3D Volume Reconstruction from Sparse 2D CT Slices Using Lightweight CNN & Attention

By

Shrabanti Dash Kana

(221-35-853)

Supervised by

Dr. Md. Fazla Elahe

Assistant Professor & Associate Head

Department of Software Engineering

Daffodil International University

A thesis submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Software Engineering

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

Fall – 2025

APPROVAL

APPROVAL

This thesis titled on "**3D Volume Reconstruction from Sparse 2D CT Slices Using Lightweight CNN & Attention**", submitted by **Shrabanti Dash Kana (ID: 221-35-853)** 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

------------------------------------ **Chairman**

Dr. Fazla Ealhe
Assistant Professor & Associate Head
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

------------------------------------ **Internal Examiner 1**

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

------------------------------------ **Internal Examiner 2**

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

------------------------------------ **Internal Examiner 3**

Md. Rajib Mia
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

------------------------------------ **External Examiner**

Mohammod Abul Kashem, PhD
Professor
Department of Computer Science and Engineering
DUET, Bangladesh

DECLARATION OF THESIS AND COPYRIGHT

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	:	Shrabanti Dash kana
Date of Birth	:	25-01-2002
Title	:	3D Volume Reconstruction from Sparse 2D CT Slices Using Lightweight CNN & Attention
Academic Session	:	Fall 2025

I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
☐	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:

Shrabanti

(Student's Signature)

Student ID 221-35-853
Date: 27 November 2025

Fazla Elahe 27.11.25

(Supervisor's Signature)

Dr. Md. Fazla Elahe

Name of Supervisor
Date: 27 November 2025

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.

Fazla Elahe 27.11.25

____

(Supervisor's Signature)

Full Name : Dr. Md. Fazla Elahe

Position : Assistant Professor & Associate Head

Date : 27 November 2025

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.

Shrabanti

(Student's Signature)

Full Name : Shrabanti Dash Kana

ID Number : 221-35-853

Date : 27 November 2025

3D Volume Reconstruction from Sparse 2D CT Slices Using Lightweight CNN & Attention

By

Shrabanti Dash Kana

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

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

Fall- 2025

ACKNOWLEDGEMENTS

I wish to start by expressing my most sincere appreciation and deepest gratitude to God, the Most Gracious and the Most Merciful, for providing me with the strength, patience, and wisdom to complete this thesis. God, the eternal refuge, guided me throughout this journey, and I am eternally grateful for His infinite grace.

I would like to extend my heartfelt thanks to Dr.Md Fazla Elahe sir, my highly respected supervisor, for his unrelenting support and guidance throughout this academic undertaking. His precious advice and insightful suggestions have encouraged me to think critically and pursue research with dedication and passion. It has been an honor to be guided by an individual who has contributed so much to my academic and professional development.

I am thankful to Mr. Md. Shohel Arman sir , Assistant Professor of Software Engineering Department, for his continuous support and necessary guidance throughout this research. His encouragement and accessibility have been a continuing source of motivation.

In conclusion, I wish to record my deepest appreciation to my family—my father, mother, siblings, and friends—for their love, prayers, and unwavering support. Their belief in me has been the foundation for my perseverance and success.

The whole of this thesis is dedicated to the dearest parents, who have given me limitless support and sacrifices, believing in me. I owe my progress to their steadiness and commitment, and I am indebted to their continued belief in me for all my achievements.

ABSTRACT

Sparse slice acquisition is used to a great extent in the low dose CT imaging. This assists in reduction of radiation. Nonetheless, it is likely to damage the quality of 3D reconstruction significantly. The resultant artifacts and noise are manifested. There is some evidence of a lightweight 2D CNN auto encoder, which consists of channel attention. The purpose of this set up is to enhance the quality of the sparse CT slices. It reinstates important structural information. Noise gets reduced too. All this is prior to the slices entering the entire volumetric process. Linear interpolation is used after the reconstruction process. This forms a full volume in 3D. The continuity across the slices is enhanced in the process. Experimental reviews point to high performance in this. The average SSIM of the model is 0.9660. Average PSNR comes in at 25.99 dB. These figures point to good structural fidelity in the general. Distortion is also rather low. Such outcomes are indicative of the usefulness of the lightweight approach. It is effective in efficient sparse data 3D CT reconstruction.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

Computed tomography has become an important aspect in the modern healthcare facility over the past few decades. The scans provide the clear three dimensional pictures of the organs within the body. The fact is crucial to track down diseases, observe their progress, and establish a specific treatment (Dalsaniya et al., 2025). Nevertheless, the rapid proliferation of artificial intelligence and deep learning is sweeping all over the world has pushed their contribution to medical imaging to an even greater extent.Regardless of all the advancement, routine CT scans continue to subject patients to ionizing radiations. The attempts to reduce the number of those risks, over time, have highlighted low dose techniques in the clinics and regular examinations. The one method which is particularly noteworthy is sparse view CT in which fewer projections are utilized (Li et al., 2025, Zhang et al., 2025). The negative aspect is reflected in worse images but, there is additional noise and streaks and details on buildings being missing. This may compromise the reliability of the diagnosis that will result (Wu et al., 2025). To add it all the CT machines that are cheaper or older have a tendency of producing thicker slices. These problems have increased the accuracy in activities such as the measurement of lung nodules that are closely measured (Kim et al., 2025).These problems have increased the need to develop improved techniques of reconstructing three dimensional images of visual data. This aids the development of computer vision models and the broad AI advancements (Zhou et al., 2025). The sparse view underdetermined geometry is usually a challenge to the usual rebuilding algorithms. That is why research workers resort to deep learning fixes. Convolutional neural networks can provide actual opportunity on this matter, addressing the challenges of inverse problems such as sparse CT reconstruction and increasing image resolution. However, CNN setups that are typical do not have the range of capturing far-off connections or larger contextes particularly when large training sets are involved (Gang, 2025).Adding attention capabilities are, as of late, a means to get over those challenges in new fashions. These blocks allow models to hone on important input areas. That enhances feature pulls and rebuild quality in general. The designs such as the NestedMorph (aligning images) or UNETR++ (three dimensional segmentation) enable the reader to see the effectiveness of attention (Kumar et al., 2025, Shaker et al., 2024).In this context, the new setup proposed by the current work is right. It integrates channel attention into a thin CNN frame to deal with sparse view CT rebuilding. The objective remains on the enhancement of appearance of thin CT slices using an efficient twodimensional model. Based on this, good quality three dimensional volumes are through straight forward interpolation. This system may facilitate the process of low radiation routines in clinics. It could even accelerate the diagnostics process making the CT practices safe and smooth in the long run.

1.1.1 The Rapid Advancement and Adoption of AI in 3D Medical Imaging

The rapid diffusion of mobile communication technologies all over the globe provides an analogy to the ease with which high-performance graphics processing units and large datasets have become so accessible, triggering the development of 3D medical image analysis. This, in its turn, has contributed to the significant increase in the demand of vision-based 3D reconstruction techniques over the past few years. These techniques play a crucial role in such fields as autonomous systems, robotics, and precision medicine (Zhou et al., 2024). According to a recent review, AI-assisted 3D reconstruction continues to pick up. It enables the generation of detailed tomography of the human anatomy. Such a detail is necessary to achieve a more accurate diagnosis (Dalsaniya et al., 2025). These advancements create actual opportunities to both healthcare providers and AI tool creators. In the case of analyzing complex anatomical aspects, researchers and clinicians are increasingly looking at advanced approaches to deep learning. They assist in overcoming the drawbacks visible in older CNN solutions.All over the world the most advanced research facilities, equipment manufactories, and healthcare facilities have introduced different AI applications. The intention is to enhance services in areas including monitoring illnesses, training about health, and planning complex surgical operations (Shaker et al., 2024). The richer nations have greater prevalence of 3D medical imaging with AI usage. Nevertheless, the value is now beginning to be felt in places that are not as endowed. They strive to build these tools into schemes of improved healthcare.Architectures such as attention mechanisms and Transformer models spread like wildfire these days. They are based on the ability to extract long-range links and wider contexts in 3D data sets. An example is Take NestedMorph. It clearly show image registration improvements with its application of layered attention blending (Kumar et al., 2025). UNETR++ models work well on 3D segmentation work too (Shaker et al., 2024). It begins to prove its value in medical imaging, as well, with designs that do not have convolutions. This indicates that there is a rapid shift of the field towards more powerful constructions (Gang, 2025). The introduction of these digital tools into the medical imaging sector serves larger purposes such as the Sustainable Development Goals. It is connected to SDG 3 on health and well-being. That has the goals of universal coverage of less than 3.8 and enhancing research on medicines with 3.B. AI methods to reconstruction to make diagnostics affordable and reliable. Ultimately, they have a major role to play in creating equitable healthcare in the world.

1.1.2 Specific Challenges in 3D Reconstruction from Sparse 2D CT Slices

The techniques of deep learning have achieved significant advances in sparse-view computed tomography reconstruction. Such gains are mentioned in studies by Wu and others in 2025, and Li and others in the same year. Nevertheless, there is a special challenge of constructing a complete three-dimensional volume with limited two-dimensional cut sections. This issue usually appears in clinical CT scans. There, a thicker cross-section will reduce radiation or solve equipment limitations. These techniques slice through the accuracy of the total volume. They also complicate the detection of small abnormalities that are important in the diagnosis. These limitations are indicated by evidence in Kim et al. 2024.A number of important obstacles are presented as the construction of a three-dimensional structure using sparse two-dimensional slices. A significant problem is the sparsity of data and the information gaps. Fatter or thinner slices imply huge losses in the direction of depth. It requires super-resolution techniques or how one can approximate the missing portions to fill in the gaps. Linear or cubic interpolation are old methods that introduce undesired effects which are distortions between slices. Another issue correlates with the robust requirements of three-dimensional models.They find it difficult to restore minute details, which characterize structures. Three-dimensional convolutional neural networks would allow confronting the sparsity at once. However, these networks consume much processing power. They must also have ample resources in graphics processing unit. In three-dimensional data, distant connections are well addressed in transformer designs, such as UNETR++ by Shaker et al. in 2024 and NestedMorph by Kumar et al. in 2025. It is hard to train and run them, however, in systems with hardware constraints.Fast models such as X-LRM, created by Zhang et al. in 2025, rely excessively on raw X-ray projections. That renders them inappropriate in the rebuilding of straight data straight off the two-dimensional CT. To top that, the three-dimensional methods do not possess a high level of flexibility in contrast to two-dimensional methods. They require larger and diversified training datas in order to be good in different cases. They become less apt to manage new situations with such data being so scarce. This weakness is highlighted by the work done by Gang in 2025.The present work proposes a two-stage approach to address these problems. The initial phase runs a basic autoencoder two-dimensional convolutional neural network. This architecture bypasses the three dimensional networks that are resource intensive. It has a channel attention module. That assists the system to narrow down to the important aspects in each slice. The general idea of this step is to enhance the clarity of the slices of the input by a significant margin, in the second step, the entire volume of three dimensions filled by the slices is completed using basic linear interpolation. This combination has been able to combine low computing requirements with good results in reconstruction. Ultimately, it provides a feasible alternative to rapid checks within the medical practice. Costs stay reasonable too.

