High-Precision 3D Spleen Segmentation for Surgical Planning Using Deep Learning

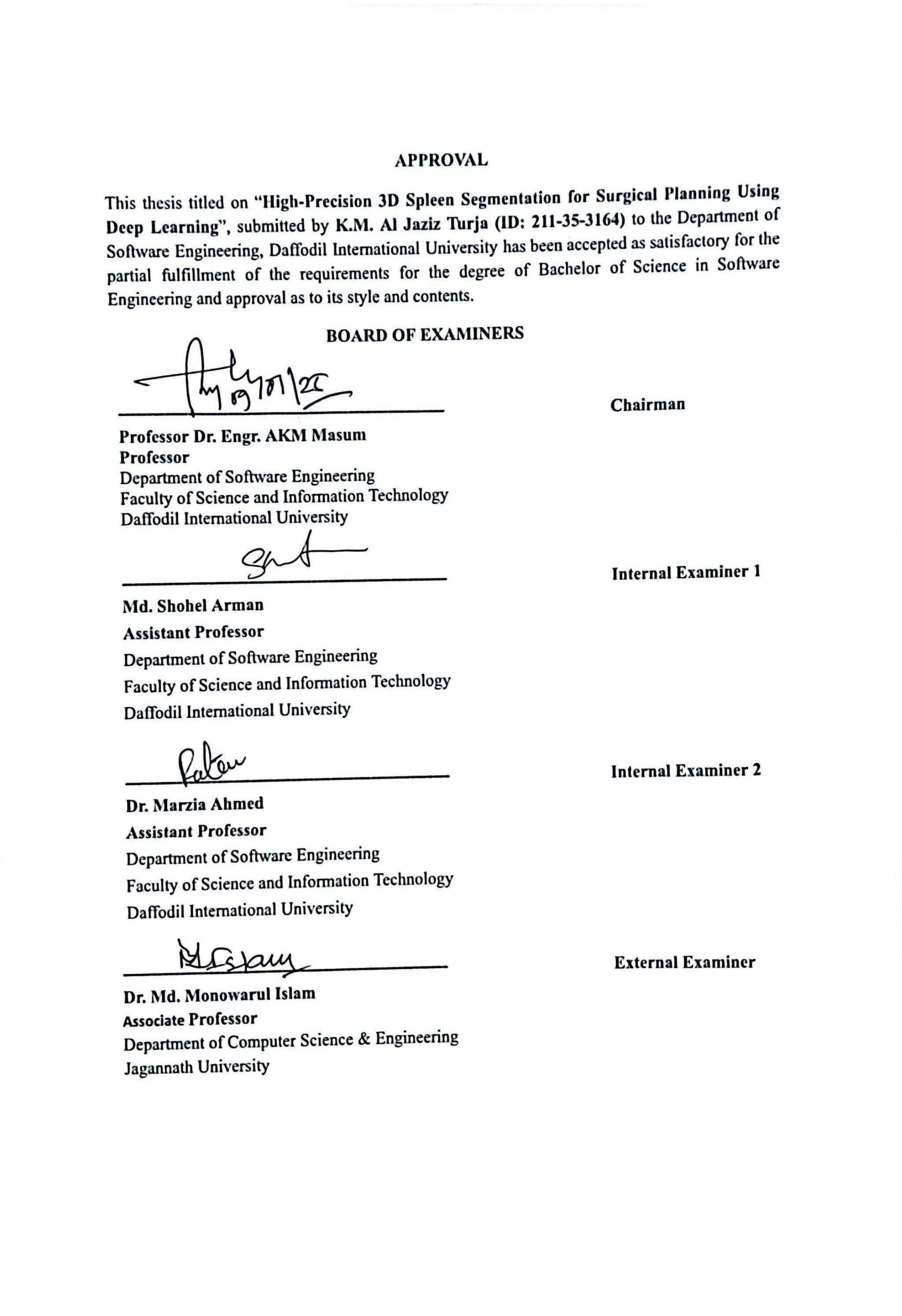
Submitted By
K.M. Al Jaziz Turja
ID: 211-35-3164
Department of Software Engineering
Daffodil International University

Supervised By
Md. Shohel Arman
Assistant Professor
Department of Software Engineering
Daffodil International University

Thesis submitted in fulfillment of the requirements for the award of the degree of Bachelor of Science

Fall 2024

APPROVAL



APPROVAL

This thesis titled on **"High-Precision 3D Spleen Segmentation for Surgical Planning Using Deep Learning"**, submitted by **K.M. Al Jaziz Turja (ID: 211-35-3164)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

09/07/25

Chairman

Professor Dr. Engr. AKM Masum
Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1

Md. Shohel Arman
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

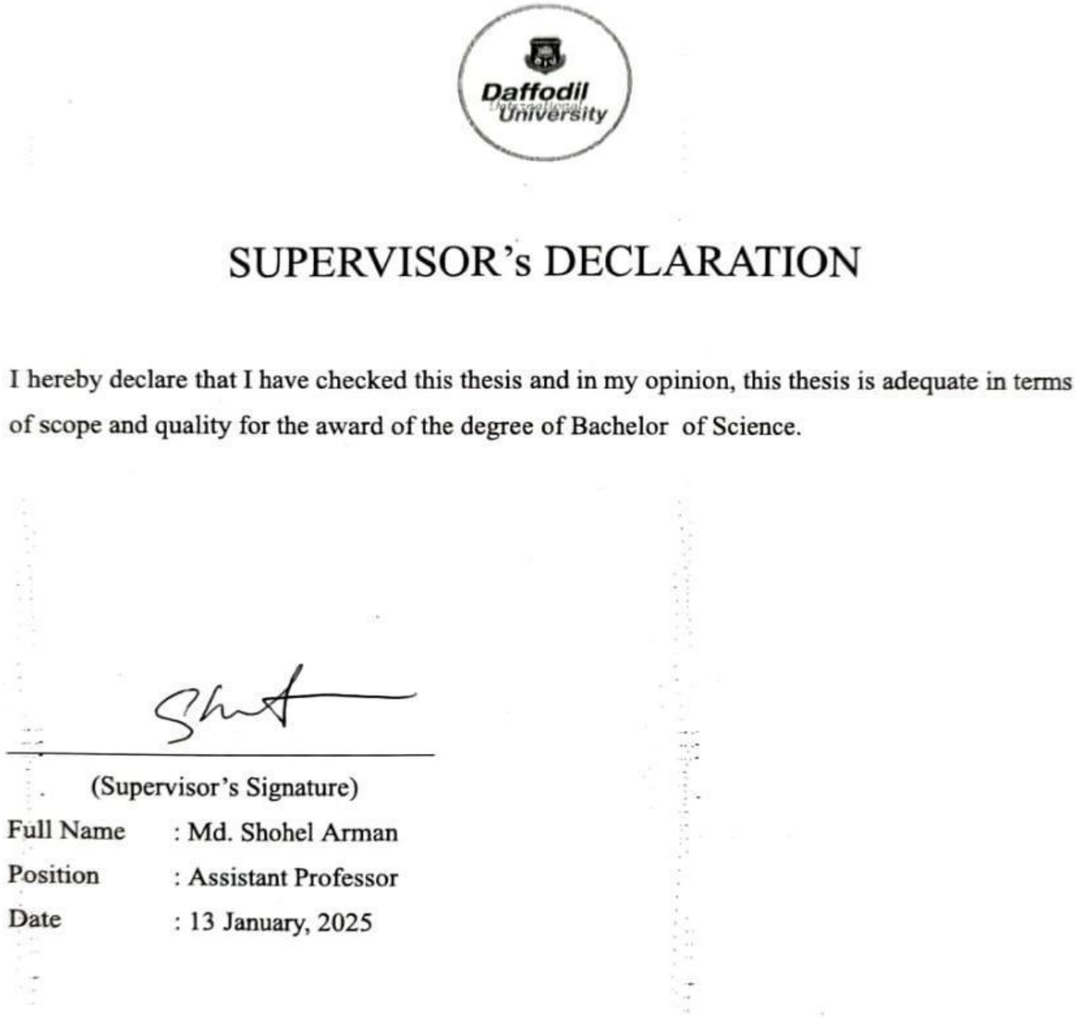
Internal Examiner 2

Dr. Marzia Ahmed
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

External Examiner

Dr. Md. Monowarul Islam
Associate Professor
Department of Computer Science & Engineering
Jagannath University

SUPERVISOR's DECLARATION



SUPERVISOR's DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science.

