEPLABV3+**Infection Segmentation**

The bottom image illustrates infection segmentation, comparing the input CT scan, ground truth mask, and predicted mask. The model accurately identifies infection regions, though minor deviations in boundaries and small regions are noted.



UI INTERFACE:

The project includes a user-friendly interface to enhance usability and provide an intuitive workflow for users.

Step 1: Segmentation Type Selection

- Users can select the type of segmentation, either lung or infection segmentation, through the interface.
- This ensures the process is customized for the specific medical imaging analysis.

Step 2: File Upload

- The interface allows users to easily upload CT scan images or grayscale medical images.
- This simplifies the process of providing input data for segmentation.

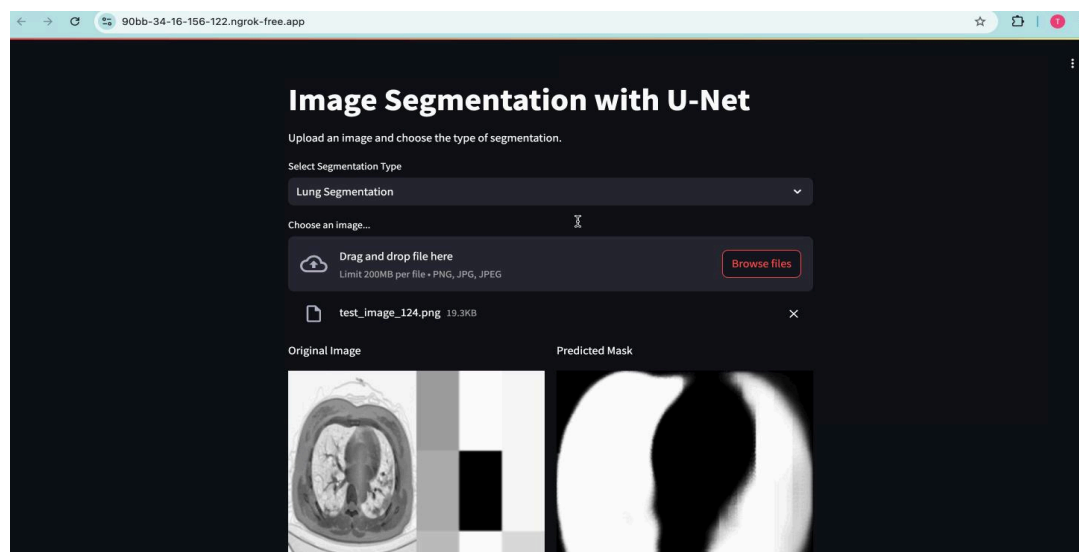
Step 3: Results Display

- The interface displays the original medical image and the predicted segmentation mask side by side.
- This enables easy visualization and comparison of results.

Step 4: Evaluation and Comparison

- Users can interact with the interface to evaluate the segmentation results and compare them for accuracy and relevance.

An image of the user interface has been included for reference, visually demonstrating how users interact with the system and showcasing its simplicity and functionality.



FUTURE SCOPE

Integration of Models: Combining the lung and infection segmentation models to develop a comprehensive pipeline for simultaneous segmentation and diagnostic analysis of CT scans.

Disease Classification: Enhancing the integrated model by incorporating additional features to classify pulmonary diseases, improving diagnostic accuracy and supporting clinical decision-making.

Infection Segmentation: While the DeepLabV3+ model was successfully trained for lung segmentation, extending the training to include infection segmentation remains an essential next step. Due to time constraints, infection segmentation using DeepLabV3+ was not carried out in this study. Future efforts can focus on training and validating the model for infection segmentation to improve its utility in detecting and analyzing diseases like COVID-19.

REFERENCES

- [1] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2020). Image Segmentation Using Deep Learning: A Survey. <https://arxiv.org/pdf/2001.05566>
- [2] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 234-241). Springer, Cham. <https://arxiv.org/abs/1505.04597>
- [3] Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation. *Nature Methods*, 18(2), 203–211. <https://doi.org/10.1038/s41592-020-01008-z>
- [4] Raichle, M. E., Glover, G. H., & Mazziotta, J. C. (2017). Neuroimaging Informatics Technology Initiative (NIfTI) and DICOM: Integration of Medical Imaging Data and Tools in Neuroimaging Research. <https://doi.org/10.3389/fnins.2017.00491>
- [5] Ma, J., Ge, C., Wang, Y., An, X., Gao, J., Yu, Z., ... He, J., 2020. COVID-19 CT Lung and Infection Segmentation Dataset (Version 1.0) .Zenodo. <https://doi.org/10.5281/zenodo.3871061>