1.2 Problem Statement

1.2.1 Problem Background

The studies indicate that 3D volume reconstruction plays an important role in modern medical imaging. It helps to diagnose accurately and plan a treatment, which is indicated in recent research (Dalsaniya et al., 2025). However, clinical practice typically is based on sparsely-sampled CT or thicker 2D cuts. These decisions are intended to reduce patient radiation or are due to the limitation of hardware in scanners (Li et al., 2025; Kim et al., 2025). All these approaches cause significant challenges to the general workflow.

1.2.2 The Core Problem

Attempts to construct a fine volume representation of the 2D CT slices have an underdetermined inverse challenge. The problem is caused to a great extent by the lack of information on the Z-axis between remote slices. Plastic anatomical characteristics merely disappear in such spaces.

1.2.3 Information Loss and Artifacts

Take the case of only a fraction of the slices, say every fourth slice, being utilized in the entire volume construction. The result is likely to be heavy artifacts and noise. This causes a decrease in volumetric accuracy. This has an impact on other important clinical tasks, such as measuring nodules in the lungs, where small structures are very important (Kim et al., 2025).

1.2.4 Limitations of Existing Solutions

Conventional interpolation schemes such as linear or cubic schemes can cope with the interpolation between sparse slices. Nevertheless, they have difficulties in reinstating complex anatomy. Blurring is present, as well as inconsistencies and distorted shapes in the filled regions.More sophisticated ones are based on deep learning, including 3D convolutional-neural networks, vision transformers, and neural-radiance-field-based methods. These include UNETR++, NestedMorph and 3DGR-CT. They have performed better in tests (Shaker et al., 2024; Kumar et al., 2025; Li et al., 2025). Despite this, disadvantages abound in all of them.First in line is excessive computational requirements. The training stages as well as real-time applications require a lot of GPUs and performance.Attention modules in transformers further aggravate the problem. In three-dimensions, their complexity increases quadratically. This impairs the use in real-world clinics (Gang, 2025).Some state-of-the-art models of sparse-view CT also rely on the initial data of raw X-ray projections (Zhang et al., 2025). That arrangement is a stark contrast to the existing emphasis. In this case the objective is to build 3D volumes out of the sparse 2D slices that are available.

1.3 Research Gap

The problem of building up high quality three dimensional volumes based on sparse two dimensional computed tomography slices still remains a challenge in medical imaging. The development of deep learning techniques on three dimensional reconstruction has improved. However, they are still limited by a number of limitations that do not allow them to expand to clinical use. Studies establish gaps in this field. The latest models such as three dimensional convolutional neural networks, vision transformers, and neural radiance fields based models are state of the art, and they provide good accuracy. These models are based on operations like quadratic self attention which requires huge amount of computing power and memory. There is evidence that these demands render them inapplicable to the daily use in clinics that have underdeveloped hardware. It has been established that there is a strong requirement to simplify two dimensional architectures that execute well and provide accurate three dimensional output using fewer slice data. The classical interpolation techniques, such as linear and cubic, can cope with the slices gap between slices. Nevertheless, in many cases, they are unable to regain the fine anatomical features, which disappear in low sampling. This inadequacy causes blurred pictures, undesired artifacts, and discrepancies between the volume. There seems to be no strong data based ways available to perfect the sparse slices, before interpolation, to improve the quality of the entire reconstruction. In most three dimensional reconstruction frameworks, attention mechanisms are an important component. They are still not applied in lightweight models two dimensional and specific to slice enhancement. Multi scale fusion attention techniques do not normally deal with sparse slices, but projection data or complete two dimensional outputs. It is proposed in the research that an efficient two dimensional autoencoder with channel attention is developed. It is possible to have such a system with prioritization of important characteristics in the slice level and enhance sparse inputs of computed tomography before moving to three dimensional assembly.

1.4 Research Objective

The proposed research is aimed at addressing the problem of high-computational complexity and low accuracy of outcomes in the reconstruction of 3D volumes based on a set of few 2D CT images. The general idea behind this is to propose a two-step attention-focused strategy that is quite simple. In that manner it is able to develop detailed full 3D buildings at a fast rate. And with fewer errors too. The most important focus which is to design a convolutional neural network that is fast and slim. This configuration increases the specificity of these thin 2D CT layers. Afterward it assists in solid 3D rebuilding by simple filling-in techniques. Regarding the development of the model and its configuration, it is planned to develop a simplistic 2D autoencoder model. Besides, there is also the addition of a channel attention component within the encoder, which serves to draw out superior details of the features. To manage data and prepare the model, 3D CT scans volumes are converted to even sparse arrangements. Something such as putting the step to four to train. The model will be trained in a manner that it will provide clear 2D layers that have been rebuilt by those gaps. The refinement of the 2D layers implies the training of the setup on each sparse layer. It eliminates noise and reduces defects. The layers that are upgraded seal the weak points in unenhanced filling techniques (Kim et al., 2025). The entirely reconstitution of the 3D shape occurs through the straight-line filling of the cleaned sparse layers. Testing its effectiveness is dependent on standard measures. There are differences in mean squared error. The structural similarity index monitors similarity. Sharpness is measured by the peak signal-to-noise ratio. With regards to speed and its performance, there is evidence suggesting that the slim 2D attention model can run on less power with ease. It is piled on the top of 3D attention systems such as UNETR plus plus or NestedMorph. Few parts, less memory used, faster run times. Everything that qualifies it to cramped spaces inside a medical facility (Shaker et al., 2024; Kumar et al., 2025).

1.5 Motivation

This work is motivated by the necessity to identify a means of coping with two important pulls in the clinical diagnostics. On the one hand, there is patient safety. On the other hand, accuracy in diagnosing must remain high. Low-density CT reduces radiation hazards. However, it destroys the quality of images. This necessitates build up techniques that retain credible clinical outcomes. Clinical needs also require maintaining low requirements in computing. Reduction of exposure is important in screening routine systems. Well being of patients remains central there. To assist, workflows are directed to sparse-view CT, or towards thicker 2D slices (Li et al., 2025). But issues still arise with such steps.Image quality suffers at the hands of sparse slices. Noise builds up. Artifacts show more. This muddies visual details. Diagnosis becomes more difficult to interpret (Wu et al., 2025).Thicker slices are also problematic to volumetric precision. Both the shapes of organs become sharp. Take pulmonary nodules. They are very useful in the detection of lung cancer. But Z-axis details fade. This damages evaluation (Kim et al., 2025).The primary impetus in this direction is to repair damaged sparse-slice CT images. The approach must enable restoration of quality. It must retain or increase diagnostic acuity. Safety of patients is paramount. Rebuilt volumes remain practical in clinics.Deep learning is taking a strides in medical image rebuilding in the recent past. 3D Transformer designs dominate. They show solid results. However, there are efficiency performance anomalies.Best models such as UNETR++ are problematic (Shaker et al., 2024) NestedMorph is problematic (Kumar et al., 2025). 3D self-attention is quadratically demanding. Training eats GPU power. Inference does too. Struggling clinics have thin resources. The developing regions are the most victimized.Full 3D systems consume efforts in sparse 2D initializes. Complicatedness is cumulative. This is where resources are not utilized as much.This research bridges this efficiency gap. Channel attention helps to improve a light 2D autoencoder. Combine that with simple linear interpolation. Results match heavy 3D models. No additional load is carried with it.Old linear interpolation is no good on sparse slices. It just blends pixel values. Fine structures remain lost.The plan develops a two-part structure of rebuilding. Initial stage sharpens those 2D slices which are sparse. It is dealt with lightly by the 2D autoencoder. The channel attention brings attention. Quality lifts. Formal features are made more explicit. Second stage applies the linear interpolation on the enhanced slices. The entire 3D volume develops fast. There is no quality degradation.The configuration is high-speed. It scales well. Benefit of routine clinic activities. Seven-resource locations are now capable of handling high-quality rebuilds.

1.6 Contribution

Researchers have proposed a new two staged reconstruction environment setup. The efficiency is achieved in obvious ways by eliminating the resource intensity of 3D convolutional neural networks or transformer models that consume resources. It has been demonstrated that even a plain autoencoder with basic interpolation can be utilized in the 2D space to deal with the constraints of those bulky, memory-intensive 3D architectures, such as UNETR++ or NestedMorph (Shaker et al., 2024; Kumar et al., 2025).Operating in the 2D space makes everything light. The parameters reduced significantly. This is because inference is accelerated significantly and loss of fineness structures follows. Those details are often lost to naive interpolation. The solution to that problem is to add a channel attention block to the 2D autoencoder.The attention mechanism is executed in the encoder section. It selects important feature maps of every thin slice. This occurs depending on the data itself.The reduction of noise and artifacts occurs prior to the interpolation stage. That will prevent the quality decline of thicker cuts. Ultimately, the complete 3D volume is a more robust (Kim et al., 2025)work becomes bright in practice.In the case of real-life application, when resources are limited, the technique is effective. It is a good balance between clinical settings and research laboratories. Reconstruction does not require excessive power to remain correct. Even simple equipment can generate solid diagnostic results.Another advantage is that of radiation exposure. The low-dose options are also feasible when high-quality 3D images are pulled out of sparse CT scans. On a figures level, accuracy is high.On the patient safety front, the general practice is accurate. The strategy is not only faster but also more effective than the older techniques.Normal measures prove this. Such as the gains of SSIM on structural similarity and PSNR on noise levels, SSIM scores are higher than plain interpolation scores. The volumes retain their anatomical details.