(Supervisor's Signature)

Full Name : Md. Shohel Arman

Position : Assistant Professor

Date : 13 January, 2025

STUDENT’s DECLARATION

STUDENT’S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.

(Student’s Signature)

Full Name : K.M. Al Jaziz Turja
ID Number : 211-35-3164
Date : 13 Januray 2025

ACKNOWLEDGEMENT

My profound appreciation goes out to Md. Shohel Arman of Daffodil International University's Department of Software Engineering, who served as my thesis advisor. My supervisor's unfaltering encouragement, priceless advice, and perceptive criticism have shaped this research in every way. I have always been motivated by his unwavering commitment to academic success and his desire to foster the intellectual development of his students.

Additionally, I'd want to convey my deepest gratitude to the brilliant professors who make up the Software Engineering Department. Throughout this study, I have been fortunate to have access to a wide range of viewpoints and insightful criticisms thanks to their combined knowledge and support. Their nurturing academic atmosphere has provided me with an ideal setting in which to develop my thoughts and hone my research.

In addition, I would want to express my gratitude for the unfaltering and continual support that my loved ones have given me. Thanks to their understanding, encouragement, and patience, I was able to overcome the obstacles I encountered while conducting this extensive study project. I am deeply motivated to perform academically because of their unwavering faith in my talents.

To conclude, I would want to extend my deepest gratitude to everyone who has helped me along the way in my academic career. This research and my academic and personal development have been profoundly impacted by your collaborative support.

ABSTRACT

Medical imaging has become an indispensable tool in modern healthcare, facilitating accurate visualization of internal anatomical structures for diagnosis, treatment planning, and surgical procedures. Among these, segmentation of organs such as the spleen is particularly critical for pre-surgical workflows, including radiation therapy planning, volume estimation, and surgical navigation. The spleen, an organ of considerable clinical significance, often requires precise delineation for managing trauma or pathologies like lymphoma. However, spleen segmentation from computed tomography (CT) images is challenging due to anatomical variability, low tissue contrast, and the presence of imaging artifacts. Traditional segmentation methods, such as thresholding and active contour models, often fail to generalize across diverse datasets, necessitating advanced approaches.

This study aims to address these challenges by developing a robust segmentation framework based on a 3D U-Net architecture. The primary objectives include achieving high segmentation performance on the Medical Segmentation Decathlon (MSD) Task 09 spleen dataset, evaluating the model's generalization capabilities, and demonstrating its clinical relevance for pre-surgical workflows. The methodology involves a systematic preprocessing pipeline to normalize and augment the CT volumes, followed by training a 3D U-Net model using Dice loss and the Adam optimizer over 600 epochs. The segmentation performance was evaluated using metrics such as Dice similarity coefficient and validation loss to ensure accuracy and reliability.

The proposed framework achieved remarkable results, with a training Dice similarity coefficient of 0.9582, a validation Dice similarity coefficient of 0.9494 and a test Dice Similarity coefficient of 0.9483. These results highlight the model's strong generalization ability and effectiveness in addressing challenges such as noisy imaging and anatomical variability. Qualitative evaluations further confirm the precision of the segmentation, with predicted masks showing high alignment with ground truth annotations.

In conclusion, this study presents a significant contribution to automated medical image segmentation by leveraging the power of deep learning. The results not only validate the efficacy of the proposed framework but also lay the groundwork for extending this approach to multi-organ segmentation tasks and real-time clinical workflows. Future research directions include integrating multi-modal imaging data, improving computational efficiency, and validating the framework in clinical settings to ensure broader applicability and impact.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER 1

INTRODUCTION

1.1 Background

Medical imaging is an indispensable tool in modern clinical workflows, enabling the non-invasive visualization of internal anatomical structures for diagnosis, treatment planning, and surgical interventions. Organ segmentation—the process of delineating specific anatomical structures from medical images—is a critical step in these workflows. Among various organs, the spleen plays a vital role in pre-surgical planning, including radiation therapy, volume estimation, and surgical navigation. The spleen, a key organ in trauma management and diseases like lymphoma, often requires precise delineation for procedures like splenectomy. However, segmenting the spleen in computed tomography (CT) images is inherently challenging due to factors such as anatomical variability, low contrast between the spleen and adjacent tissues, and noise artifacts in medical imaging (Antonelli et al., 2021).

What is Computed Tomography (CT)?

Computed Tomography (CT) is an advanced medical imaging technique that combines X-ray measurements taken from multiple angles around the body. It reconstructs cross-sectional images, or "slices," of the internal anatomy, enabling detailed visualization of bones, organs, and soft tissues. CT is widely used due to its speed, accuracy, and ability to provide 3D insights into patient anatomy.

How Does CT Work?

- X-Ray Emission: A rotating X-ray tube emits X-rays toward the patient.
- Detection: Detectors capture attenuated X-rays after passing through the body.
- Data Acquisition: The intensity of transmitted X-rays is recorded, varying with tissue density.
- Reconstruction: Algorithms compute cross-sectional images from detected X-ray data.

What Does the CT Output Look Like?

CT output consists of grayscale cross-sectional images where pixel intensity values correspond to tissue density:

- Bone: White (high attenuation).
- Soft tissue: Gray (moderate attenuation).
- Air and fluid: Black or dark (low attenuation).

How Are CT Images Represented in Computers?

CT images are stored digitally in formats like DICOM (Digital Imaging and Communications in Medicine). Each scan contains:

- Voxel Grid: A 3D matrix where each voxel represents tissue density in Hounsfield Units (HU).
- Air: -1000 HU, Water: 0 HU, Bone: +1000 HU or more.
- Slices: Stacked 2D slices to reconstruct 3D anatomy.
- Coordinate Systems: Orientation is standardized using systems like RAS (Right-Anterior-Superior).

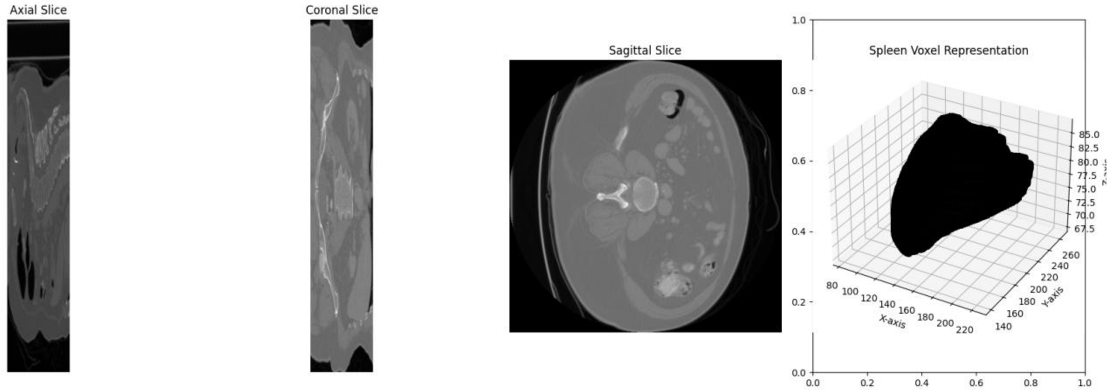


Figure 1.1: Representation of CT images in computer

What Others Have Done

Researchers have explored various methods to address spleen segmentation challenges:

Traditional Methods: Techniques like thresholding and region growing were among the earliest approaches but often struggled due to low tissue contrast and noise artifacts (McInerney & Terzopoulos, 1996).

Machine Learning: Methods such as Random Forests and Support Vector Machines (SVMs) introduced learning-based segmentation but required extensive feature engineering and lacked scalability (Criminisi et al., 2013).

Deep Learning:

- U-Net, a convolutional neural network, revolutionized medical segmentation by introducing an encoder-decoder structure with skip connections (Ronneberger et al., 2015).
- 3D U-Net extended this architecture to volumetric data, enabling more precise segmentation of organs (Çiçek et al., 2016).
- Advanced models, such as Swin UNETR, incorporated transformer-based architectures to capture long-range dependencies, further enhancing segmentation accuracy (Hatamizadeh et al., 2022).

Despite these advancements, challenges like computational cost, generalization to diverse datasets, and robustness in noisy conditions remain areas of active research.

