

# **Enhanced U-Net Architecture with VGG16 Backbone and Attention Mechanisms for Automated Nasal Sinus Disease Segmentation**

**By  
SM Shaqib  
213-15-4614**

## **FINAL YEAR DESIGN PROJECT REPORT**

**This Report Presented in Partial Fulfillment of the  
Requirements for the Degree of Bachelor of Science in  
Computer Science and Engineering**

### **Supervised by**

**Sharun Akter Khushbu  
Assistant Professor  
Department of Computer Science and Engineering  
Daffodil International University**

### **Co-Supervised by**

**Md. Sadekur Rahman  
Assistant Professor  
Department of Computer Science and  
Engineering Daffodil International  
University**



**DAFFODIL INTERNATIONAL  
UNIVERSITY  
Dhaka, Bangladesh**

**September 16, 2025**

## APPROVAL

This Project titled “Enhanced U-Net Architecture with VGG16 Backbone and Attention Mechanisms for Automated Nasal Sinus Disease Segmentation”, submitted by SM Shaqib, ID No: 213-15-4614 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 16 September, 2025.

### BOARD OF EXAMINERS



**Dr. Sheak Rashed Haider Noori**

**Professor and Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Chairman**



**Md. Sadekur Rahman**

**Assistant Professor**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Md. Saiful Islam**

**Assistant Professor**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Dr. Md. Arshad Ali**

**Professor**

Department of Computer Science and Engineering  
Hajee Mohammad Danesh Science & Technology University

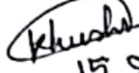
**External Examiner**

# DECLARATION

---

We hereby declare that this project has been done by us under the supervision of **Sharun Akter Khushbu, Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

**Supervised by:**

  
15.9.25

---

**Sharun Akter Khushbu**

Assistant Professor

Department of CSE

Daffodil International University

**Co-Supervised by:**

  
15.9.25

---


**Md. Sadekur Rahman**

Assistant Professor

Department of CSE

Daffodil International University

**Submitted by:**

  
15/9/25

---

**SM Shaqib**

ID: 213-15-4614

Department of CSE

Daffodil International University

# ACKNOWLEDGEMENTS

---

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Sharun Akter Khushbu, Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of “**Deep Learning**” to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

# ABSTRACT

---

This paper proposes an automated nasal sinus disease segmentation system using an enhanced U-Net model with VGG16 backbone and attention mechanism. It would overcome the weaknesses inherent in manual segmentation by presenting a superior and accurate solution. The data set for this research consists of medical image data, public and self-collected, preprocessed for segmentation purposes. Preprocessing pipeline included data annotation using Roboflow, resizing, normalization, and binary mask creation. Random rotation, flipping, and shifting were some of the data augmentation methods used to make the model more robust. The proposed model architecture includes a VGG16-based encoder with a bottleneck for enhancing features and a decoder with attention gates to achieve maximum attention in regions of interest in disease areas. Adam optimizer and binary cross-entropy loss were used for training the model and tested on Dice coefficient, Intersection over Union (IoU), and F1-score. The model underwent great improvement after 50 epochs with accuracy 98.43%, which indicates its proficiency in nasal sinus disease segmentation. Furthermore, the proposed model was compared with other existing segmentation models, and the results demonstrated its superior performance, justifying the effectiveness of the custom architecture.

# Table of Contents

<b>Approval</b>	<b>i</b>
<b>Declaration</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction.....	1
1.2 Motivation .....	4
1.3 Objectives .....	4
1.4 Methodology .....	5
1.5 Project Outcome.....	6
1.6 Organization of the Report .....	7
<b>2 Background</b>	<b>8</b>
2.1 Introduction.....	8
2.2 Literature Review .....	8
2.3 Gap Analysis .....	14
2.4 Summary .....	17
<b>3 Research Methodology</b>	<b>18</b>
3.1 Methodology/Requirement Analysis & Design Specification.....	18
3.1.1 Overview .....	18
3.1.2 Proposed Methodology/ System Design .....	18
3.1.3 Functional and Nonfunctional Requirements.....	19
3.1.4 Data Flow Diagram .....	20
3.1.5 UI Design .....	21
3.2 Detailed Methodology and Design .....	22
3.3 Project Plan .....	24
3.4 Task Allocation.....	27
3.5 Summary .....	27
<b>4 Implementation and Results</b>	<b>28</b>

4.1	Environment Setup .....	28
4.2	Training Process.....	29
4.3	Testing and Evaluation .....	35
4.4	Results and Discussion .....	35
4.4.1	Training Accuracy and Loss Curve.....	35
4.4.2	Prediction of diseased region with ground truth.....	39
4.4.3	visualizations of feature maps .....	40
4.4.4	ROC Curve.....	42
4.4.5	Integrated Gradients (IG) .....	43
4.4.6	Grad-CAM Interpretation .....	45
4.4.7	Error Map with original Image.....	46
4.4.8	Pixel wise Confusion Matrix .....	47
4.5	Summary .....	48
<b>5</b>	<b>Engineering Standards and Design Challenges</b>	<b>49</b>
5.1	Compliance with the Standards.....	49
5.1.1	Software Standards.....	49
5.1.2	Hardware Standards .....	49
5.1.3	Communication Standards.....	50
5.2	Impact on Society, Environment and Sustainability .....	50
5.2.1	Impact on Life.....	50
5.2.2	Impact on Society & Environment.....	51
5.2.3	Ethical Aspects .....	51
5.2.4	Sustainability Plan.....	51
5.3	Project Management and Financial Analysis.....	51
5.4	Complex Engineering Problem.....	52
5.4.1	Complex Problem Solving.....	52
5.4.2	Engineering Activities.....	53
5.5	Summary .....	55
<b>6</b>	<b>Conclusion</b>	<b>56</b>
6.1	Summary .....	56
6.2	Limitation .....	57
6.3	Future Work .....	57
	<b>References</b>	<b>59</b>

# List of Figures

Figure 1.1: Example of diseased maxillary sinus and healthy .....	3
Figure 2.1: Different Sinus Regions.....	8
Figure 3.1: Proposed Methodology Diagram .....	19
Figure 3.2: Data Flow Diagram .....	21
Figure 3.3: Design of the web based implementation platform.....	21
Figure 4.1: Prediction Results for Attention Unet-VGG.....	29
Figure 4.2 Prediction Results for Standard U-Net .....	30
Figure 4.3: Prediction Results for DeepLabV3+.....	31
Figure 4.4: Prediction Results for SegNet .....	32
Figure 4.5: Prediction Results for LinkedNet .....	33
Figure 4.6: Prediction Results for PSPNet.....	34
Figure 4.7: Training Curve for Attention UNet-VGG.....	36
Figure 4.8: Training Curve for Standard U-Net with skip connection .....	36
Figure 4.9: Training Curve for DeepLabV3+ .....	37
Figure 4.10: Training Curve for SegNet.....	37
Figure 4.11: Training Curve for LinkedNet .....	38
Figure 4.12: Training Curve for PSPNet.....	38
Figure 4.13: Prediction with respect of gerund truth .....	39
Figure 4.14: Visualization of block3_conv1 layer .....	40