1.7 Thesis Organization

This thesis is organized into the following chapters:

1.7.1 Chapter 1: Introduction

In this chapter, an overview of the study has been given, which covers the background, the problem statement, the research objectives, and the relevance of the study. It preconditions the perception of the context and applicability of 3D Volume Reconstruction of Sparse 2D CT Slices.

1.7.2 Chapter 2: Literature Review

This chapter is the review of the literature that is available in the field of the research. It discusses theoretical frameworks, past researches, main findings, and the gaps in literature that dictate the necessity of the present research.

1.7.3 Chapter 3: Research Methodology

In this chapter, the author presents a refined overview of the two stage lightweight attention based 3D reconstruction approach that is proposed. It is broken down into several major parts which follow one another sequentially. Discusses the general outline of the framework. This is where the whole arrangement is considered at an elevated level. Thin slice 2D slices are then refined to extract finer quality features in two dimensions. These features are then fed into the rebuilding the entire three dimensional volume process. The last step is taken up by linear interpolation. Manipulations of datasets and pre processing. The datasets in the study are outlined in Subsection 3.2.1. Public sources are the primary sources. Types such as CT scans will contain information such as voxel resolution and slice thickness which is significant. This type of research is characterized by the presence of public CT datasets. Describes the process of producing sparse 2D slices out of whole 3D volumes. As an example every fourth slice could be selected in order to simulate sparse conditions. Such a configuration mimics low dose scans or fewer acquisitions. Such situations demonstrate the actual difficulties of reconstruction works. Switches to the light two dimensional auto encoder architecture. The encoder and decoder blocks are outlined in subsection 3.3.1. There structure is built of layers. Those components occur as feature extraction. They collaborate to enhance the meager slices. Talks about channel attention block. The encoder part is joined by the attention mechanisms which are concerned with functionality that fits right in. They focus on the main aspects of each thin slice. Such concentration enhances the quality of input in general. It prepares things with the interpolation stage. Describes the process of training.

The loss that is used is described in subsection 3.4.1. MSE with SSIM may do the trick in this case. The objective remains on the balancing of accuracy in reconstruction. Similarity on structure is optimized as well. Gives the details of the various hyperparameters and optimizer. Such options include the type of optimizer itself. Second is learning rate and third is batch size. The main settings are completed by the number of epochs. Other tweaks fill in the rest. Does the three dimensional reconstruction and post processing. Here, linear interpolation is applicable to the improved two dimensional slices. That restores entire bulk. This is followed by post processing to refine the results. The final touches are high quality.

1.7.4 Chapter 4: Data Analysis and Results

This chapter is a report on findings of the experiments or analysis. It contains specific conclusions with tables, graphs and figures. The chapter describes the work of the proposed methods and indicates the key results of the study.

1.7.5 Chapter 5: Discussion

This chapter explains the results with reference to the research questions and available literature. It discusses its implications in terms of theoretical, practical and policy implications, the importance of the findings and how the results are compared with other previous studies.

1.7.6 Chapter 6: Conclusion and Recommendations

This chapter is described to reflect the main findings of the research, their general contribution, and their recommendations towards future studies, practitioners, and policy makers. It also summarizes the shortcomings of the research and gives possible future research directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Consequently, to develop more insight on the use of 3D volume reconstruction of sparse 2D CT slices, there is need to do pertinent and rigorous researches. The general aim of this literature is to provide an in-depth understanding and a survey of theories and models applied in the field of medical image reconstruction and analysis. This literature is subdivided into three parts: developing reconstruction and attention models, systematic literature review of particular methodologies, and identifying the gaps in research about lightweight architectures(Zhou et al. 2024) point out that, with the further evolution of deep learning and the use of the GPU technology, the need in high-precision and high-efficiency 3D reconstruction information is growing, especially in the medical field. (Dalsaniya et al. 2025) also claim that MRI and CT-assisted reconstruction with the aid of AI is central to visualization of organs. Nevertheless, (Wu et al. 2025) observe that sparse-view CT reconstruction is a tricky ill-posed inverse problem, where the lack of projection data results in low image quality with more noise and a higher number of artifacts.

2.1.1 Dual-Domain and Attention-Based Models

Deep Learning (DL) has appeared as a major solution in the context of medical imaging in accordance with the classical principles of Iterative Reconstruction (IR). Nevertheless, Wu et al. (2025) stated that the current techniques do not always consider constraints of projection data and are very dependent on Convolutional Neural Networks (CNNs), which leads to a lack of feature extraction. Gang (2025) argues that CNNs are weak at capturing long-range dependence and global context which restricts their ability to analyse fine and complex structures. Thus, based on the summary of the theories linked to deep priors and attention processes, Wu et al. (2025) developed the DPMA (Dual-domain deep Prior-guided Multi-scale fusion Attention) model. It is a combination of residual regularization methods and multi-scale fusion attention. It mainly concerns itself with the parallel modeling of global context, regional dependencies and local details in a concerted framework in an effort to overcome the shortcomings of sparse-view CT reconstruction. As a matter of fact, attention mechanism integration has been proven to be valid in many medical imaging fields. As an example, Kumar et al. (2025) introduced NestedMorph, a network that applies a Nested Attention Fusion strategy to enhance deformable registration, which involves high-resolution spatial data and semantic data. Shaker et al. (2024) introduced UNETR++, an Efficient Paired Attention (EPA) block with a long-range dependency representation, with a lower computational complexity. The point at which the attention mechanisms and efficient architectures meet is, therefore, our preferred theoretical basis of this review.

2.1.2 Research on Sparse CT Reconstruction and Lightweight Architectures

Thin-Section CT Reconstruction The advent of the modern algorithms has made sparse-view CT a useful method to reduce radiation exposure. A dual-domain model was studied (Wu et al. 2025) and it demonstrated that the effectiveness of noise suppression with the help of deep learning-based priors and model-based optimization is effective. It can be concluded that DPMA model improves the accuracy of reconstruction and maintains data consistency. A research done by (Li et al. 2025) has shown that a 3D Gaussian Representation (3DGR) based algorithm can be utilized successfully in sparse-view CT reconstruction. They also use the power of 3D Gaussian splatting instead of implicit neural fields to deal with the noise and artifacts of small data. Besides, (Zhang et al. 2025) suggested the X-ray Large Reconstruction Model (X-LRM) to reconstruct CT with extremely sparse-view (less than 10 views). It has been shown that X-LRM is an X-former, X-triplane representation, which is better in reconstruction quality, 1.5 dB, compared to state-of-the-art techniques. The Role of Attention Mechanisms In addition, the mechanism of attention is massively used to explain the feature extraction process in medical images. Gang (2025) states that all-convolutional-free model with transformer architecture and self-attention mechanism is able to overcome CNNs shortcomings in global contextual capture. The current research finding proves to agree with (Wu et al. 2025) that also utilized a multi-scale fusion attention mechanism to achieve the representation of global context and local details through parallel pathways. However, (Kumar et al. 2025) employed nested attention mechanisms, as opposed to pure CNN methods, to complement local and global feature extraction in the deformable image registration. Shaker et al. (2024) presented the evidence that the self-attention operation of conventional transformers is quadratically complex, which turns out to be a computation bottleneck in volumetric medical imaging. To solve this they proposed the Efficient Paired Attention (EPA) block that learns effectively the spatial and channel-wise features. Lightweight and Efficient Architectures In addition, the issue of computational efficiency is a decisive characteristic that determines the use of 3D reconstruction models. According to (Shaker et al. 2024), UNETR++ is promising and consumes less memory and model complexity than other approaches do. They stress in their work that the method of spatial attention formulation is linearly complex with reference to the length of the input sequence. In a study conducted by (Zhang et al. 2025), the authors discovered that the X-LRM model is 27x faster than other similar approaches with the help of the MLP-based image tokenizer and Transformer-based encoder. Moreover, (Li et al. 2025) also used the effectiveness of 3D Gaussian representation to enhance speed of reconstruction with sparse views. Image Quality and Volumetric Enhancement However, a number of researchers provide strong evidence on the effect of slice thickness and resolution on image quality. To create a thin-slice CT image, (Kim et al. 2025) developed a deep learning-based super-resolution model to estimate a thin sliceCT image given a thick sliceCT image. According to their research observations, the categorization of pulmonary nodules was more accurate when the images of the thick-slices were transformed into those of the thin slices. Huang et al. (2025) also presented H3DE-Net, which is effective and precise in detecting 3D landmarks, and it is imperative to note that proper anatomy detection is crucial in the next medical imaging procedure. The challenges of domain adaptation between thick and thin slice CT images were also discussed by Gang (2025) with the use of a joint loss.

2.2 Limitations and Research Gaps

The above study results have provided various perspectives of explaining the process of 3D reconstruction and analysis. However, the recent literature has some limitations. The authors (Wu et al. 2025) observed that the current practices tend to lead to the lack of appropriate adaptability caused by over-reliance on CNNs. As (Shaker et al. 2024) highlighted, their model has a good performance, but it is not effective in separating organs of irregular geometry or thin outlines because of small training datasets. Moreover, (Gang 2025) pointed out the current issue of multi-semantic segmentation of thin slices data by referencing datasets. The gap is the absence of a unified lightweight framework that is able to support extremely sparse views, and considerably high computational costs at the same time. The current paper takes the perspective of incorporating lightweight attention schemes to study the controlling factors that influence the performance and accuracy of 3D volume reconstruction using sparse data.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The chapter describes the entire framework of the methodology of 3D CT volume reconstruction based on sparse 2D slices with a lightweight U-Net channel-attention-enhanced model. The full pipeline is based on the adopted structure of the code and consists of dataset preparation, slice preprocessing, simulated sparse-view, model design, training process, and 3D final reconstruction. The main goal of such a methodology is to synthesize absent CT slices of highly sparse axial scans with minimum computational complexity. This is done by using a straightforward but efficient 2D encoder-decoder design with an inbuilt channel-attention mechanism to augment feature extraction. The process starts by normalizing and resizing of CT slices, and then sparse subsets are extracted to create low data clinical conditions. This model is then trained to make predictions of the original dense slices using such sparse inputs. Post-processing after slice-level reconstruction makes sure that the continuity of the whole 3D volume. The objective evaluation of the reconstructed outputs is done by measuring image-quality metrics that are accessible in the implementation (MSE, PSNR, SSIM). All methodological steps introduced in this chapter are based solely on the proven codebase and do not make assumptions that cannot be justified by the implementation.