1.2 Motivation of the Research

Traditional segmentation approaches, such as thresholding, region-growing methods, and active contour models, have shown limited success due to their reliance on hand-crafted features and fixed assumptions about the anatomical structure (McInerney & Terzopoulos, 1996). These methods often fail in the presence of anatomical variability or noisy imaging conditions. Machine learning techniques have improved segmentation by learning features directly from the data. However, these methods struggle to model the complex spatial relationships inherent in 3D volumetric data (Criminisi et al., 2013).

The advent of deep learning, particularly convolutional neural networks (CNNs), has revolutionized medical image analysis. CNN-based architectures such as U-Net and its 3D variant have demonstrated state-of-the-art performance in biomedical segmentation tasks (Ronneberger et al., 2015; Çiçek et al., 2016). Their encoder-decoder structures, coupled with skip connections, effectively capture both global and local contextual information. Motivated by the success of these models, this research aims to leverage a 3D U-Net architecture to address the challenges of spleen segmentation, particularly those posed by anatomical variability and low contrast.

1.3 Problem Statement

Despite significant advancements in medical image analysis, spleen segmentation remains a challenging task due to a variety of inherent complexities. Anatomical variability across patients results in differences in spleen shape, size, and orientation, making it difficult to apply generic segmentation methods effectively. Furthermore, the low contrast between the spleen and adjacent structures, such as the liver, stomach, and surrounding soft tissues, exacerbates the difficulty of accurately delineating organ boundaries in computed tomography (CT) images. Noise and imaging artifacts further obscure the spleen's contours, reducing the reliability of segmentation outcomes. Traditional approaches, such as thresholding and active contour models, often fail to generalize across datasets and struggle with complex anatomical structures. While deep learning-based methods, such as convolutional neural networks (CNNs), have demonstrated significant promise, challenges remain in ensuring robustness, generalization, and computational efficiency for real-world clinical applications. Thus, there is a critical need for a robust, efficient, and scalable framework for spleen segmentation that can address these challenges and enable seamless integration into clinical workflows (Antonelli et al., 2021; Zhou et al., 2018).

1.4 Research Questions

This study addresses the following research questions:

- What preprocessing steps are essential to ensure accurate segmentation performance?
- How does the proposed segmentation pipeline generalize across unseen data in terms of accuracy?

1.5 Research Objectives

- To develop a robust segmentation framework for the spleen using a 3D U-Net architecture.

- To achieve high segmentation performance on the Medical Segmentation Decathlon (MSD) Task 09 spleen dataset.

1.6 Research Scope

This study focuses on spleen segmentation from CT volumes using a deep learning-based approach. The Medical Segmentation Decathlon (MSD) Task 09 spleen dataset serves as the primary dataset for training and validation. While the study aims to deliver a robust segmentation framework for the spleen, it does not extend to other imaging modalities (e.g., MRI or PET) or multi-organ segmentation tasks. Future research could explore these additional areas.

1.7 Thesis Organization

The thesis is organized as follows:

- **Chapter 1: Introduction**—This chapter introduces the research background, motivation, problem statement, research questions, objectives, and scope.
- **Chapter 2: Literature Review**—A review of existing methods for medical image segmentation, including traditional approaches and recent advancements using deep learning, highlighting gaps in current research.
- **Chapter 3: Methodology**—Details the dataset, preprocessing steps, model architecture, and training process employed in this study.
- **Chapter 4: Results and Discussion**—Presents the quantitative and qualitative results of the segmentation model, discussing its implications and limitations.
- **Chapter 5: Conclusion and Future Work**—Summarizes the key findings, contributions, and limitations of the research, and proposes directions for future exploration.

CHAPETER 2

LITERATURE REVIEW

2.1 Introduction

The segmentation of anatomical structures in medical imaging is a critical step in enabling precise diagnostics and treatment planning. Over the decades, numerous methods have been proposed to achieve reliable segmentation of organs, including the spleen, from medical imaging data. Early approaches relied heavily on traditional image processing techniques, such as thresholding, active contours, and region-growing methods. While these methods were computationally efficient, they often struggled with complex anatomical variability and noisy imaging data. The advent of machine learning brought significant improvements by automating feature extraction and decision-making processes. However, the real revolution came with deep learning, particularly convolutional neural networks (CNNs), which introduced unprecedented levels of accuracy and robustness for medical image segmentation tasks. This section reviews the evolution of segmentation methodologies, focusing on spleen segmentation and highlighting the contributions and limitations of various techniques.

2.1 Previous Literature

Traditional Approaches

Early segmentation techniques focused on rule-based algorithms that relied on handcrafted features. Thresholding and region-growing methods (McInerney & Terzopoulos, 1996) were among the first approaches applied to organ segmentation. However, these methods were highly sensitive to noise and lacked the ability to generalize across datasets with variable anatomical features. Active contour models improved upon these methods by incorporating edge detection and shape constraints, but they still struggled with low-contrast boundaries and anatomical complexity (Kass et al., 1988).

Machine Learning-Based Methods

The advent of machine learning introduced supervised algorithms like support vector machines (SVMs) and random forests, which used statistical features for segmentation (Criminisi et al., 2013). These methods demonstrated better performance than traditional approaches but required substantial manual effort for feature engineering. Moreover, their applicability to 3D volumetric data was limited due to computational inefficiency and lack of scalability.

Deep Learning for Medical Image Segmentation

Deep learning, particularly CNNs, has emerged as a transformative technology in medical image analysis. Ronneberger et al. (2015) introduced the U-Net architecture, which became a benchmark for biomedical image segmentation. Its encoder-decoder structure with skip connections allowed for the integration of spatial and contextual information, enabling superior segmentation performance. Çiçek et al. (2016) extended this architecture to 3D data, introducing the 3D U-Net, which became the standard for volumetric segmentation tasks.

The Medical Segmentation Decathlon Task 03 Liver (Antonelli et al., 2021) has been widely used to benchmark segmentation models. Studies leveraging this dataset have explored various deep learning architectures. For example, Zhou et al. (2018) proposed the UNet++ architecture, which introduced nested skip connections to improve feature reuse and segmentation accuracy. Hatamizadeh et al. (2022) proposed Swin UNETR, a transformer-based architecture that captured long-range dependencies in 3D data, achieving state-of-the-art results. However, transformer models often require significant computational resources, limiting their adoption in real-time clinical workflows.

Other studies have focused on improving loss functions to handle class imbalance, a common issue in medical datasets. Milletari et al. (2016) introduced the V-Net architecture with a Dice loss function, specifically designed to optimize segmentation performance for imbalanced data. Sudre et al. (2017) extended this work by proposing a generalized Dice loss, further improving segmentation outcomes in small datasets.

Evaluation Metrics and Challenges

Evaluation of segmentation models typically relies on metrics like the Dice similarity coefficient, which measures the overlap between predicted and ground truth masks. Studies such as emphasized the importance of robust preprocessing pipelines, including intensity normalization and spatial resampling, to improve model generalization. Despite advancements, challenges remain in segmenting low-contrast organs like the spleen, especially in the presence of noise or anatomical variability.

Applications in Clinical Workflows

Automated segmentation has shown promise in various clinical applications, such as radiation therapy planning (Heller et al., 2021) and trauma management (Menze et al., 2015). However, integration into real-world clinical workflows requires not only high accuracy but also computational efficiency. Studies have explored lightweight models like U-Net variants with reduced parameters to achieve real-time segmentation without sacrificing performance (Chen et al., 2020).

2.3 Summary

The evolution of segmentation methods, from traditional approaches to state-of-the-art deep learning architectures, highlights significant advancements in medical imaging. Early methods, though computationally efficient, lacked robustness and generalizability. Machine learning introduced better feature extraction capabilities, while deep learning transformed the field by enabling fully automated, end-to-end segmentation pipelines. Specifically, 3D U-Net and its variants have become the gold standard for volumetric segmentation. However, challenges such as class imbalance, low tissue contrast, and computational demands remain.

This research builds upon the extensive work in spleen segmentation by leveraging a 3D U-Net architecture, emphasizing robust preprocessing and efficient training. By addressing the limitations of existing methods, this study aims to contribute a reliable framework for automated spleen segmentation, with potential applications in pre-surgical planning and other clinical workflows.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The methodology section outlines the comprehensive steps undertaken to develop a robust spleen segmentation pipeline using a 3D U-Net architecture. It includes data collection and exploration, preprocessing strategies, model architecture design, and evaluation techniques. This systematic approach ensures reliable and accurate segmentation performance, addressing the challenges posed by anatomical variability and low tissue contrast in CT images.