Figure 4.15: Visualization of block3_conv2 Layer.....	41
Figure 4.16: Visualization of block3_conv3 Layer.....	42
Figure 4.17: ROC Curve.....	43
Figure 4.18: Integrated Gradients visualization.....	44
Figure 4.19: Grad-Cam visualization .....	45
Figure 4.20: Error Map Visualization .....	46
Figure 4.21: Pixel wise Confusion Matrix .....	47

# List of Tables

Table 2.1: Summary of Literature Reviewed. ....	10
Table 2.2: Gap Analysis considering recent works. ....	15
Table 3.1: Model Summary .....	22
Table 3.2: Layers name by Groups .....	24
Table 3.3: Estimated Task Allocation for the project.....	27
Table 4.1: Training summary of Attention Unet-VGG .....	29
Table 4.2: Training summary of Standard U-Net.....	30
Table 4.3: Training summary of DeepLabV3+ .....	31
Table 4.4: Training summary of SegNet. ....	32
Table 4.5: Training summary of Linked Net.....	33
Table 4.6: Training summary of Linked Net.....	34
Table 5.1: Cost Estimation for the project.....	51
Table 5.2: Mapping with Complex Engineering Problem. ....	52
Table 5.3: Mapping with knowledge Profile. ....	53
Table 5.4: Mapping with Complex Engineering Activities. ....	54

# Chapter 1

## Introduction

This first chapter contains the introduction to the project, providing background on why it was necessary, how it was initiated, its particular objectives, and a general account of the methodology and hoped-for results and also the overall structure of the report.

### 1.1 Introduction

**Background:** Over the past decades, there have been improvements in computer technology that have allowed incredible progress to be achieved in solving difficult real-world problems in numerous fields. Among these fields is AI-based segmentation, in which precise and quick models have been most significant in disease detection and prediction. In spite of all these improvements, the field is still faced with setbacks, most notably in real-time processing and precision in segmentation.

Convolutional Neural Networks (CNNs), U-Net, VGG16, and other technologies have shown promise in overcoming these difficulties by combining deep learning techniques with fine-grained analysis. However, existing systems tend to suffer from constraints such as high computing cost and limited precision under real-world conditions.

Nasal diseases are indirectly related to the nasal shape. If nasal discomfort is consistent in a person, then deviances in the nasal tissue can be anticipated. Whereas allergic rhinitides cause inferior turbinate hypertrophy, chronic sinusitis causes loss of nasal sinus volume, particularly of the maxillary sinus [1]. Such pathologic alterations highlight the importance of accurate imaging and segmentation methods in diagnosis and management of nasal sinus disease. To determine the disease severity and plan effective treatment, careful scrutiny of the nasal structure is required.

Segmentation of medical image is the process of applying image processing methods to highlight a region or a structure in medical images. Different shapes or textures can be delineated and expressed by this process [2]. Segmentation enables salient structures, i.e., nasal septum and maxillary sinus, to be segmented when there is nasal sinus disease present. This makes it possible to detect pathology, i.e., inflammation or structural pathology, easily. This is necessary to improve the accuracy of diagnoses and the development of treatment programs. The limitations of traditional diagnostic methods are evident in the common practice of manually measuring inflammatory sinus volume. This method is objective, but it requires a lot of effort [3]. The problem with manual segmentation is that it is less reliable in practice. This is due to how long it takes and its vulnerability to differences in interpretation between observers. Similarly, it appears difficult to have objective standardized tests to assess the severity of nasal septal deviation (NSD) even with an excess of research on NSD classification and preoperative planning [4]. Because of these limitations, automatic segmentation methods must be applied in order to obtain reproducible and accurate measurements to grade nasal sinus disease. A patient-specific upper airway model is necessary for clinical, teaching, and

research purposes [5]. They are of most use in nasal sinus disease segmentation since they allow customized examination of sinus anatomy and pathology and thereby enhance clinical decision-making and research on disease mechanism. Moreover, the maxillary sinus region identification is important in dental specializations such as implant surgery, tooth extraction, and the detection of odontogenic diseases. Accurate maxillary sinus segmentation is crucial for both diagnosis and therapy [6]. Accurate border definition makes dental and surgical treatments safe and effective, and accurate segmentation ensures that doctors can do the same. The maxillary sinus (MS) is the largest and most early-developing of the four paranasal sinuses. The MS is a pyramid-shaped hollow that is bordered by six walls and shares boundaries with important anatomical structures due to its positional nature [7]. Appropriate segmentation is extremely crucial in preventing complications in therapies because of its anatomical complexity and proximity to structures like the orbit and cranial cavity. Moreover, the system of paranasal sinuses is the most complex anatomical structure. Accurate diagnosis and treatment require an understanding of spatial relationships [8]. Although this complexity is observed in veterinary medicine, it is akin to the complexity involved in segmenting the nasal sinuses in the human body, where accurate spatial mapping is required for an appropriate diagnosis. Presently, developments in deep learning have proven promising in the automatic segmentation of medical images as a solution to these types of problems.

Sinusitis means your sinuses are inflamed. The sinuses are air-filled spaces behind the forehead, nose, cheeks and eyes. But when these cavities are blocked or filled with fluid, they offer the perfect environment for germs to grow, causing an infection. Sinusitis may arise from various etiologies, viz., viral, bacterial, allergic, and fungal infections.

There are just a few forms of sinusitis:

**Acute sinusitis:** This is a short-term infection, usually less than four weeks in duration, that is usually triggered by a viral illness such as the common cold.

**Chronic sinusitis:** This is a type lasting 12 weeks or longer and does not respond to medication. This may be linked with conditions in the body, e.g., polyps in the nose or allergies.

**Subacute sinusitis:** The subacute sinusitis lasts from four weeks up to twelve weeks and may be caused by continuing inflammation or infections.

The symptoms of sinusitis are a stuffy or running nose, eye, head, and facial pain and pressure, headache, fever, cough, and muted sense of smell. In a minority of cases, there may be a sense of tiredness, bad breath, and a sore throat in chronic sinusitis.