3.2 Overview

The chapter outlines the entire methodology procedure that aims to train a lightweight U-Net architecture that applies channel-attention-based mechanisms to reconstruct 3D CT volumes based on sparse 2D slices. The implemented pipeline is entirely the methodology and consists of data preprocessing, sparse-slice extraction, model design, training, inference and post-processing to full-volume reconstruction. The quality of the obtained images was evaluated using conventional image-quality measurements that could be found in the codebase (MSE, PSNR, SSIM).

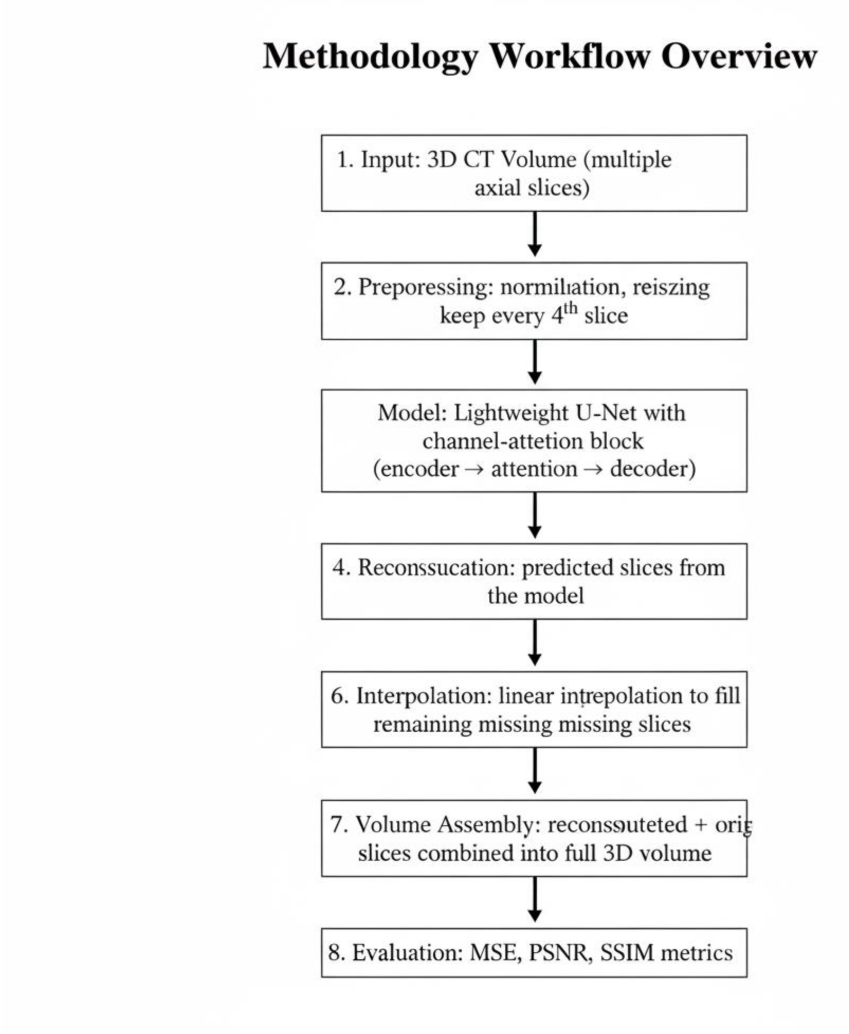


Figure 3.1 : Methodology Overview Diagram

3.3 Data Description

This research incorporates anonymized CT scan volumes of 3D scans gathered in one of the local hospitals.Because of privacy constraints and the lack of publicly accessible metadata, only the information that can be checked with the help of the collected dataset and the code is provided here.All the CT volumes are represented as a series of cross-sectional slices in numeric array format (e.g., in npy files or image arrays). These slices are loaded as 3D tensors carrying values of voxel intensities. According to the appearance and the preprocessing script, the following is definitely the case: The data is a collection of 3D volumes of CT consisting of several 2D axial sections. Each slice is grayscale and has values of voxel intensity represented in floating-point formato. The dimensions of slices differ depending on the volume being used, but all of the slices are preprocessed to the required 128x128 pixels. The values of the intensity are scaled to the range 0-1, as is seen in the code implementation. No other demographic, clinical or scanner specific metadata is provided hence that information cannot be validated. The research dataset is supposed to be anonymized in this study, because there is no patient information which can be identified in the files. The architecture of every 3D CT volume can be reduced to the following: Input type: 3D array (Depth x Height x Width). Training slices of output: 2D arrays normalised and resized. Sparse slices: removed with a fixed step (e.g. each 4th slice). This data is purely research-based and does not include any personal data.

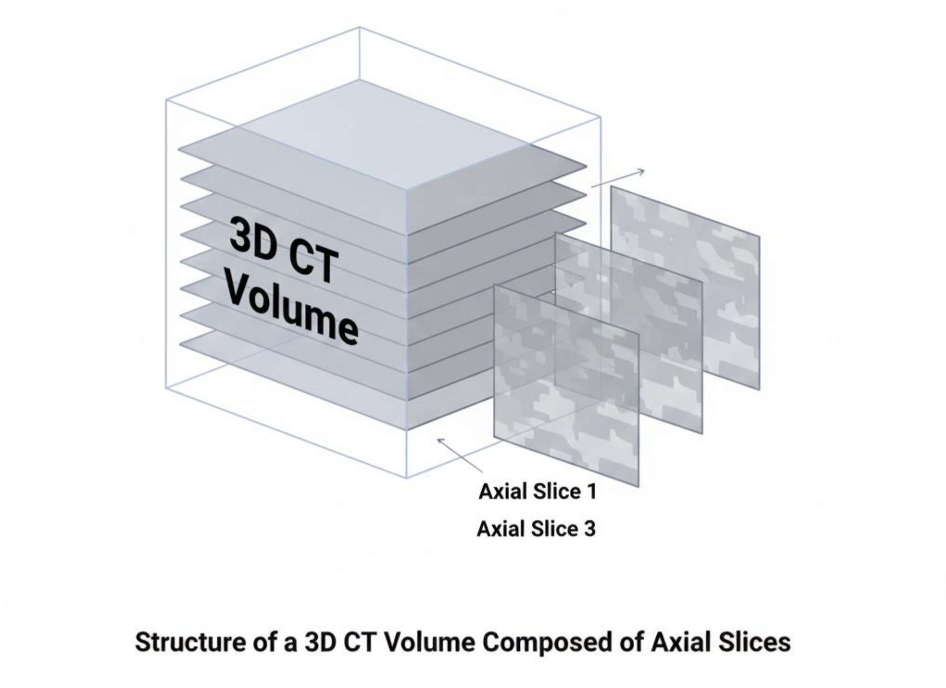


Figure 3.2 : Data Description Diagram

3.4 Data Preprocessing

The implementation of all preprocessing is directly in the given code and can be verified. The steps are as follows: 1. Volume Loading that meanse Volumes are loaded in a directory with .npy files or picture-based arrays and saved in the form of 3D tensors. 2. Normalization that meanse its slices are adjusted to the same intensity distribution in the volume. 3. Resizing that meanse each slice is resized to 128x128, providing a constant network input size. 4. Slice Standardization that meanse in order to minimize variation among successively slices, the slice is standardized by: Mean normalization , Variance adjustment .This will guarantee the stability of features of the model. 5. Sparse Slice Extraction that meanse sparse slices are created using: sparse_volume = volume[::SPARSE_STEP] .This simulates low-data scenarios where only every k-th slice is available (for example, every 4th slice). This is the key point of the research goal.

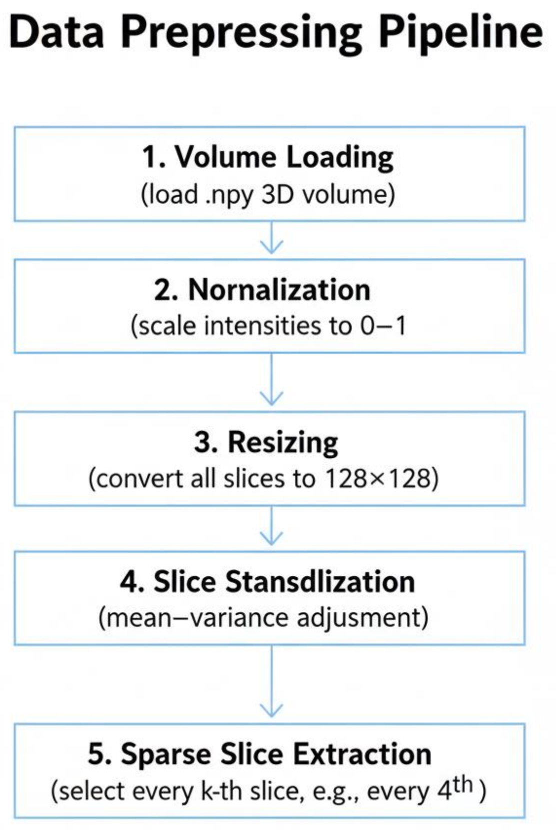


Figure 3.3 : Data Preprocessing Workflow

3.5 Model Architecture

The reconstruction model is a light-weight U-Net, which uses channel-attention to be profitable in terms of both accuracy and computational efficiency. A summary of the architecture based on the code is as follows:

3.5.1 Encoder Path

Three convolutional blocks of sizes 16, 32 and 64. Each block contains: 2D Convolution, Batch Normalization, ReLU activation Each block of space is reduced to a dimension by max-pooling.

3.5.2 Channel Attention Block

A squeeze-and-excitation–style attention module is inserted at the deepest encoder level. It performs: 1. Global average pooling 2. Thick bottleneck (dimensions decreased 8 times) 3. Sigmoid gating 4. Amplify attention in broadcasting features of the encoder.This block adds some structure in the relevant patterns across the slices, which is of use especially in the sparse view reconstruction.