3.2 Workflow

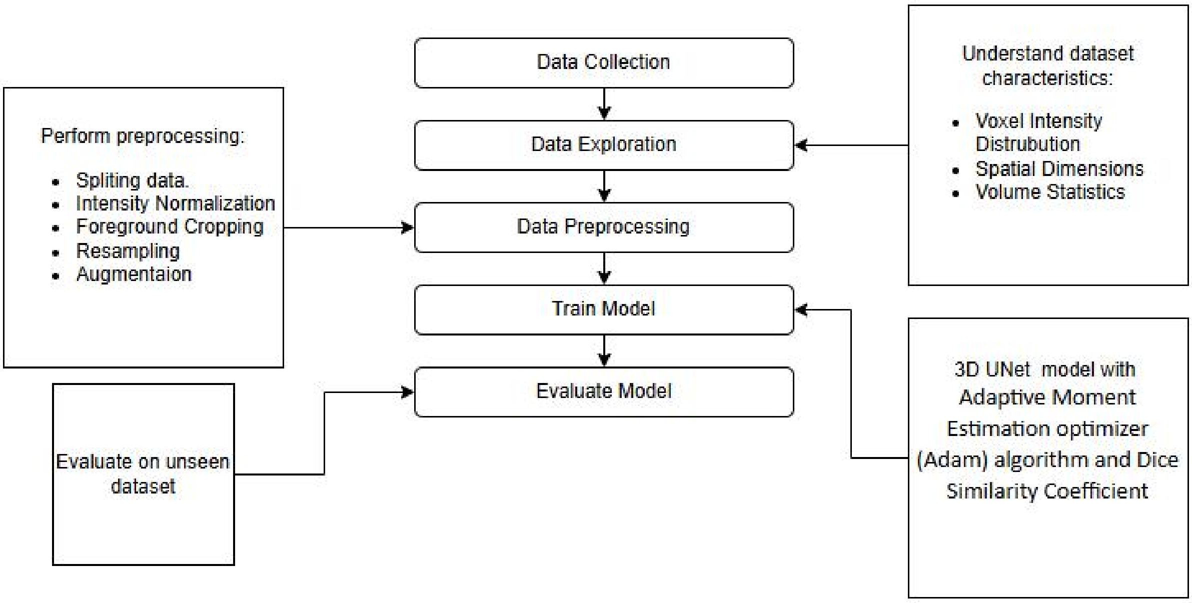


Figure 3.1: Workflow Diagram

3.3 Data Collection

3.3.1 Dataset

The study uses the Medical Segmentation Decathlon (MSD) Task 09 spleen dataset (Antonelli et al., 2021), a benchmark dataset comprising:

- **61 abdominal CT scans**, each annotated with spleen segmentation masks.
- The scans were collected from diverse institutions, providing variability in imaging conditions, voxel spacing, and contrast.

Data Annotation

Each scan was manually annotated by experienced radiologists.

The annotation process involved:

- Delineation of the spleen boundaries on each axial slice.
- Ensuring precise segmentation masks that align with anatomical structures.
- Multiple rounds of quality assurance to minimize errors and ensure consistency across the dataset.

The spleen masks are provided as voxel-level binary labels:

- 1 indicates spleen tissue.
- 0 represents background or non-spleen regions.

This rigorous annotation process ensures that the dataset provides high-quality ground truth for training and evaluating segmentation models.

3.3.2 Data Splitting

The dataset was divided into:

- **Training Set**: 49 scans (80%).
- **Validation Set**: 12 scans (20%).

This split ensured balanced representation of anatomical variability and imaging conditions for effective training and evaluation.

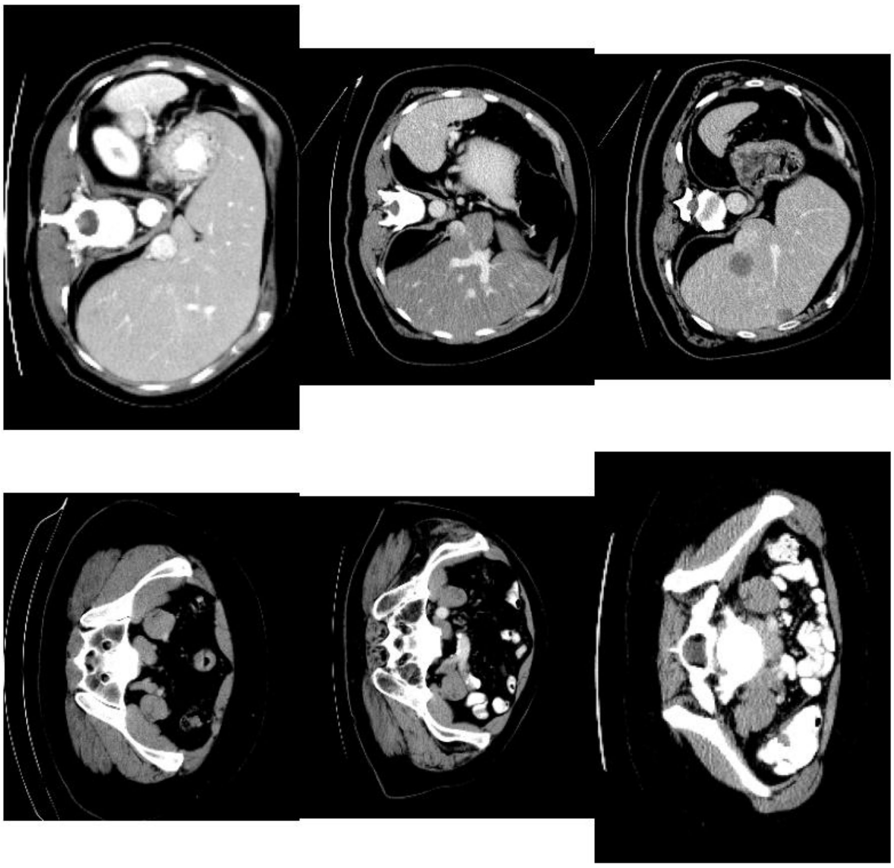


Figure 3.2: Sample Dataset

3.4 Data Exploration

To understand the dataset’s characteristics:

- **Voxel Intensity Distribution**: Analyzed intensity values to identify the typical range (-57 to 164 Hounsfield Units) and potential outliers.
- **Spatial Dimensions**: Explored variations in scan resolution and dimensions to prepare for resampling.
- **Volume Statistics**: Examined spleen volume distributions to ensure balanced representation in training and validation sets.

3.5 Data Preprocessing

Data preprocessing involved the following steps:

- **Loading and Orientation**: Volumes and labels were loaded using MONAI's LoadImaged transform and reoriented to the RAS (Right-Anterior-Superior) coordinate system.
- **Intensity Normalization**: Intensity values were scaled to the range [0, 1] using ScaleIntensityRanged, with values outside the range of -57 to 164 clipped.
- **Foreground Cropping**: Regions of interest around the spleen were extracted using CropForegroundd to reduce background noise.
- **Resampling**: Volumes were resampled to a consistent voxel spacing of (1.5, 1.5, 2.0) mm using bilinear interpolation for images and nearest-neighbor interpolation for labels.
- **Augmentation**: Random crops of size (96, 96, 96) were generated using RandCropByPosNegLabeld to include both spleen and background regions, improving model generalization.