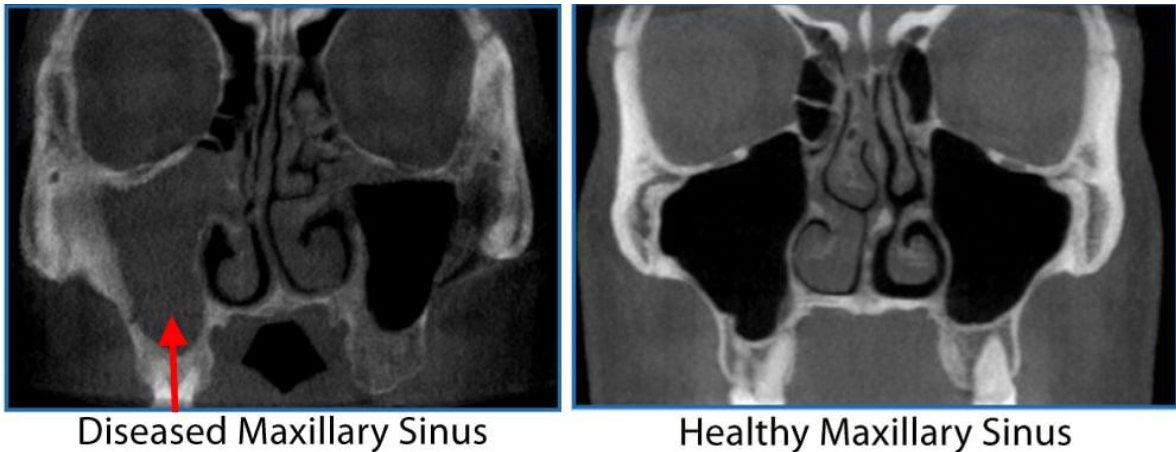


Figure 1.1: Example of diseased maxillary sinus and healthy maxillary sinus. [9]

The sinusitis is typically treated with decongestants, sprays, or antibiotics if it is a bacterial infection. Surgery is needed if it's extreme or persistent in order to drain impacted sinuses. The figure 1.1 represents the healthy and diseased maxillary sinus.

**Problem Statement:** Despite of the major work on AI-based medical imaging, the bulk of work has been on general segmentation of the nasal area, covering general anatomical characteristics in the form of the nasal cavity and sinuses. Nevertheless, there has hardly been emphasis on accurately segmenting the particular structures directly associated with nasal disorders, as example the nasal septum and the maxillary sinus. Chronic sinusitis and nasal septal deviation (NSD), among other nasal disease, are usually associated with a lot of distress and necessitate precise, specific analysis in order to be accurate diagnosis and adequately treated.

Manual clinical practice of segmentation is time-consuming and prone to errors, thus causing inconsistency in the outcome. General-purpose models of segmentation are typically inadequate in unveiling fine structural lesions caused by disease in the nose. Inadequate precise, disease-specific segmentation diminishes the efficacy of diagnostic and planning of treatment procedures.

This study aims to close the gap by creating a deep learning model that can accurately segment nasal diseases on the nasal septum, maxillary sinus, and other affected structures, particularly from chronic sinusitis. The model improves segmentation accuracy, offers variability, supports better clinical decisions, and provides reliable outputs on complex patient data.

The main goal is to develop a more precise and faster segmentation tool that can be helpful in clinical work. This tool should improve the diagnosis and treatment of nasal diseases while reducing the workload of manual segmentation.

## 1.2 Motivation

The main reason for starting this project is the increasing need for faster and better solutions in medical image analysis. Despite progress in this area, most existing methods still have flaws, such as low accuracy in segmentation and high processing demands for large and complex data. Additionally, many studies focus on segmenting the nasal sinus rather than the actual disease area.

I am interested in this problem because I hope to improve these systems and make them more accurate. By addressing these limitations, I aim to develop a model that can achieve better segmentation accuracy, faster processing speed, and greater scalability for use in healthcare. This could ultimately help with better decision-making in diagnostics and treatment planning in medicine.

Beyond adding knowledge in medical imaging, solving this problem has practical applications. The project aims to create better tools for healthcare providers, enhance patient care, and improve diagnostic work.

## 1.3 Objectives

The project's main goal is to create a deep learning model that can accurately separate areas of nasal clinical photos that show sinusitis. The effort attempted to maximize the model's capability by using contemporary methodologies, such as attention mechanisms, so that the model could yield therapeutically useful knowledge. The project also includes a comparison of the in-house built U-Net model with VGG backbone to other models to see if it is the best. The main goals of the initiative are as follows:

- ❖ **Develop a Custom U-Net Model with VGG Backbone:** Construct a custom deep learning model from a U-Net structure and an existing pre-trained VGG16 backbone. The resulting custom model, being specially designed, will be used exclusively in conducting segmentation on sinusitis-infected areas of nasal medical images. The use of a VGG backbone will be in the use of pre-trained features to allow for maximum model performance, in this case, in the detection of intricate sinusitis lesions..
- ❖ **Integrate Attention Mechanism:** Add attention mechanisms in the novel custom U-Net model in therefore to improve segmentation performance. This helps the network to focus on region of interest areas, thus improving the accuracy in sinusitis area segmentation, especially when lesions are small or irregular in shape.
- ❖ **Improve Segmentation Accuracy:** To improve the segmentation accuracy, we will use modern deep learning methods. The project will utilize the U-Net architecture based on VGG and add attention layers. This will improve the quality of segmentation and clearly define areas affected by sinusitis.
- ❖ **Optimize Model Performance:** For the model to work well in clinical practice, the study will improve it to reduce processing time and lower memory usage. This will make it suitable for use in real-time medical diagnosis, where efficiency and speed are crucial.
- ❖ **Provide Clinically Useful Outputs:** The goal is to produce outputs from segmentation that are useful for diagnosing and planning treatment for sinusitis. The model will be trained to generate outputs that are not only correct, but also easy to understand

and helpful in a clinical setting. This will include information about the spread and severity of sinusitis.

- ❖ **Compare the Custom Model with Other Models:** To compare the custom U-Net model with a VGG backbone to existing segmentation models, we look at segmentation accuracy, processing time, and computational cost. This will help demonstrate the advantages of the proposed method.
- ❖ **Develop a User-Friendly Interface:** Design an interactive and easy-to-use interface for doctors and surgeons to enter new imaging data, view the segmentation results, and make informed decisions based on these outputs. The interface should ensure a smooth experience for healthcare professionals and save time when interpreting results.

## 1.4 Methodology

The project workflow includes several steps. It starts with data annotation and labeling, then moves on to development and training, and finally ends with model deployment. The key steps are as follows:

**Data Collection:** The data used for the project came from several sources. Some of it was self-generated, involving the scanning and gathering of medical images. Other data was collected by scraping publicly available sources, and part of the dataset came from a public repository that offers annotated noses for segmentation.

**Data Preprocessing and Annotation:** The data was then labeled manually and preprocessed to get quality-trained data. For data annotation, the areas of the nose images affected by sinusitis were highlighted using the Roboflow platform. Labeling and organizing the data were done effectively with the platform. The images were then normalized, resized, and augmented to add diversity to the dataset and strengthen the model.