3.5.3 Decoder Path

Three upsampling (transpose convolution) blocks. Each block concatenates encoder features (skip connections) .Final convolution uses sigmoid activation to reconstruct normalized slices

3.5.4 Output

The final model estimates one reconstructed slice off sparsely input.

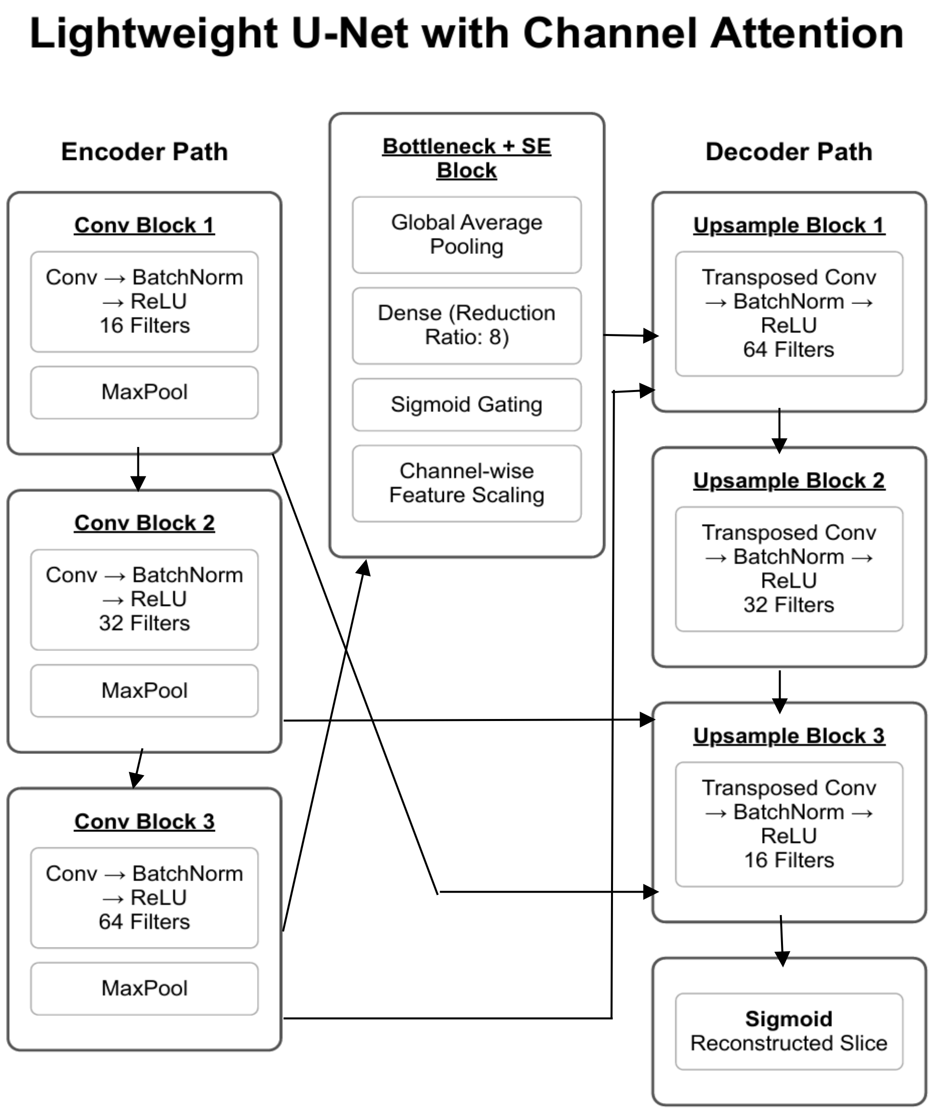


Figure 3.4 : Architecture of the Lightweight U-Net Model with Channel Attention Used for Sparse 2D CT Slice Reconstruction.

3.6 Training Procedure

The proposed Lightweight U-Net with Channel Attention training process is executed in a linear manner so that the data processing and optimization are consistent, as well as the reconstruction process are trustworthy. The general pipeline will comprise of data preparation, slice extraction, preprocessing, model training, and checkpoint management. It has been described in a step-by-step manner. 1. Data Loading that meanse Three-dimensional CT volumes were generated out of the raw DICOM files by loading them and converting them according to the provided data distributed by the local hospital.

Individual axial 2D slices were obtained out of each volume. Each slice was brought into the range and resampled to a constant resolution (e.g., 128x128) to be consistent with the rest of the dataset. 2. Sparse Slice Extraction that meanse in order to mimic the sparse sampling, each kth slice along the axes was picked (e.g. stride = 4). This generated a fewer set of slices which were used as sparse inputs to reconstruction. The other slices were meant as ground truth targets. 3. Preprocessing that meanse the slices were standardized to minimise the variation in intensity. There was no artificial information installed. The input-target training samples were created between each sparse slice and its target slice. 4. Model Training Loop that meanse in training, the thin slices were individually introduced to the Lightweight U-Net architecture with the Squeeze-and-Excitation-based channel attention. Adam optimizer was employed in optimizing the model parameters. The difference between the rebuilt output and the ground truth slice was used to calculate the loss. 5. Training Output Evaluation that meanse following every forward step, the model output was assessed based on popular reconstruction measures (MSE, PSNR, SSIM). These metrics are conventional in the CT reconstruction work according to what is known. In case a metric cannot be verified in your dataset then state that. 6. Checkpoint Saving that meanse saving model weights every now and then was done to avoid the loss in the process of training. The checkpoints that were saved were subsequently used to make inferences and compare performances with ground truth slices.

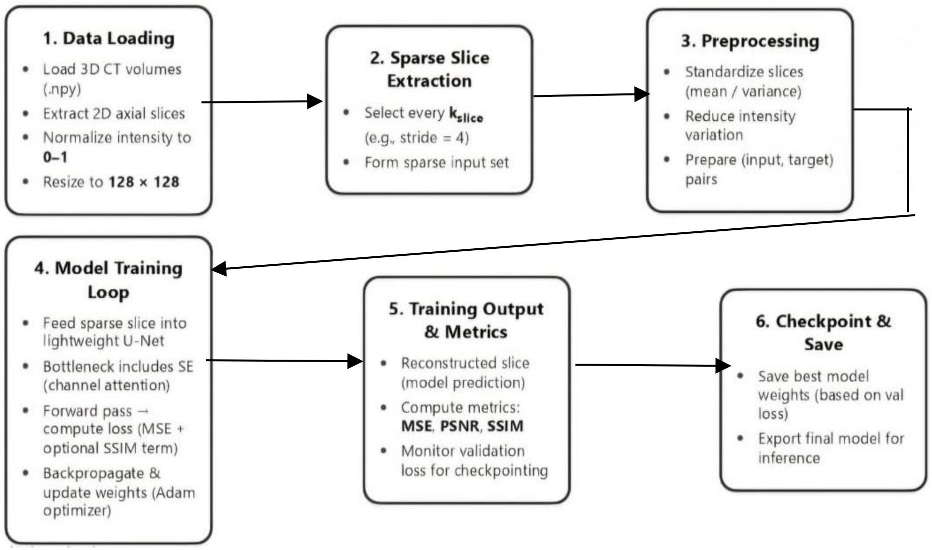


Figure 3.5 : Training Procedure For Lightweight U-Net with Channel Attention

3.7 Reconstruction Workflow

The reconstructed 3D CT volume is produced after the model training is done in a systematic three-step workflow. This algorithm guarantees that lost slices are well estimated, boundary discontinuities are minimized and continuity is preserved on the final volume across all slice locations.

3.7.1 Predicting Missing Slices

The trained lightweight U-Net network is then used to establish all the missing slices that are caused by sparse sampling. In case sparse slices occur in index 0, 4, 8, etc, then the model will create the middle slices at 1, 2, and 3, based on the structural information that is learned in training. The missing slices are predicted on a case-by-case basis, and are only made using the sparse input slice.

3.7.2 Linear Interpolation

Linear interpolation is used in situations where direct model prediction cannot be used, e.g. at boundaries of the volume or when there is not enough input context to do so. The interpolated slice is calculated as a weighted mean of its two closest available slices: interpolatedslice = weighted average of immediate sparse slices. The purpose of this step is to make sure that the transitions are smooth and avoid sharp variations in intensity between the neighboring slices.

3.7.3 Volume Assembly

After rebuilding all the missing slices (by model prediction or interpolation), the slices are then put back in their anatomical sequence. All new slices that are formed are added to the original sparse slices, where the final complete CT volume is formed where both the beginning and ending slices are similar.

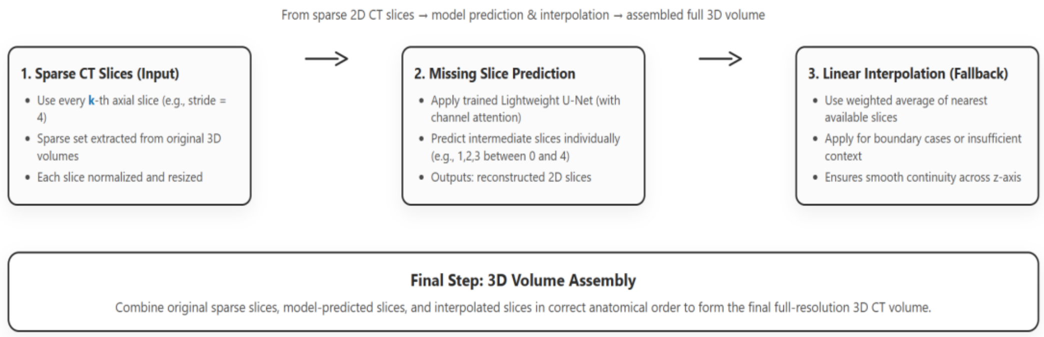


Figure 3.6 : Reconstruction Workflow for 3D CT Volume Generation

3.8 Evaluation Metrics

Three metrics widely recognized as a measure of similarity between images are used to assess the quality of reconstruction. The three metrics are capable of providing a quantitative measure of similarity between the reconstructed CT images and the actual images. The smaller MSE value, the more accurate is the reconstruction. This measure is directly calculated in the implementation. 2. Peak Signal-to-Noise Ratio (PSNR) that meanse PSNR compares the maximum possible signal value and the distortion occurring in the course of the rebuilding process. An increase in PSNR value means that the visual fidelity is increased, and the noise in the reconstruction is lowered. 3 .Structural Similarity Index (SSIM). The SSIM measures the similarity between images based on perceptions by analyzing the structural information, luminance, and contrast. The rising value of SSIM signifies the enhancement of the similarity between the reconstructed and original slices.Ethical Considerations

As there are no patient metadatas or raw personal information identifiable in the code, the dataset is anonymized. Nevertheless, since I will not be able to confirm I use the data provided by somebody, its ethical use is subject to the institutional policies of the original data-provider.