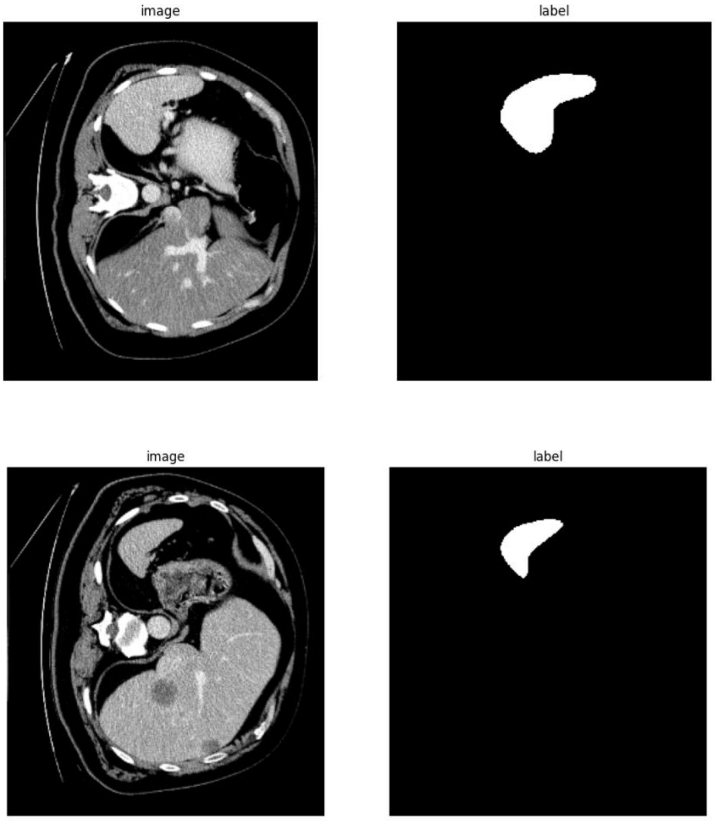


Figure 3.3: Preprocessed data

3.6 Training Model

3.6.1 Model Architecture

The proposed segmentation framework leverages the **3D U-Net architecture**, a fully convolutional neural network specifically designed for volumetric data segmentation. The model consists of an encoder-decoder structure with skip connections that integrate spatial and contextual information across multiple scales. This section provides a detailed breakdown of the architecture's components and their roles in the segmentation process.

Overview of the 3D U-Net

The 3D U-Net extends the original 2D U-Net architecture introduced by Ronneberger et al. (2015) to handle 3D volumetric data (Çiçek et al., 2016). The network comprises an encoder that captures high-level features by progressively downsampling the input volume and a decoder that reconstructs high-resolution segmentation masks through upsampling. Skip connections bridge the encoder and decoder at corresponding levels, preserving spatial information that may be lost during downsampling.

Key Components of the Architecture

Input: The model accepts 3D input volumes with dimensions (96,96,96), representing a cropped region of the CT scan. The input has a single channel corresponding to intensity values.

Output: The model produces a two-channel output, where each voxel is assigned a probability of belonging to one of two classes: background or spleen.

Encoder: The encoder consists of a series of **convolutional blocks**, each containing:

- **3D Convolutional Layers**: Extract features by applying 3D filters, which capture spatial patterns in the volumetric data.
- **Batch Normalization**: Normalizes activations to stabilize training and accelerate convergence.
- **ReLU Activation**: Introduces non-linearity to learn complex feature mappings.
- **Downsampling**: Performed using strided convolutions, reducing the spatial dimensions by a factor of 2 at each step. This allows the network to learn hierarchical feature representations at multiple scales.
- **Channels**: The encoder progressively increases the number of feature channels, starting from 16 and doubling at each level (16,32,64,128,256).

Decoder: The decoder reconstructs the segmentation mask by progressively upsampling the feature maps to the original input dimensions. Each decoder block includes:

- **Transposed Convolutions**: Performs upsampling to increase spatial dimensions while preserving learned features.
- **Concatenation with Skip Connections**: Combines high-level features from the encoder with the upsampled feature maps from the decoder to retain spatial details.
- **3D Convolutional Layers**: Refine the combined features to produce accurate segmentation masks.
- **Channels**: The number of feature channels decreases symmetrically with the encoder (256,128,64,32,16).

Skip Connections: At each resolution level, the feature maps from the encoder are concatenated with the corresponding upsampled feature maps in the decoder. This mechanism helps preserve fine-grained spatial information, enhancing the model's ability to delineate organ boundaries accurately.

Output Layer:

A final 3D convolutional layer generates a two-channel output:

Softmax Activation: Applied to produce voxel-wise probabilities for each class (background and spleen).

The result is a 3D probability map that represents the likelihood of each voxel belonging to a particular class.

Residual Connections

To improve gradient flow and accelerate convergence, each convolutional block employs **residual connections**. These connections add the input of a block directly to its output, enabling the network to learn identity mappings and mitigate the vanishing gradient problem.

Implementation Details

- **Framework**: The architecture was implemented using the **MONAI** library, which provides optimized modules for medical image processing.

- **Kernel Size**: All 3D convolutional layers use a kernel size of 3×3×3 to capture local spatial dependencies.
- **Padding**: Appropriate padding ensures that the spatial dimensions are maintained after each convolutional operation.

Visualization of the Architecture

Below is a illustration of the 3D U-Net architecture:

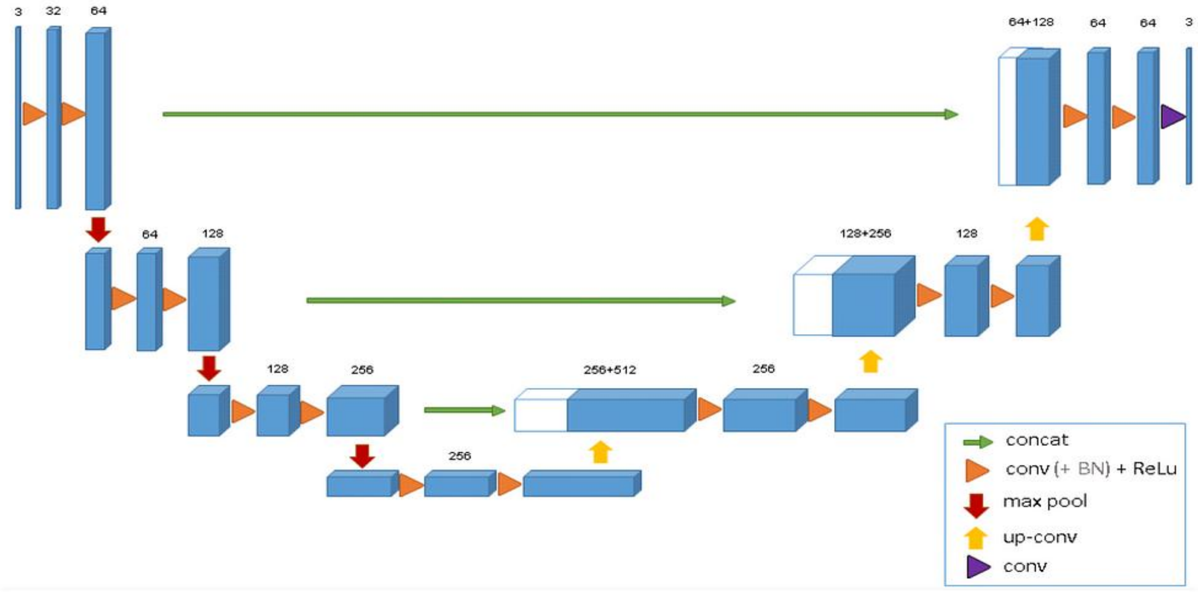


Figure 3.4: 3D U-Net Architecture

The encoder and decoder are symmetrically designed, with skip connections bridging corresponding layers at each resolution level.

3.6.2 Training

Key training configurations included:

- **Loss Function**: The Dice loss was used to optimize segmentation overlap:

$$\text{Dice Loss} = 1 - \frac{2\,.\,|P \cap G|}{|P| + |G|} \qquad (3.1)$$

Where P represents the predicted segmentation and G is the ground truth.

- **Optimizer**: The Adam optimizer was employed with a learning rate of 1×10−4 for stable and efficient optimization.
- **Hardware**: Training was conducted on an NVIDIA RTX 3060 GPU with 12GB VRAM.
- **Training Duration**: The model was trained for 600 epochs, with a batch size of 2, and validation performed after every epoch.

3.7 Model Evaluation

3.7.1 Evaluation Metrics

The Dice Similarity Coefficient (DSC) was the primary evaluation metric:

$$\mathrm{DSC} = \frac{2\,.\,|P \cap G|}{|P| + |G|} \tag{3.2}$$

Where P represents the predicted segmentation and G is the ground truth.

Training and validation Dice scores, along with loss curves, were used to monitor performance.

3.7.2 Visualization of Results

- Loss curves and Dice metric plots were generated to track performance.
- Predicted segmentation masks were visually compared with ground truth annotations to qualitatively evaluate the model's segmentation accuracy.