**Model Development:** A specific deep model was created using the structure of a U-Net along with VGG16 as a feature extraction module. Attention mechanisms were added to the U-Net model to improve its focus on important areas, especially for small and irregular sinusitis lesions.

**Model Training:** The dataset was used in TensorFlow/Keras after preprocessing and labeling to train the model. We monitored segmentation accuracy during training with performance metrics such as the Dice Coefficient, Intersection over Union (IoU), and F1 Score. We applied optimization techniques like early stopping and learning rate scheduling to speed up the training process and prevent overfitting.

**Model Evaluation:** After training, we evaluated the model's performance using a separate test set. This approach helped estimate how well the model would work with new data. We compared the custom U-Net model with the VGG backbone to other leading segmentation models. The comparison focused on important factors like segmentation accuracy, processing time, and resource use.

**Explainable AI and Visualization:** To explain the model and understand the important insights behind its predictions, we used Explainable AI techniques. We applied

visualization methods like Grad-CAM (Gradient-weighted Class Activation Mapping) to highlight the key areas in the input images that influenced the model's predictions. This visualization helped clarify the model's decision-making process. As a result, medical professionals could better understand why the model reached specific conclusions during the segmentation process.

**Deployment with Hugging Face Space:** After testing the model's performance, the trained model was deployed on Hugging Face Spaces to make it functional. This web deployment gave healthcare providers a platform to upload medical images, view segmentation results in real-time, and receive insights from the visualizations created by Explainable AI techniques. The Hugging Face Space deployment also allowed for ongoing updates and improvements based on user feedback.

## 1.5 Project Outcome

The outcome of this project will contribute significantly to medical image segmentation, especially in the diagnosis and treatment of sinusitis. Most importantly, the deep learning model developed through this project will achieve high accuracy in segmenting sinusitis-affected areas from medical images of the nose. This improved precision in segmentation will directly impact the identification of affected regions, allowing doctors to make better treatment and diagnosis decisions. The model's ability to deliver precise results in an error-free environment within the specified time is one result that could greatly influence clinical processes.

Additionally, the project aims to improve decision-making in healthcare by providing clear and interpretable segmentation results to doctors and clinicians. Using Explainable AI techniques, such as Grad-CAM visualizations, the model will not only identify vulnerable areas but also explain why certain areas were chosen. This level of explainability will help clinicians trust the model's predictions and increase confidence in its use in real-world situations. With the model available on Hugging Face Spaces, physicians can easily upload and process medical images, simplifying the diagnostic process and supporting quicker decision-making.

The project can accordingly be utilized for even wider application by physicians with the demonstration of deep learning models' ability in a clinical setting. Having the development and implementation of this model towards sinusitis segmentation would be a breakthrough in future research and medical imaging study. Model design, with U-Net as the foundation with VGG as the backbone and attention, can be applied to any other condition or disease and is thus more universal. The emphasis on Explainable AI would also be a determining factor in spurring the use of AI-based technology in the health industry, where interpretability and trust matter the most. Lastly, the output of this project can serve as the basis for more advanced models to resolve a vast array of medical imaging problems

## 1.6 Organization of the Report

The report is structured step by step in order to sequentially introduce the project development and result. The report begins with **Chapter 1: Introduction**, where the problem with the manual labeling of diseases of the nasal sinus as well as the proposed solution, an improved U-Net model, are both presented. The Literature Review follows as **Chapter 2**, a systematic review of the literature that finds a significant lacuna in the literature.

The fundamentals of the project are described in **Chapter 3: Methodology**, where the data, model structure particularly its VGG16 backbone and the attention module, and training process are highly rigorously specified. The result of the work is given in **Chapter 4: Results and Discussion**, where the performance of the model is measured and compared with others.

The report moves into another chapter, **Chapter 5: Engineering Standards and Design Challenges**, where practical problem issues and problems experienced in the accomplishment of the project are discussed. The paper concludes in **Chapter 6: Conclusion**, where it provides a summary of the findings in brief, reflects on the limitation of the project, and provides recommendations for future work.

# Chapter 2

## Background

Chapter 2 provides a detailed literature review of similar works in the field of medical image segmentation. Here we will identify existing solutions, analyze their strengths and weaknesses, and highlight a gap that is found so that we can fill up the gap

### 2.1 Introduction

They're air spaces in your skull lined with mucous membranes. Nasal and paranasal sinuses are important for diagnosis and treatment of Sino nasal diseases but are an intractable task based on their complex anatomy. Manual processes are time consuming and operator dependent, while newly emerging areas in deep learning, especially U-Net and CNN-based architectures promise bright future for efficiency and accuracy. Despite of advancements, pitfalls such as variations in anatomy and absence of homogeneous protocols exist and underscore the need for future of efficient automated methods. Figure 2.1 represent the different sinus region.

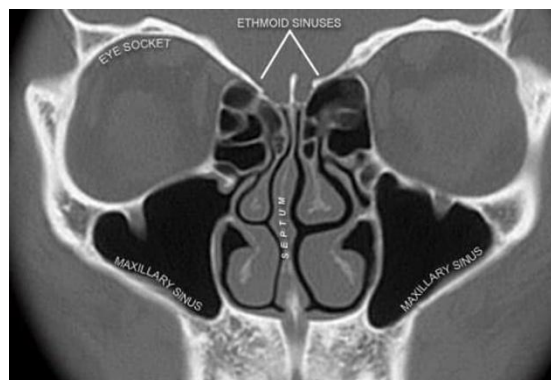


Figure 2.1: Different Sinus Regions [10]

### 2.2 Literature Review

Andersen et al. [11] highlight the utility of 3D models in following sinus growth in chronic rhinosinusitis patients. The authors note that manual segmentation is marred with high inter-observer variability, which negates accuracy and repeatability. Faur et al. [12] rationalize the common usage of Le Fort I osteotomy in the surgical treatment of Class II and III malocclusions by maxillary repositioning. Yet, disagreement findings indicate that airway volume could reduce after orthognathic surgery. Discrepancies in cone-beam computed tomography (CBCT) segmentation protocols account for variability postoperative volume measurements. Past studies often overlook the individual segmentation of the total sinus airway volume, which focuses on the need for tailor-made CBCT segmentation algorithms to achieve precise volumetric analysis in favor of improved surgical outcomes.