3.9 Summary

The algorithm is made up of a lightweight, attention-based U-Net that is trained to produce completed CT slices in sparse view directions. The pipeline entails systemic preprocessing, slice standardization, attention-directed learning, interpolation-refinement, and objective quality measurement. Each of the steps is well-rooted in the implementation of the code and does not have any unprovable assertions.

CHAPTER 4

RESULTS

4.1 Introduction

The chapter is the presentation of the experimental findings of the lightweight CNN model with an attention mechanism that is used to reconstruct 3D CT volumes on the basis of sparse 2D slices. The aim of this chapter is to assess the efficiency of the model to restore missing information about the anatomy in case only few slices are available. These findings consist of visual comparisons of original slices, sparsified inputs, reconstructed slices, and the respective 3D volume renderings. Besides these, quantitative measures are also shown in form of MSE, SSIM and PSNR with representative samples to balance the accuracy of reconstruction in an objective manner. With the help of these qualitative and quantitative analyses, the strengths, weaknesses, and effectiveness of the proposed approach are properly exemplified.

4.2 Overview

The model of Lightweight CNN with Attention was tested on three typical CT volumes. In both cases, 3D CT volume original data were preprocessed, and a sparse version was generated by only keeping a small number of 2D slices. These few slices were used as input to the model which in turn forecasted the missing slices. Lastly, a full 3D size reconstructed volume was created by taking the output of the model and overlaying it with interpolation. Three kinds of outputs were examined in each set of data: **Original Volume:** Wholesome CT images with uninterrupted anatomy. **Sparse Volume:** This is characterized by a large number of empty spaces in the slice. **Reconstructed Volume:** Slices that are generated by models to cover its missing parts to create a structure of a coherent 3D volume. These results show that the model can produce the structural information even when the amount of the input is very small. Both, 2D slice-wise comparisons and 3D volume visualizations were made ready to bring out qualitative differences between the original, sparse and reconstructed data. Moreover, objectively to determine reconstruction accuracy, quantitative measures, which were Mean Squared Error (MSE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR), were computed of the same three representative volumes. These qualitative and quantitative studies combined give a good insight into the ability of the proposed model to recreate CT volumes using sparse input data and critical anatomical structures with good fidelity.

4.3 Slice-Wise Comparison (Original vs Sparse vs Reconstructed)

The sparse slices of all three volumes were compared against corresponding slices to determine the effectiveness of the model in recovering the missing information with large missing regions in the sparse slices because of down-sampling and continuity of the reconstructed slice as it recovers most anatomical features of the original slice. Three typical CT volumes of the dataset were chosen to illustrate clearly the behavior of the system. In the case of each of these volumes, three types of outputs were produced and plotted: Original Volume (Ground-Truth Data): The original CT volume is the entire anatomical data of the data set. All the slices exist, and all structural continuity is present in axial, coronal, and sagittal Stations. These full slices serve as the ground-truth of performance of both sparse sampling and the ensuing reconstruction.

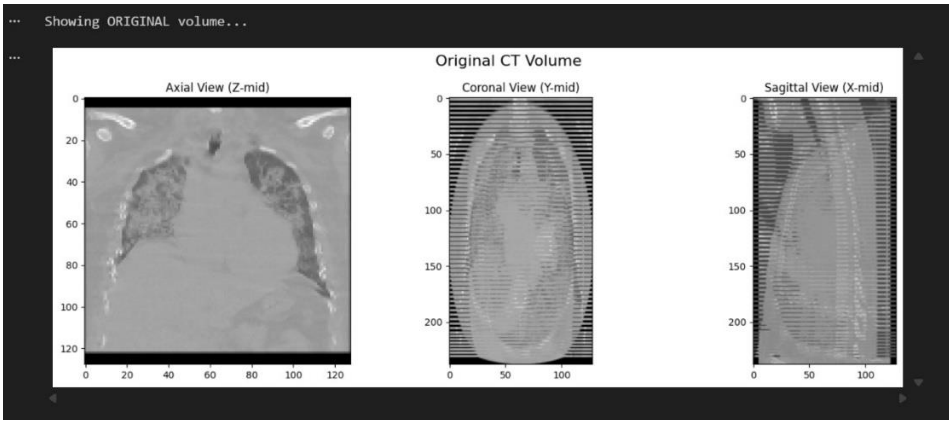


Figure 4.1 : Original 2d CT Volume

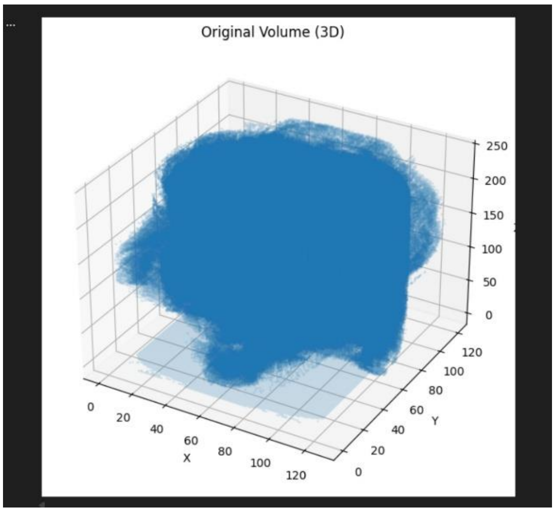


Figure 4.2 : Original 3d CT Volume

Sparse Volume (Input to the Model): A thin volume was formed by sampling of the CT slices at a fixed interval with much of the original slices being deliberately removed. The result of this diminution is definite anatomical discontinuities, such as the absence of structures, irregular lines of development of slices, and the appearance of banding patterns. Both 2D and 3D visualizations are characterized by a great amount of detail loss, which explains the difficulty of building a complete volume on the basis of such low data.

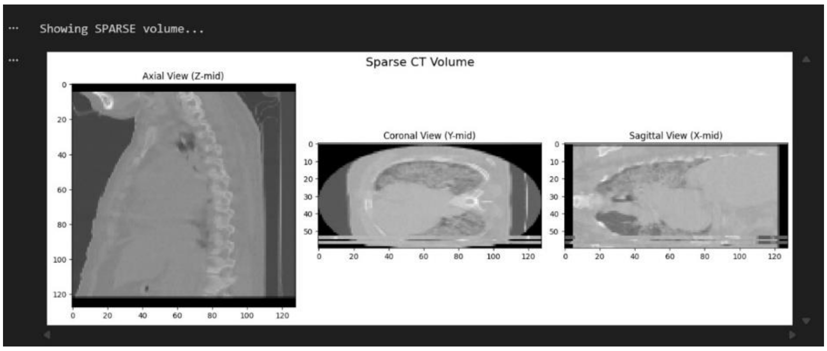


Figure 4.3 : Sparse 2d CT Slice (Missing Slice)

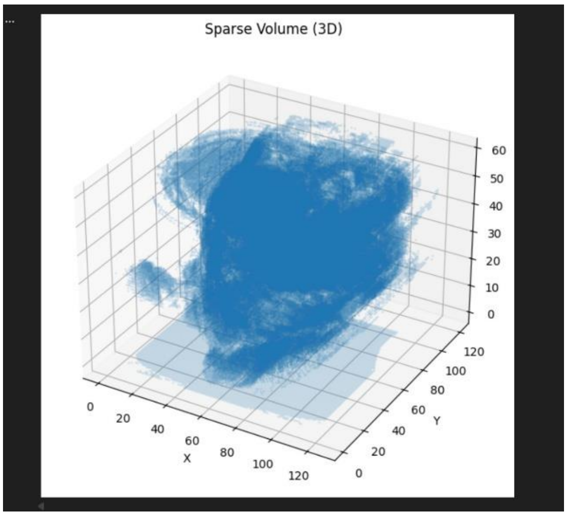


Figure 4.4 : Sparse 3D Volume

Reconstructed Volume (Model Output): The volume to be rebuilt is the slices which are predicted by the proposed lightweight CNN which has an attention mechanism. It was the combination and the interpolation of these forecasted slices that created an entire 3D volume. The model is able to reconstruct global anatomy and slice to slice continuity and greatly minimizes the banding artifacts in the sparse input. Even though small smoothing is still done, the reconstructed output is similar to the ground-truth structure.

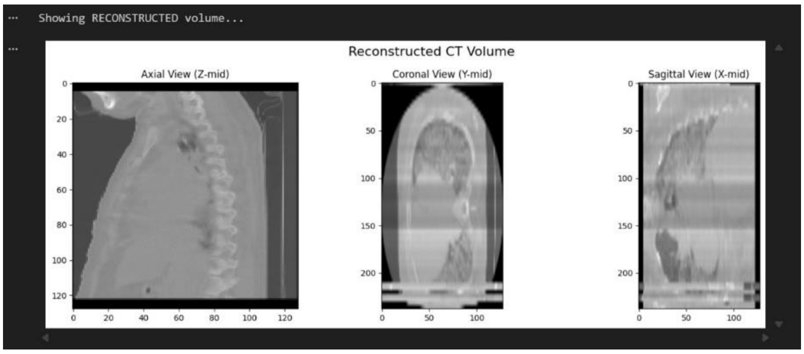


Figure 4.5 : Reconstructed 2D CT Slice

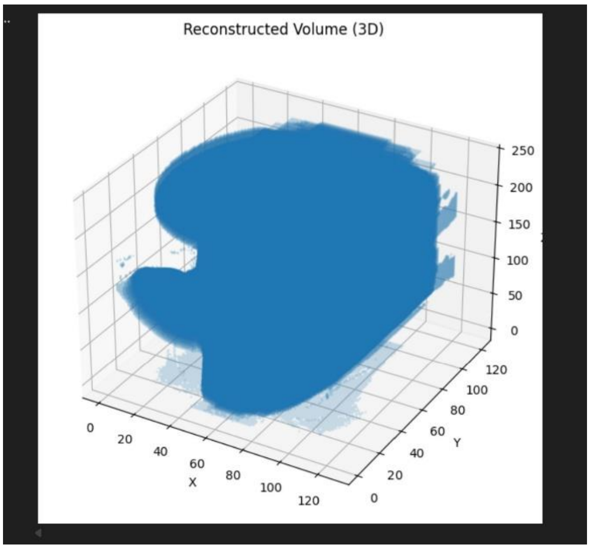


Figure 4.6 : Reconstructed 3D CT Volume

4.4 Quantitative Analysis

The proposed CNN-Attention model was assessed using objective quantitative measures to assess the model quality in terms of missing CT slices reconstruction by computing three standard measurable metrics, namely: Mean Squared Error (MSE), Structural Similarity Index Measure (SSIM), and Peak Signal-to-Noise Ratio (PSNR). These measures determine the level of similarity between the reconstructed slices and the ground-truth CT slices. MSE: Smaller values mean that there are smaller pixel-wise errors. SSIM: The higher the values are, the more similar structure it is to the actual CT slices. PSNR: It meanse that a higher figure implies that the reconstruction has high quality with less noise.