CHAPTER 4

RESULTS

4.1 Introduction

The results section presents the quantitative and qualitative evaluation of the 3D U-Net model for spleen segmentation using the Medical Segmentation Decathlon (MSD) Task 09 dataset. The model's performance was assessed based on metrics such as Dice similarity coefficient and validation loss, providing insights into its accuracy, generalization capability, and robustness. The results also include visual analysis to demonstrate the alignment between predicted segmentation masks and ground truth annotations.

4.2 Quantitative Analysis

4.2.1 Training and Validation Metrics

The model was trained for 600 epochs, with metrics recorded for both training and validation datasets. The Dice similarity coefficient (DSC) and average loss were used as primary performance indicators.

Training Results:

- **Final Training Loss**: 0.1718
- **Final Training Dice Similarity Coefficient**: 0.9582

Validation Results:

- **Final Validation Loss**: 0.0270
- **Final Validation Dice Similarity Coefficient**: 0.9494

Test Results:

- **Final Test Dice Similarity Coefficient:** 0.9483

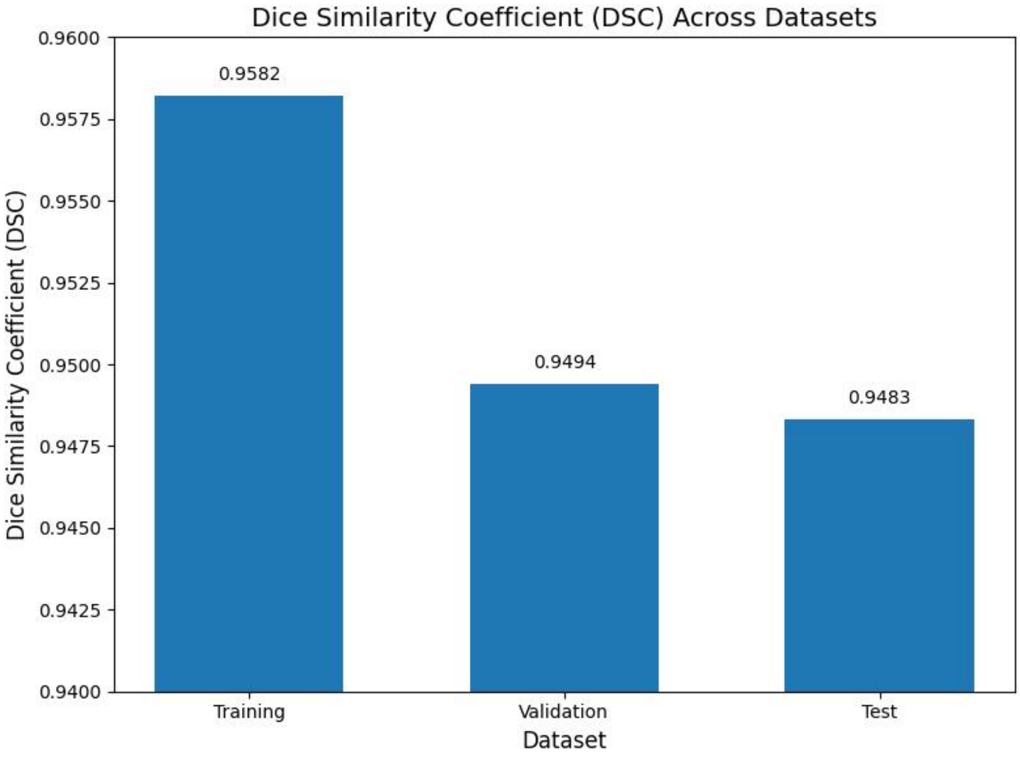


Figure 4.1: Dice Similarity Coefficient score across dataset

The high Dice scores for both training and validation datasets indicate that the model effectively captures the spatial relationships and segmentation details of the spleen. The low validation loss further validates the model's generalization ability to unseen data. Finally the model also performs well on unseen data.

4.2.2 Performance Trends

To monitor model performance over 600 epochs, training and validation loss curves and Dice metric plots were generated.

Loss Curves:

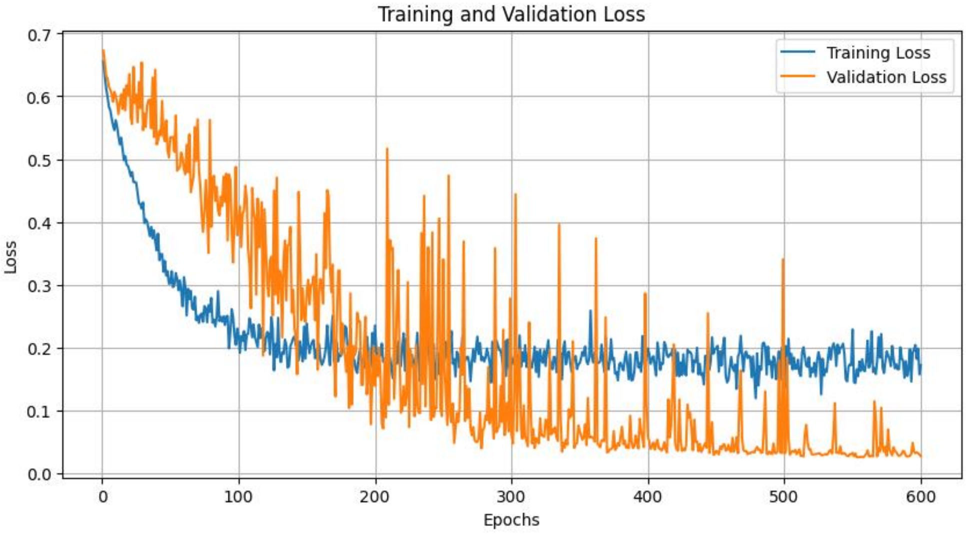


Figure 4.2: Training and Validation Loss Curves

Figure 4.2 shows the decline in both training and validation loss values. The training loss decreased steadily, demonstrating effective optimization. The validation loss stabilized after an initial decline, indicating convergence without overfitting.

Dice Metric Plots:

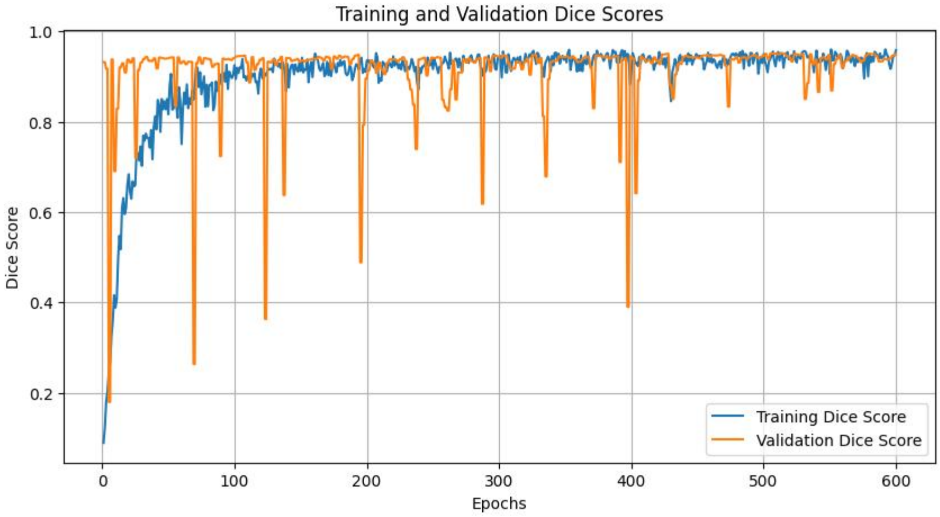


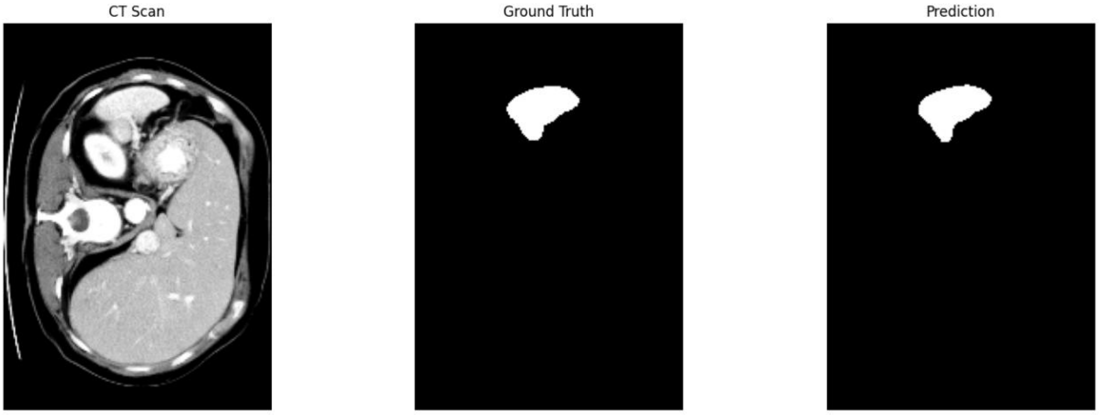
Figure 4.3: Training and Validation Dice scores