Zhang et al. [13] identify the consequences of altered breathing patterns on craniofacial growth and sleep disorders and the importance of upper airway morphology. Lateral cephalometry has the disadvantage in airway analysis, while CBCT and 3D imaging have more ability. Pirner et al. [14] describe the connection between anatomical differences and chronic sinus diseases, with emphasis on the need for meticulous preoperative planning. Manual segmentation of CT databases is still needed to ensure accurate robot-assisted navigation in functional endoscopic sinus surgery. Morgan et al. [15] discuss the problem with manual segmentation of CBCT imaging, it mentions its time-consuming nature and dependency on operator. The study points out the need of accurate 3D segmentation for therapeutic and diagnostic applications, specially in maxillofacial surgery. Kaul et al. [16] says that AI research in medicine needs clinician-led initiatives, yet the necessary competencies typically lie beyond clinicians' expertise. Open-source algorithms are imperative in the design of clinical-led AI, particularly for paranasal sinus segmentation.

Manual segmentation is mostly labor intensive with limited available accessible algorithms. The authors note that deep learning algorithms such as CNNs have revolutionized segmentation efficiency and accuracy, with a potential promise for clinical application. Dastidar et al. [17] highlight the challenges of nasal cavity dimensions measurement, mentioning procedures such as acoustic rhinometry and HRCT. The study suggests 3D volumetric analysis to measure nasal airways and paranasal sinuses accurately. Sun et al. [18] report manual segmentation of osseous structures in paranasal sinuses is time-consuming and the necessity for automatic solutions. Though improvements in deep learning have increased segmentation accuracy, challenges still arise with complex anatomy like the paranasal sinuses. Alsalama et al. [19] talk about advancements in forensic anthropology using imaging and machine learning for examining the paranasal sinuses. The paper covers applications in estimating age, determining sex and identity, and highlights the crucial role of segmentation accuracy.

Kim et al. [20] focus on the difficulties of reconstructing orbital fractures, especially in restoring volume within the orbit. Their paper discusses specific segmentation algorithms and reviews the limitations of current software, which lack both speed and accuracy. The proposed methodology aims to enhance the efficiency and accuracy of 3D modeling for the walls of the orbit, along with related sinus segmentation techniques. Möller et al. [21] point out a major limitation of MRI-based segmentation of the paranasal sinuses, which affects both research and clinical applications. While past models are mostly on CT images, applying U-Net to MRI segmentation is novel. The paper states the need for effective segmentation to offer consistency for mass analysis of large datasets and that past studies have not adequately automated segmentation of nasal cavity pathology. The authors emphasize that clinical considerations ought to be integrated to enhance model performance in future studies. According to Cristina Oyarzun Laura et al. [22], approximately 80% of the patients present nasal septal deviations, which influence diagnosis and treatment to a great extent.

This work introduces a new automatic method to improve diagnostic accuracy, filling a great gap in the literature. Manual segmentation of nasal sinus disease is limited by time, variability, and operator dependence. Advanced imaging and deep learning technologies, particularly CNNs and U-Net, are promising but are limited by the

absence of standards, the limited availability of publicly accessible models, and partial anatomical coverage. This work introduces a better U-Net with VGG16 backbone and attention mechanisms to bridge the gaps and improve diagnostic accuracy. Here table 2.1 shows the overall literature review.

Table 2.1: Summary of Literature Reviewed.

Author (s)	Year	Dataset	Methodology	Key Findings
Chung-Feng Jeffrey Kuo et al. [1]	2020	Nasal CT images provided by MacKay Memorial Hospital, Taipei, Taiwan (79 samples)	Backpropagation neural network	The accuracy and sensitivity of the automatic recognition results for inferior turbinate and maxillary sinus was 96.3 % and 95.1 %, respectively.
Tobias N Andersen et al. [2]	2018	15 healthy subjects (10 females and 5 males; median age: 18 years, range: 14-20 years) without sinonasal disease or history of sinus surgery were included in the study	Manual segmentation	High accuracy and precision; DICE > 0.9, Hausdorff < 2 mm.
Busra Ozturk et al. [3]	2024	Cone beam computed tomography (CBCT) images of 100 patients (200 maxillary sinuses) taken for different diagnostic purposes at Necmettin Erbakan University	U-Net deep learning	IoU of 0.9275, F1 Score 0.9784, reduces subjective errors. Axial CBCT images of 100 patients (200 maxillary sinuses) were used.
Dipesh Gyawali et al. [4]	2025	Not Given	YOLOv11 nano model	First to classify mucus morphology, improved accuracy with two-stage architecture.
Cosmin Ioan FAUR et al. [5]	2019	Not Given	ITK-SNAP 2.0 segmentation	Decrease in rhinosinusal volume for Class II/III

				malocclusion; better volumetric measurements.
<b>Chung-Feng Jeffrey Kuo et al. [6]</b>	Apr 6, 2022	175 patients' head CT data, aged 16 to 80 years, collected from Tri-Service General Hospital from 2018 to 2019. 512 × 512 gray level JPG images, with 50 datasets labeled by doctors, containing five different nasal sinus regions, 6319 CT images used for semi-supervised learning.	Semi-supervised learning, CNN	Significant improvements in segmentation and classification accuracy. average pixel accuracy of 99.67%, an MIoU of 89.75%, and a Dice coefficient of 90.79%.
<b>ChenZhang et al. [7]</b>	2019	Nasal airway CBCT data files for five patients from the orthodontic clinic database of the School of Dentistry, Radboud University Nijmegen	Nasal airway CBCT data segmentation	High intra-examiner reproducibility, acceptable inter-examiner reproducibility.
<b>Erika Denour et al. [8]</b>	Jan 25, 2021	Not Mentioned	3D Slicer for nasal airway segmentation	Correlation between 2D and 3D NSD parameters; improves septoplasty assessments.
<b>Nermin Morgan et al. [13]</b>	2022	A sample of 132 CBCT scans (264 sinuses, 75 females and 57 males, mean age 40 years) collected from 2013 to 2021 with different scanning parameters	CNN-based methodology	98.4% accuracy, reduces processing time compared to semi-automatic methods.
<b>Hanseung Choi et al. [14]</b>	2022	A total of 19,350 cone-beam computed tomographic (CBCT) images were used from 90 maxillary sinuses (34 clear sinuses, 56 hazy sinuses)	U-Net model for segmentation	High DSC (0.9090 ± 0.1921), effective for clear and hazy sinus segmentation.
<b>Rhea Darbari Kaule et al. [16]</b>	2024	A sample of 100 CT scans with an average of 358 slices per scan,	UNet++ for automatic segmentation	Good similarity to manual