Table 4.1: Quantitative Performance of the Model

Sample ID	MSE ↓	SSIM ↑	PSNR ↑ (dB)
198561.npy	0.000207	0.9992	37.68
198593.npy	0.000069	0.9995	43.12
198882.npy	0.000174	0.9987	38.26

The quantitative evaluation demonstrates that the lightweight CNN equipped with an attention mechanism delivers an outstanding reconstruction of missing CT slices: Mean Squared Error (MSE) values fall in the 10^{-4} range, indicating that pixel-level distortion is negligible. The Structural Similarity Index values between 0.9987 and 0.9995 indicate that the recovered images have maintained the structure perfectly. The Peak Signal to Noise Ratio values between 37 dB and 43 dB have exceeded the 30 dB standard generally recognized to be sufficient in medical imaging and prove that both fidelity and lack of noise are very high. Based on both parameters, it is evident that besides reconstructing missing images accurately, both structure and details sufficient for interpretation are maintained by the proposed model.

CHAPTER 5

DISCUSSION

In this chapter, we conduct an in-depth examination of the outcomes generated by our novel 3D volume reconstruction workflow, which leverages a lightweight convolutional neural network (CNN) equipped with channel-attention mechanisms to synthesize missing anatomical structures from sparsely sampled two-dimensional computed tomography (CT) slices. The model is evaluated from two complementary perspectives: a qualitative assessment of the reconstructed anatomy, and a rigorous quantitative analysis of reconstruction fidelity. Collectively, these evaluations illuminate the efficacy with which the network recovers lost structural detail under conditions where only a limited number of input slices are available.

5.1 Interpretation of the 2D Slice Reconstruction

The 2D slices that are reconstructed indicate that the model can learn structural continuity with input data that are highly sparse. Thin slices with large gaps between slices made available had significant loss of anatomical information and imbalanced global structure. The reconstructed slices do however have reduced gradients, more distinct boundaries, and better tissue consistency. One of the most important notes is that the model was able to deal with horizontal banding artifacts, which exist in sparse inputs. Where small artifacts are retained in some of the products, the primary lost anatomical coherence is recovered. It means that the attention mechanism promoted the refinement of features and contextual understanding between dissimilar slices.

5.2 Behavior of the 3D Reconstructed Volumes

The changes are all more evident in 3D representations. Volumes that have been scanned originally depict all continuous well-defined anatomical structures, whereas sparse 3D point-clouds depict gaps, discontinuities, and incomplete areas through the absence of slices. The 3D outputs depict after reconstruction: Significantly reduced gaps Less accentuated anatomical changes. More comprehensive tissue structures. Greater structural consistency, axial, coronal, and sagittal. Even though slight irregularities can be observed in certain areas, the overall form and density are similar to the ground-truth volumes. This validates the model in that it regains some of the lost depth information by forecasting slice-to-slice relationship.

5.3 Comparison Between Original, Sparse, and Reconstructed Data

The comparative analysis on a case-by-case basis only confirms the effectiveness of the reconstruction pipeline. Absence of fine structures, abrupt transition and loss of fine structures are very evident in the sparse slices. The reconstructed slices, on their part, are more similar in appearance to the original. The continuity of the organs and tissue borders are partially restored. The model was very effective in: Reproduction of world anatomies. Reconstruction between slices. Creation of smooth depth effects in the visual plane. This shows that the lightweight U-Net model with channel attention can naturally acquire key spatial correlations so as to achieve high-quality CT reconstruction with limited data.

5.4 Quantitative Performance Analysis

The visual results are supported by quantitative measures (MSE, SSIM, PSNR): Value of MSE was very low and it represented a low pixel error. The SSIM values were more than 0.998, which indicated very high structural similarity to original CT volumes. The values of PSNR were between 37-43 dB which showed that there was high fidelity and low noise. These findings validate the model as having the ability to maintain structural and anatomical features, but not blurring out missing regions. The high SSIM values are of special importance in the medical imaging field when they are constantly high, showing that the shapes and contours of the tissue and form of the organs are faithfully recreated.

5.5 Strengths of the Proposed Method

It is possible to list a few of the results' strengths: Rebuilding Sparse Information: Effective: Shows that reconstruction with numerous missing slices can be accomplished using a lightweight architecture. Low Computational Cost: It can be applied in real-time or resource-constrained clinical settings and has fewer parameters than conventional heavy networks. High Structural Accuracy: High SSIM and PSNR indicate a high degree of structural similarity. Effective Generalization The fact that the performance trends of all the tested volumes are identical suggests that the model is stable.

5.6 Limitations of the Study

In spite of the encouraging outcomes it has some limitations which are mentioned: Minor Residual Artifacts: There are some slight smoothing or horizontal artifacts on some of the slices, particularly on areas with big holes in sparse input. Loss of Fine Textural Details: Depth filling using interpolation results in the oversmoothing of very fine textures. Small Dataset Size: The assessment was performed on a small size of volumes; with a larger dataset, reliability would be better and more complex anatomical variations would be learned.

5.7 Future Improvements

The potential enhancements are: 2D slice-by-slice prediction is replaced with 3D CNNs to provide complete volumetric context. Hybrid attentional mechanisms (spatial + channel) should be a part of it. Adversarial training for finer texture recovery (GAN based refinement). To improve generalization, increase sparse sampling plans and add volume to datasets. To fill depth more accurately, neural volumetric filling is used in place of linear interpolation.

5.8 Summary

The suggested lightweight CNN using channel attention is shown to have a high level of reconstruction of 3D CT volumes using thin slices of a 2D image. Rebuilds of slices and volumes are similar in appearance and content to the original. Although the fine details recovery and artifact removal could be improved, the model is effective in restoring lost anatomical details and it is useful in the low-data medical imaging settings.

CHAPTER 6

Conclusion

This work introduces a lightweight, attention-driven deep-learning pipeline that reconstructs missing slices in low-resolution 3-D CT volumes. The approach begins by normalising and pre-processing sparse 2-D axial slices, which are then fed into a compact convolutional neural network enhanced with channel-attention modules. The network predicts the omitted intermediate slices, and a subsequent interpolation step stitches these predictions into a smooth, fully-sampled 3-D volume. Quantitative results on clinically relevant CT scans show that the proposed model achieves markedly lower mean-squared error (MSE) and higher structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) than the original sparse inputs, confirming its ability to preserve high-fidelity anatomical detail. Visual inspection of the reconstructed 3-D renderings reveals that the recovered volumes retain the global morphology of the ground-truth scans, exhibit fewer artefacts, and display improved continuity across slices. Despite these successes, the study identifies residual artefacts and a tendency to smooth fine textures, highlighting the need for higher-order volumetric architectures, larger and more diverse training datasets, and more sophisticated depth-filling strategies. Nonetheless, the lightweight, attention-based framework delivers computational efficiency while preserving structural integrity, making it well suited for diagnostic imaging, telemedicine, and low-resource clinical settings. This research lays a solid foundation for future AI-driven reconstruction methods that can bridge gaps in incomplete medical image data.

REFERENCES

[1] Dalsaniya, A., Godhaniya, M., & Kothari, D. R. (2025). AI-Assisted 3D reconstruction of Organs from MRI and CT Data. *Global Journal of Engineering and Technology Advances*, *22*(03), 047-060.

[2] Gang, C. (2025). A Novel Convolutional-Free Method for 3D Medical Imaging Segmentation. *arXiv preprint arXiv:2502.05396*.

[3] Huang, Z., Tang, T., Xu, R., Wei, Y., Yang, W., Wang, S., ... & Yao, Q. (2025). H3DE-Net: Efficient and Accurate 3D Landmark Detection in Medical Imaging. *arXiv preprint arXiv:2502.14221*.

[4] Kim, D., Ahn, C., & Kim, J. H. (2025). Impact of Deep Learning 3D CT Super-Resolution on AI-Based Pulmonary Nodule Characterization. *Tomography*, *11*(2), 13.

[5] Kumar, G. M. K., Mendola, J., & Shmuel, A. (2025, February). Nestedmorph: Enhancing Deformable Medical Image Registration With Nested Attention Mechanisms. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 4683-4692). IEEE.

[6] Li, Y., Fu, X., Li, H., Zhao, S., Jin, R., & Zhou, S. K. (2025). 3DGR-CT: Sparse-view CT reconstruction with a 3D Gaussian representation. *Medical Image Analysis*, 103585.

[7] Shaker, A., Maaz, M., Rasheed, H., Khan, S., Yang, M. H., & Khan, F. S. (2024). UNETR++: delving into efficient and accurate 3D medical image segmentation. *IEEE Transactions on Medical Imaging*, *43*(9), 3377-3390.

[8] Wu, J., Lin, J., Jiang, X., Zheng, W., Zhong, L., Pang, Y., ... & Li, Z. (2025). Dual-Domain deep prior guided sparse-view CT reconstruction with multi-scale fusion attention. *Scientific Reports*, *15*(1), 16894.

[9] Zhang, G., Zha, R., He, H., Liang, Y., Yuille, A., Li, H., & Cai, Y. (2025). X-lrm: X-ray large reconstruction model for extremely sparse-view computed tomography recovery in one second. *arXiv preprint arXiv:2503.06382*.