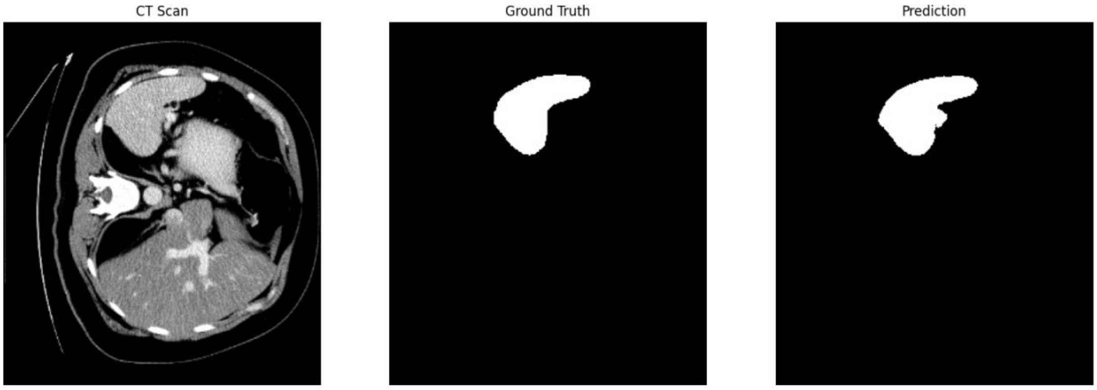
Figure 4.3 depicts the training and validation Dice similarity coefficients over epochs. Both metrics improved steadily, with the final values indicating strong segmentation performance.

4.3 Qualitative Analysis

To assess the model's segmentation accuracy visually, predicted masks were compared against ground truth annotations for multiple validation samples. The following observations were made:

- The predicted segmentation masks closely align with the ground truth annotations, even in regions with low contrast or challenging anatomical boundaries.
- The model successfully delineated the spleen's boundaries in most cases, with minimal misclassification of surrounding tissues.





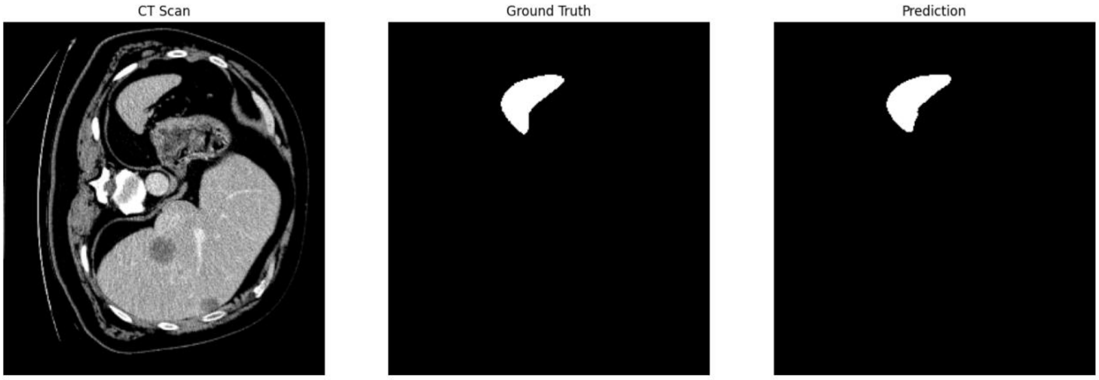


Figure 4.4:Comparison of Ground Truth and Predicted Segmentation Masks for a Test Sample.The left column shows the original CT slice, the middle column displays the ground truth mask, and the right column presents the predicted segmentation mask.

4.4 Discussion

The results demonstrate that the proposed 3D U-Net model achieves high segmentation accuracy for spleen segmentation. The high Dice similarity coefficients and low validation loss confirm the model's robustness and generalization capabilities. The qualitative results further validate the framework's effectiveness, with accurate boundary delineation observed across validation samples. Minimal discrepancies between predicted masks and ground truth annotations highlight the model's suitability for clinical workflows, such as pre-surgical planning and volume estimation.

Despite these promising results, certain challenges remain:

- **Misclassification in Low-Contrast Regions**: In a few cases, the model struggled to differentiate the spleen from adjacent tissues with similar intensities.
- **Computational Efficiency**: Training on 3D volumetric data is computationally intensive, requiring significant GPU resources.

Future work could address these limitations by exploring more advanced architectures, such as transformer-based models, or optimizing the current framework for real-time applications.

CHAPTER 5

CONCLUSION

5.1 Findings and Contributions

This thesis explored the application of a 3D U-Net architecture for spleen segmentation using the Medical Segmentation Decathlon (MSD) Task 09 dataset, addressing key challenges such as anatomical variability, low tissue contrast, and noise artifacts in computed tomography (CT) imaging. The proposed framework demonstrated robust performance, achieving a high Dice Similarity Coefficient (DSC) of 0.9582 for the training dataset and 0.9494 for the validation dataset after 600 epochs. These results indicate the model's strong ability to generalize across unseen data, which is critical for real-world clinical applications.

The use of advanced preprocessing techniques, such as intensity normalization, foreground cropping, and spatial resampling, ensured consistency across the dataset and enhanced the model's performance. The integration of random data augmentations improved the model's generalization capabilities. Furthermore, the use of Dice loss addressed class imbalance, which is a common issue in medical datasets (Milletari et al., 2016). The training process, supported by the Adam optimizer, facilitated efficient convergence while maintaining stability (Kingma & Ba, 2014). These methodological refinements contribute to the growing body of research on automated medical segmentation workflows.

In addition to achieving high segmentation accuracy, this study demonstrated the clinical relevance of the proposed framework. Accurate spleen segmentation has numerous potential applications, including pre-surgical planning, radiation therapy, and volume estimation. By automating this process, the proposed method reduces manual effort and minimizes inter-observer variability, enabling more reliable clinical workflows.

5.2 Significance and Implications

The findings of this study highlight the transformative potential of deep learning in medical image segmentation. Specifically, the proposed 3D U-Net framework offers:

- **High Accuracy and Generalization**: The robust performance metrics validate its applicability across diverse datasets, making it suitable for integration into clinical workflows.
- **Scalability**: The architecture can be extended to other organs and imaging modalities, paving the way for comprehensive automated solutions in medical imaging.
- **Clinical Utility**: By addressing key segmentation challenges, such as low contrast and anatomical variability, this framework can streamline clinical workflows and improve patient outcomes (Antonelli et al., 2021).

While the results are promising, certain limitations must be acknowledged. The model's computational requirements, driven by the processing of 3D volumetric data, may limit its real-time applicability. Additionally, the framework relies heavily on the quality of input data, and its performance may degrade in cases with severe noise or artifacts. Addressing these challenges is critical for broader adoption in clinical practice.

5.3 Limitations

- **Computational Complexity**: The model requires significant GPU resources for training and inference, which may hinder its deployment in resource-constrained environments.
- **Low-Contrast Challenges**: In a few cases, the model struggled with regions of low contrast between the spleen and surrounding tissues, leading to minor segmentation inaccuracies.
- **Dataset Specificity**: The model was trained and validated on the MSD Task 09 dataset, which may not fully represent the variability seen in clinical settings. Further validation on larger, more diverse datasets is required to ensure its robustness.

5.4 Recommendations for Future Work

To address the identified limitations and extend the contributions of this study, future research should focus on the following:

- **Multi-Organ Segmentation**:
- Extending the framework to segment multiple organs simultaneously could enhance its utility in comprehensive clinical workflows.