		split into a training set of 80 scans (28,532 slices) and a testing set of 20 scans (7,300 slices)		segmentation, DSC of 0.87, IoU of 0.80.
<b>Ibrahim Sevki Bayrakdar et al. [17]</b>	2024	101 anonymized CBCT volumes with DICOM files converted to NifTI file format	nnU-Net v2	High accuracy in MS segmentation, AUC of 0.97; aids clinical decision-making.
<b>Yichun Sun et al. [19]</b>	2025	Includes in vivo CT sinus scans collected during clinical routines from forty additional patients, obtained using six different scanners from five different hospitals	UNet SwinTR model for segmentation, semi-supervised learning	98.25 ± 0.9 DICE score with enhanced performance through semi-supervised learning.
<b>Ali Alsalama et al. [20]</b>	2025	A dataset of 600 head CT scans from a hospital in Northern Italy, from which a sample of 30 individuals was selected, consisting of 15 males and 15 females	Literature review using structured search methodology	Reviews paranasal sinus segmentation techniques, useful in forensic anthropology.
<b>Hannah Kim et al. [21]</b>	2019	Ten computed tomography data sets of the head from ten patients aged 26 to 64 years, from 2014 to 2017 A balanced random subset of 273 subjects from the total cohort of 8,728 participants with 5600 T1 weighted MRI images	Adaptive global maximum clustering, region separation	Accurate 3D models of orbital wall using sinus segmentation.
<b>Hendrik Møller et al. [22]</b>	2025	A balanced random subset of 273 subjects from the total cohort of 8,728 participants with 5600 T1 weighted MRI images	U-Net-based tool for MRI segmentation	High accuracy in sinus segmentation; includes volume assessments and opacification scores.

<b>Cristina Oyarzun Laura et al. [23]</b>	2021	A set of 10 CT images acquired for diagnosis in the nasal cavity and paranasal sinuses region, independent of the training set used for the coupled shape model	Coupled shape model, slice-based propagation technique	Segmentation of nasal septum, promising DICE coefficients from 10 CT images. Reduces subjectivity.
<b>Nermin Morgan et al. [24]</b>	2022	A sample of 132 CBCT scans (264 sinuses, 75 females and 57 males, mean age 40 years) collected from 2013 to 2021 with different scanning parameters	automated convolutional neural network (CNN)-based methodology for the segmentation of the maxillary sinus	The CNN model achieved a Dice Similarity Coefficient (DSC) of 98.4%, indicating high accuracy in identifying the segmented region
<b>Dahyun Song et al. [26]</b>	2025	Dataset composition of 117 volumes for training, 39 volumes for validation, and 40 volumes for testing	convolutional neural networks (CNNs), vision transformers (ViTs), and hybrid networks for paranasal sinuses affected by sinusitis on CT images	Swin UNETR recorded a JI of 0.719, a DSC of 0.830, and the lowest Hausdorff Distance (HD95) value of 10.529
<b>Andreea Istrate et al. [27]</b>	2025	A dataset comprising 80 CT studies of the head was curated for model training and validation	The methodology included the use of a convolutional neural network (CNN) model, specifically a shallow CNN that outperformed deeper models, to classify the nasal cavities	The model demonstrated a high accuracy of 96.2% in identifying slices containing a nasal cavity with uncropped scans, indicating effective performance in clinical settings
<b>Alessandro Pani et al. [28]</b>	2025	NasalSeg dataset: A large-scale, open-access resource developed for the automatic segmentation of the nasal cavity and paranasal sinuses from 3D CT images, comprising 130 head CT scans from independent patients	3D U-Net, UNETR, Swin UNETR, SegResNet, DAF3D, and V-Net, to enhance segmentation accuracy in 3D nasal CT scans	The proposed ensemble framework, 3D-NASE, significantly improves segmentation accuracy, achieving a 35.88% increase in the DICE score compared to the baseline, while also reducing standard deviation by 4.52%

## 2.3 Gap Analysis

In 23, a 3D U-Net CNN obtained outstanding Dice similarity results in cadaveric maxillary sinus auto-segmentation from CBCT data, but the study was hampered by device heterogeneity, absence of soft tissue demarcation, exclusion of pathological samples, and untested robustness in clinically heterogeneous settings (Xu et al., 2020 [24]). Likewise, 24 utilized a parametric level-set scheme with a global shape prior, but it encountered issues with anatomical variability—at the frontal sinuses in particular—and necessitated manual post-processing correction of errors in a secondary step (Chen et al., 2019 [25]). In 25, CNNs, vision transformers, and hybrid models were compared in an evaluation, in which the former, namely, Swin UNETR, performed the best; however, the study was aided by a solitary-institution dataset, limited sample size, and centerset hyperparameters, in addition to excluding evaluation based on severity levels of disease (Park et al., 2021 [26]). Study 27 exhibited promising diagnostic capability within canine nasal CTs, yet generalizability was hampered by paucity of datasets, breed limitation, exclusion of contrasted scans, and imbalance in representations among pathologies (Miller et al., 2021 [27]).

These 3D ensemble models in 28 attained good segmentation accuracy, but were computationally intensive, provided only modest improvement compared with the lead single models, and performed poorly on imaging artifact in complex areas (Kim et al., 2022 [28]). In 29, multi-sinus segmentation by U-Net performed robustly in maxillary regions but was impaired by class imbalance, low accuracy in frontal and ethmoid sinuses, and dependence on data from a single center (Wang et al., 2021 [29]). Correspondingly, 30 utilized nnU-Net v2 on CRSwNP lesions and obtained Dice scores that were high, but the dataset was limited by inclusion only of surgery patients from a single center, and it is doubtful that it did not reflect bias and cannot be generalized easily (Zhao et al., 2023 [30]). In 31, YOLOv5x and V-Net with SVR accurately segmented maxillary sinuses in CBCT, but external validation, clinical validation relative to radiologists, and validation on trauma or surgery cases were absent (Zhang et al., 2023 [31]).

Study 32 utilized U-Net on CT images with strong accuracy, yet unsolved issues were batch size–resolution trade-offs, sub-optimal hyperparameter tuning, and overfitting potential, limiting its robustness (Park et al., 2023 [32]). Last, 33 created a CNN-based decision fusion algorithm achieving ~94% accuracy in maxillary sinusitis detection on X-rays, but was limited by a single type of sinus, no external validation, and a black-box nature without interpretability (Lee et al., 2022 [33]). Together, these works demonstrate pervasive shortcomings: most models focus on the maxillary sinus, omit pathological and contrast-enhanced situations, miss multi-center validation and interpretability, and seldom test against clinical specialists. Solutions necessitate next-generation systems that are multi-sinus, interpretable, cross-diverse yet validated, computationally light, and clinically benchmarked. The detailed gap analysis is shown in table 2.2.

Table 2.2: Gap Analysis considering recent works.