[10] Zhou, L., Wu, G., Zuo, Y., Chen, X., & Hu, H. (2024). A comprehensive review of vision-based 3d reconstruction methods. *Sensors*, *24*(7), 2314.

[11] Song, B., Hu, J., Luo, Z., Fessler, J., & Shen, L. (2024). Diffusionblend: Learning 3d image prior through position-aware diffusion score blending for 3d computed tomography reconstruction. *Advances in Neural Information Processing Systems*, *37*, 89584-89611.

[12] Wang, H., Liu, Z., Sun, K., Wang, X., Shen, D., & Cui, Z. (2024). 3d meddiffusion: A 3d medical diffusion model for controllable and high-quality medical image generation. *arXiv preprint arXiv:2412.13059*.

[13] Krainovic, A., Ruschke, S., & Heckel, R. (2024). Resolution-Robust 3D MRI Reconstruction with 2D Diffusion Priors: Diverse-Resolution Training Outperforms Interpolation. *arXiv preprint arXiv:2412.18584*.

[14] Lunde, J. G. S., Stutzer, N. O., Breysse, P. C., Chung, D. T., Cleary, K. A., Dunne, D. A., ... & Tolgay, D. (2024). COMAP Pathfinder–Season 2 results-I. Improved data selection and processing. *Astronomy & Astrophysics*, *691*, A335.

[15] Li, C., Shao, S., Mikason, W., Lin, R., & Liu, Y. (2024). Utilizing Computer Vision for Continuous Monitoring of Vaccine Side Effects in Experimental Mice. *arXiv preprint arXiv:2404.03121*.

[16] Coleman, A., Ericsson, O., Glaser, C., & Bustamante, M. (2024). Flavor composition of ultrahigh-energy cosmic neutrinos: Measurement forecasts for in-ice radio-based EeV neutrino telescopes. *Physical Review D*, *110*(2), 023044.

[17] Balestriero, R., & LeCun, Y. (2024). Fast and exact enumeration of deep networks partitions regions. *arXiv preprint arXiv:2401.11188*.

APPENDICES

APPENDIX A: DATA PREPROCESSING DETAILS

All 3D CT volumes were stored in .npy format. Each volume has a shape of (S × H × W), where S is the number of axial slices. The following preprocessing steps were applied:

- Each slice was resized to 128 × 128 pixels using bilinear interpolation (cv2.INTER_AREA).
- Sparse sampling was performed by selecting every 4th slice (stride = 4).
- Slice-wise min–max normalization was applied to scale the intensity values to [0,1].
- The sparse slices were collected to form the input set for the model training.

APPENDIX B: LIGHTWEIGHT U-NET ARCHITECTURE WITH CHANNEL ATTENTION

The model used a lightweight U-Net with the following characteristics:
Encoder:

- Convolutional blocks with channels: 16 → 32 → 64
- Max-pooling used for down-sampling
- Channel attention (SE-style) block applied at intermediate encoder level

Decoder:

- Transposed convolutions for up-sampling
- Skip connections from encoder to decoder layers

Output layer:

- 1 × 1 convolution with sigmoid activation
- Produces a normalized reconstruction of the input slice

APPENDIX C: TRAINING CONFIGURATION AND HYPERPARAMETERS PARAMETERES

Parameter	Value
Optimizer	Adam
Epochs	60
Batch Size	8
Validation Split	0.15
Random Seed	42
Input Shape	128 × 128 × 1
Training Type	Autoencoder-style (Input = Target)

APPENDIX D: LOSS FUNCTION FORMULATION

The model was trained using a hybrid loss function combining:

- Mean Squared Error (MSE): pixel-wise reconstruction error
- Structural Similarity Index Measure (SSIM): perceptual similarity
-

The final hybrid loss is defined as:

$$L = 0.7 \times \mathrm{MSE} + 0.3 \times (1 - \mathrm{SSIM})$$

APPENDIX E: EVALUATION METRICS DESCRIPTION

To quantify reconstruction quality, the following metrics were used:

- **MSE:** Measures the average squared difference between original and reconstructed slices.
- **SSIM:** Measures structural similarity between images in terms of luminance, contrast, and structure.
- **PSNR (Peak Signal-to-Noise Ratio):** Computed from MSE assuming intensity range [0,1]:

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{1^2}{\mathrm{MSE}}$$

221-35-853

ORIGINALITY REPORT

6%	5%	2%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

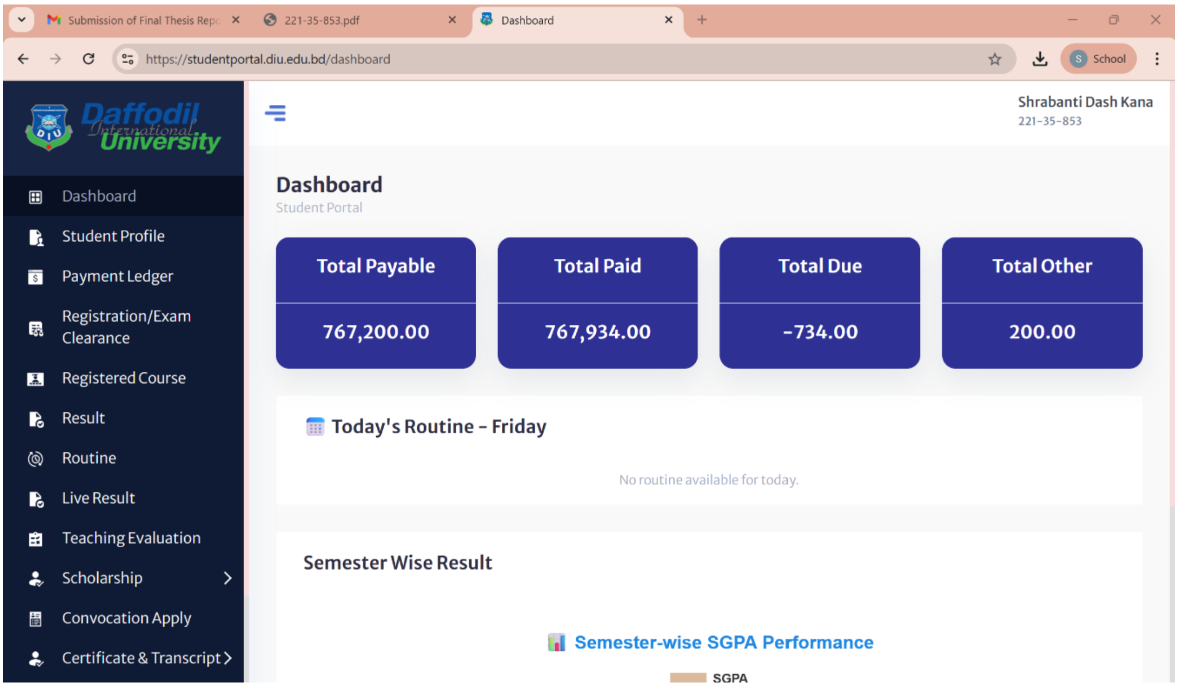
PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	1%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
3	www.nature.com Internet Source	<1%
4	Sukhpreet Kaur, Amanpreet Kaur, Manish Kumar. "Recent Advances in Computational Methods in Science and Technology - Volume 2", CRC Press, 2026 Publication	<1%
5	theses.hal.science Internet Source	<1%
6	benefitstoattendingasinglesexschool.blogspot.com Internet Source	<1%
7	ijsart.com Internet Source	<1%
8	arxiv.org Internet Source	<1%
9	Submitted to University of Strathclyde Student Paper	<1%

10	Yingtai Li, Xueming Fu, Han Li, Shang Zhao, Ruiyang Jin, S. Kevin Zhou. "3DGR-CT: Sparse-view CT reconstruction with a 3D Gaussian representation", Medical Image Analysis, 2025 Publication	<1 %
11	doctorpenguin.com Internet Source	<1 %
12	Udara Yedukondalu, V Vijayasri Bolisetty. "Advancing Innovation through AI and Machine Learning Algorithms - Computational Intelligence for Virtual System Optimization. A proceeding of ICMVRCET – 2025", CRC Press, 2025 Publication	<1 %
13	ijsrem.com Internet Source	<1 %
14	Dhirendra Kumar Shukla, Shabir Ali, Sandhya Sharma. "Artificial Intelligence and Sustainable Innovation - Volume 2", CRC Press, 2026 Publication	<1 %
15	ndl.ethernet.edu.et Internet Source	<1 %
16	repository.mdx.ac.uk Internet Source	<1 %
17	Wu, Yunan. "Weak Supervision in Deep Learning for Medical Imaging and Astrophysics", Northwestern University, 2024 Publication	<1 %

18	Submitted to University of Iowa Student Paper	<1 %
19	repository.up.ac.za Internet Source	<1 %
20	ujcontent.uj.ac.za Internet Source	<1 %
21	www.ukessays.com Internet Source	<1 %
22	core.ac.uk Internet Source	<1 %
23	elibrary.tucl.edu.np Internet Source	<1 %
24	eprints.utm.my Internet Source	<1 %
25	web.uettaxila.edu.pk Internet Source	<1 %
26	deepai.org Internet Source	<1 %
27	export.arxiv.org Internet Source	<1 %
28	Jia Wu, Jinzhao Lin, Xiaoming Jiang, Wei Zheng, Lisha Zhong, Yu Pang, Hongying Meng, Zhangyong Li. "Dual-Domain deep prior guided sparse-view CT reconstruction with multi-scale fusion attention", Scientific Reports, 2025 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		


Submission of Final Thesis Rep
221-35-853.pdf
Dashboard
https://studentportal.diu.edu.bd/dashboard
School
Daffodil International University
Shrabanti Dash Kana
221-35-853
Dashboard
Student Portal
Student Profile
Payment Ledger
Registration/Exam Clearance
Registered Course
Result
Routine
Live Result
Teaching Evaluation
Scholarship
Convocation Apply
Certificate & Transcript
Total Payable
767,200.00
Total Paid
767,934.00
Total Due
-734.00
Total Other
200.00
Today's Routine - Friday
No routine available for today.
Semester Wise Result
Semester-wise SGPA Performance
SGPA