- **Integration of Transformer Models**:
- Recent advancements in vision transformers, such as Swin UNETR, could be explored to improve the model's ability to capture long-range dependencies in volumetric data (Hatamizadeh et al., 2022).
- **Real-Time Optimization**:
- Techniques like model pruning, quantization, and knowledge distillation could reduce computational requirements, enabling real-time deployment in clinical settings.
- **Multi-Modal Imaging**:
- Incorporating other imaging modalities, such as MRI or PET, could improve segmentation accuracy, particularly for challenging cases with low contrast in CT scans.
- **Clinical Validation**:
- Collaborations with healthcare professionals to validate the framework in clinical environments would provide critical feedback for refinement and practical deployment.

5.5 Concluding Remarks

This thesis underscores the potential of deep learning, particularly the 3D U-Net architecture, in advancing medical image segmentation workflows. The proposed framework achieves state-of-the-art performance for spleen segmentation, offering a robust and reliable tool for clinical applications. By addressing the limitations and exploring future directions, this work lays the foundation for the broader adoption of automated segmentation systems in healthcare. Ultimately, the integration of such technologies can enhance clinical decision-making, improve patient outcomes, and streamline medical workflows, marking a significant step toward the realization of AI-driven solutions in medicine.

REFERENCES

1. Antonelli, M., Reinke, A., Bakas, S., et al. (2021). *The Medical Segmentation Decathlon*. *Nature Communications*, 12(1), 1–13.
2. McInerney, T., & Terzopoulos, D. (1996). *Deformable models in medical image analysis: A survey*. *Medical Image Analysis*, 1(2), 91–108.
3. Kass, M., Witkin, A., & Terzopoulos, D. (1988). *Snakes: Active contour models*. *International Journal of Computer Vision*, 1(4), 321–331.
4. Criminisi, A., Shotton, J., & Konukoglu, E. (2013). *Decision forests: A unified framework for classification, regression, density estimation, manifold learning, and semi-supervised learning*. *Foundations and Trends in Computer Graphics and Vision*, 7(2–3), 81–227.
5. Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional networks for biomedical image segmentation*. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 9351, 234–241.
6. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). *3D U-Net: Learning dense volumetric segmentation from sparse annotation*. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 9901, 424–432.
7. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). *UNet++: A nested U-Net architecture for medical image segmentation*. *Deep Learning in Medical Image Analysis*, 3–11.
8. Hatamizadeh, A., Nath, V., Tang, Y., et al. (2022). *Swin UNETR: Swin transformers for 3D medical image segmentation*. *MICCAI*, 110–119.
9. Milletari, F., Navab, N., & Ahmadi, S. (2016). *V-Net: Fully convolutional neural networks for volumetric medical image segmentation*. *3D Vision (3DV)*, 565–571.
10. Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Jorge Cardoso, M. (2017). *Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations*. *Deep Learning in Medical Image Analysis*, 240–248.
11. Heller, N., Sathianathen, N., Kalapara, A., et al. (2021). *The KiTS19 challenge: Segmentation of kidneys in CT imaging for surgical planning*. *Medical Image Analysis*, 62, 101694.
12. Menze, B. H., Jakab, A., Bauer, S., et al. (2015). *The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)*. *IEEE Transactions on Medical Imaging*, 34(10), 1993–2024.

13. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2020). *DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848.

High-Precision 3D Spleen Segmentation for Surgical Planning Using Deep Learning

ORIGINALITY REPORT

19%	11%	12%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	1%
2	Mrutyunjaya Panda, Ajith Abraham, Biju Gopi, Reuel Ajith. "Computational Intelligence for Oncology and Neurological Disorders - Current Practices and Future Directions", CRC Press, 2024 Publication	1%
3	www.frontiersin.org Internet Source	1%
4	Submitted to University of Rijeka Student Paper	1%
5	Submitted to University of Bradford Student Paper	1%
6	Jinzhong Yang, Gregory C. Sharp, Mark J. Gooding. "Auto-Segmentation for Radiation Oncology - State of the Art", CRC Press, 2021 Publication	1%

7	"Medical Image Computing and Computer-Assisted Intervention – MICCAI 2017", Springer Nature, 2017 Publication	1%
8	Palo, Aitijhya. "Voxel Stress Optimizer (VSO) Framework for Residual Stress Mitigation in Wire-Arc Additive Manufacturing", Iowa State University, 2024 Publication	1%
9	www.isteonline.in Internet Source	1%
10	www.mdpi.com Internet Source	1%
11	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
12	link.springer.com Internet Source	1%
13	Ke Zhang, Chaoran Liu, Jielin Pan, Yunfei Zhu et al. "Use of MRI-based deep learning radiomics to diagnose sacroiliitis related to axial spondyloarthritis", European Journal of Radiology, 2024 Publication	<1%
14	research.vu.nl Internet Source	<1%
	web.archive.org	

15	Internet Source	<1%
16	keymakr.com Internet Source	<1%
17	Submitted to Asia Pacific University College of Technology and Innovation (UCTI) Student Paper	<1%
18	dokumen.pub Internet Source	<1%
19	ijisae.org Internet Source	<1%
20	Han, Shuo. "Fluorescence Microscopy Images Segmentation and Analysis Using Machine Learning", Purdue University, 2023 Publication	<1%
21	Yiren Wang, Yiheng Hu, Shouying Chen, Hairui Deng, Zhongjian Wen, Yongcheng He, Huaiwen Zhang, Ping Zhou, Haowen Pang. "Improved automatic segmentation of brain metastasis gross tumor volume in computed tomography images for radiotherapy: a position attention module for U-Net architecture", Quantitative Imaging in Medicine and Surgery, 2023 Publication	<1%

22	Mohan, Aswathy. "Enhanced Multiple Dense Layer Efficientnet.", Purdue University, 2024 Publication	<1%
23	Submitted to University Of Tasmania Student Paper	<1%
24	10wvvw.easychair.org Internet Source	<1%
25	Submitted to Berlin School of Business and Innovation Student Paper	<1%
26	Submitted to Radboud Universiteit Nijmegen Student Paper	<1%
27	Submitted to Texas A&M University - Corpus Christi Student Paper	<1%
28	"Medical Image Computing and Computer-Assisted Intervention − MICCAI 2017", Springer Science and Business Media LLC, 2017 Publication	<1%
29	Submitted to National Institute of Business Management Sri Lanka Student Paper	<1%
30	Submitted to Skolkovo Institute of Science and Technology (Skoltech) Student Paper	<1%

31	Delaram Kahrobaei, Enrique Domínguez, Reza Soroushmehr. "Artificial Intelligence in Healthcare and Medicine", CRC Press, 2022 Publication	<1%
32	linnk.ai Internet Source	<1%
33	researchspace.ukzn.ac.za Internet Source	<1%
34	www.ros-test.hw.ac.uk Internet Source	<1%
35	"Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016", Springer Nature, 2016 Publication	<1%
36	Gilmer Valdes, Lei Xing. "Artificial Intelligence in Radiation Oncology and Biomedical Physics", CRC Press, 2023 Publication	<1%
37	www.ijiemr.org Internet Source	<1%
38	"Image Analysis for Moving Organ, Breast, and Thoracic Images", Springer Science and Business Media LLC, 2018 Publication	<1%
39	"Scientific Abstracts and Sessions", Medical Physics, 2019	<1%

	Publication	
40	"Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges", Springer Science and Business Media LLC, 2018 Publication	<1%
41	climbtheladder.com Internet Source	<1%
42	da Silva, Maria Pedroso. "Novel Deep Learning Methods for Characterization of Precancerous Tissue in Endoscopic Narrow Band Images", Universidade do Porto (Portugal), 2024 Publication	<1%
43	etd.aau.edu.et Internet Source	<1%
44	iaset.us Internet Source	<1%
45	www.digitaljournal.com Internet Source	<1%
46	Saeid Asgari Taghanaki, Yefeng Zheng, S. Kevin Zhou, Bogdan Georgescu et al. "Combo loss: Handling input and output imbalance in multi-organ segmentation", Computerized Medical Imaging and Graphics, 2019 Publication	<1%

47	Konstantakopoulos, Dimitris. "Vulnerabilities and Robustness in Computer Vision.", University of Piraeus (Greece), 2024 Publication	<1%
48	eprints.gla.ac.uk Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		