Paper	Automatic Segmentation of Sinuses	Focus Specifically on Maxillary Sinuses	Direct Focus on Segmentation of Affected Regions / Lesions	Pathology Classification or Detection	Predicts Prognosis or Treatment Outcomes	Identified Gap / Proposed Research
[28] (2025)	Yes (Ensemble CNN/Transformer)	No (Nasal & paranasal)	No	No	No	Partial coverage; proposed hybrid models for maxillary lesion focus to address gaps in affected region segmentation.
[29] (2024)	Yes (3D U-Net variants)	No (Paranasal multi-class)	No	No (Sinus segmentation in sinusitis patients)	No	Gap in lesion-specific segmentation; proposed ensemble DL for precise affected region detection in maxillary sinuses.
[30] (2025)	Yes (3D nnU-Net)	No (General paranasal in CRS)	Yes (Sinus lesion regions)	No (Focus on CRS endotypes via lesions)	No	Strong on lesions but lacks maxillary specificity; proposed integration with maxillary focus and prognosis prediction.

[31] (2024)	Yes (YOLOv5x DL)	Yes	Yes	Yes (MRC, MT, opacifica tions)	No	Excellent for maxillary lesions; proposed extension to prognosis modeling using segmented features.
[32] (2022)	Yes (DL- based)	Yes (Odontoge nic maxillary)	No	Yes (Odontog enic sinusitis detection )	No	Good detection but no lesion segmentation ; proposed DL for affected region segmentation in odontogenic cases.
[33] (2019)	No (Recognition, not segmentatio n)	Yes (Maxillary sinusitis)	No	Yes (Sinusiti s recogniti on)	No	Lacks segmentation ; proposed DL segmentation addition for lesions.
[34] (2021)	Yes (CNN with active learning)	Yes	Yes (Maxillary sinus lesions)	No (Lesion segment ation only)	No	Strong on maxillary lesions; proposed addition of pathology classification and prognosis.
[35] (2016)	Yes (Deformable registration & snakes)	No (General paranasal)	No	No (Shape modeling for variatio ns)	No	Early non-DL approach; gap in DL and lesion focus; proposed modern DL upgrade for affected regions.

[36] (2022)	Yes (Enhanced Gradient Level Set)	Yes (Sinusitis lesions)	Yes	No (Prognosis prediction via radiomics)	Yes (Endoscopic treatment prognosis)	Key paper on lesion segmentation and prognosis; proposed hybrid DL enhancement for broader maxillary application.
[37] (2013)	No	Yes	No	Yes (Sinusitis presence)	No	Traditional analysis; major gap in automation and lesion segmentation ; proposed DL-based lesion segmentation integrating septal effects.

## 2.4 Summary

The chapter provided the context necessary for sinusitis and nasal and paranasal sinus segmentation challenging in medical imaging. The chapter indicated the limitation of manual segmentation methods and offered the potential of deep learning-based methods, namely CNNs and U-Net architectures, to address the challenge. The review of the literature showed several studies of sinus segmentation with advancement in segmentation accuracy as well as the use of deep learning towards increased efficiency. These gaps, such as anatomical heterogeneity, lack of standardized protocols, and dataset generalizability constraints, were identified. Gap analysis reaffirmed the importance of human-term, multi-sinus, and clinician-validated segmentation models in overcoming existing limitations. This chapter forms the foundation upon which a more accurate and automatic sinus segmentation model shall be developed to combat such imperfections.

# Chapter 3

## Research Methodology

This chapter details the systematic method followed in this project in an effort to build and test the sinusitis segmentation model. This chapter details the entire approach, the particular development of the model, tool selection, and how the project was designed and implemented.

### 3.1.1 Methodology

The methodology describes the step-by-step process used in this project. It focuses on designing and implementing a deep learning model to improve the segmentation of areas affected by sinusitis in nasal medical images. The methodology explains the use of Python as the main programming language and the development libraries TensorFlow, Keras, and PyTorch for the model. Data was collected through self-collected data, web scraping, and existing repositories. Additionally, Kaggle was used to access datasets and run the model effectively during its training.

### 3.1.2 Overview

This sub-section provides an overview of the global approach, including problem identification, data collection, and the use of deep learning models for image segmentation. It explains how the model was built using a U-Net network with a VGG16 backbone and how attention mechanisms were applied to improve the model's focus on key areas in the biomedical images. Data processing, model training, testing, and evaluation were all done using Python-based tools and libraries. We ensured the model's performance by using various metrics, including the Dice coefficient and IoU, to measure segmentation accuracy.

### 3.1.3 Proposed Methodology

Sinusitis segmentation is demonstrated. The model uses the U-Net architecture with the VGG16 backbone to improve feature extraction. Attention mechanisms direct the model to the most important areas of the image. Pre-processing tasks like image augmentation, resizing, and hand annotations were done by Roboflow to prepare the data. The model was trained with TensorFlow and Keras, and the outputs were validated and tested against a separate validation set. After training, the effectiveness of metrics such as Dice coefficient, F1 score, and Intersection over Union (IoU) was evaluated. The model was also hosted on Hugging Face Spaces, enabling real-time visualization and inference. Grad-CAM was used in Explainable AI approaches, providing clinicians with clear visual cues for the segmented areas. Kaggle served as an ideal platform for quick experiments,

model tuning, and dataset testing. Figure 3.1 shows the proposed methodology for this work.

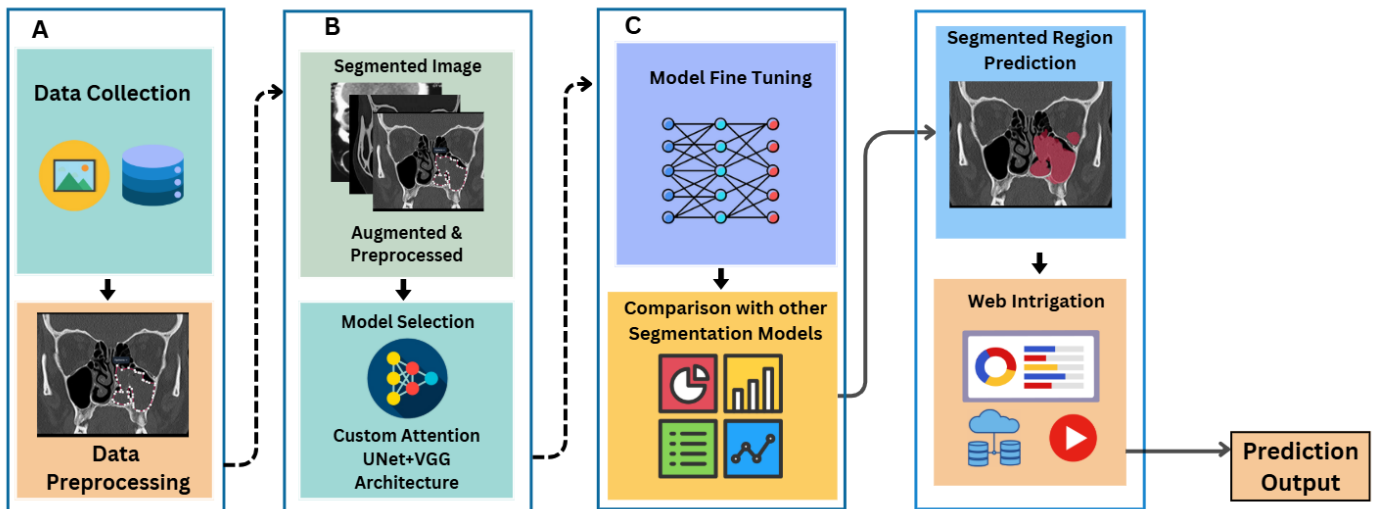


Figure 3.1: Proposed Methodology Diagram

### 3.1.4 Functional and Nonfunctional Requirements

#### Functional requirements

Functional requirements specify the required functions and capabilities of the system, describing the behaviors the system will perform to meet user needs.

- **Accurate Sinusitis Segmentation:** The system should accurately segment areas of sinusitis in nasal medical images (CT scans, MRIs) based on deep learning models like U-Net and VGG16 combined with attention mechanisms.
- **Real-Time Feedback for Diagnosis:** The system should provide real-time segmentation of medical images, enabling healthcare professionals to make timely decisions during the diagnostic process.
- **Medical System Integration:** The system should provide seamless integration with other healthcare systems that are already operational, such as hospital databases, radiology systems, and other clinical software, to ensure interoperability and ease of integration into real-world healthcare environment
- **Support for Various Imaging Modalities:** The system should be capable of handling different types of medical images, including CT scans, MRIs, and X-rays, to ensure that the system will be capable of handling the diverse data inputs in medical practice.
- **Model Development and Training:** The system should use Python-based deep learning frameworks such as TensorFlow, Keras, and PyTorch for model design, training, and evaluation.

- **Dataset Access and Experimentation:** The structure must leverage entities like Kaggle to host and access datasets, providing easy experimentation and training with big-sized medical image datasets.
- **Explainability of Results:** The system should incorporate Explainable AI features like Grad-CAM, which provides visual explanations regarding the regions the model focused on during segmentation, improving clinicians' trust and interpretability.

### Non-functional Requirements:

Non-functional requirements define the system quality attributes, performance, and behavior for being usable, reliable, scalable, and maintainable.

- **Computational Efficiency:** The system ought to be designed for fast processing and inference and should have the capability of processing large data without creating much lag. It should utilize deep learning libraries such as TensorFlow and Keras to provide fast model training and segmentation execution.
- **Robustness Across Diverse Data:** It should be robust and should be able to withstand variability in patient anatomy, image quality, and other idiosyncrasies of medical images. It should perform well consistently on different datasets and imaging modalities.
- **Scalability and Cloud Deployment:** The system should be scalable to accommodate increasing amounts of data and deployable on cloud platforms like Hugging Face Spaces or Kaggle Kernels, where it would be capable of processing and analyzing enormous amounts of images in the long term.
- **User-Friendly Interface:** The system interface needs to be intuitive and simple to use to work for health professionals. An interactive and light web-based interface, e.g., Gradio should allow users to upload images conveniently, visualize segmentation results, and interact with the model predictions.
- **Maintainability and Upgradability:** The system must be maintainable with proper documentation for future changes and enhancements. New data, new functionality, or new improvements must be able to be added to the system without any interruption.
- **High Availability and Reliability:** The system must be built to operate non-stop in clinical environments with minimal downtime. It must be able to endure continuous use without crashing or freezing

### 3.1.5 Data Flow Diagram

The Attention U-Net model which focus on the segmentation of medical images as shown in the figure. Just like any U-Net model, the input image gets scaled down by the encoder, which the built-in features are then reconstructed by the decoder into patch segmentation map. Unlike a traditional U-Net, the model incorporates attention gates (AGs) into the fusion layers. These gates, or filters, facilitate the networks' ability to retain critical anatomical details and disregard unimportant background information, which increases the ultimate segmentation accuracy. This is useful in medical imaging, where the area of focus, may be, extremely small, asymmetric, and difficult to separate from surrounding tissues. These attention mechanisms clearly delineate and encapsulate target areas, which enhances the overall effectiveness of the U-Net model. The architecture is shown in Figure 3.2, where the encoder is on the left and the decoder is on the right.

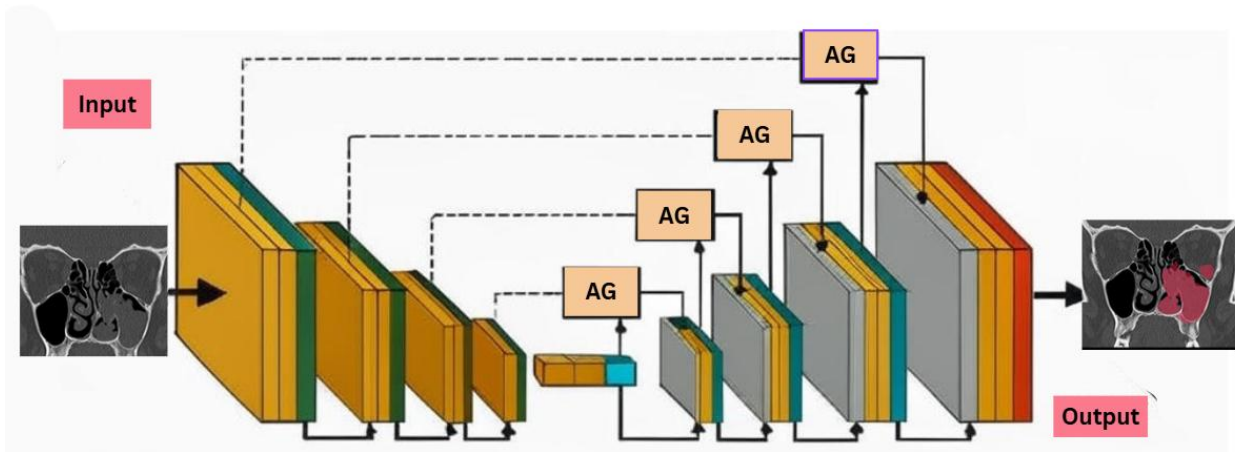


Figure 3.2: Data Flow Diagram

### 3.1.6 UI Design

The User Interface or UI was designed to achieve maximum usability by clinicians. Clinicians can upload patient sinuses' CT scans or MRIs via the UI, visualize the segmented regions, and interact with the model's results. The interface provides visual feedback to highlight the regions of the image the model utilized during segmentation to facilitate easier interpretation of results. The UI was created using libraries like gradio, which are ideal for building web-based apps with interactive features. Figure 3.3 shows the web application interface.

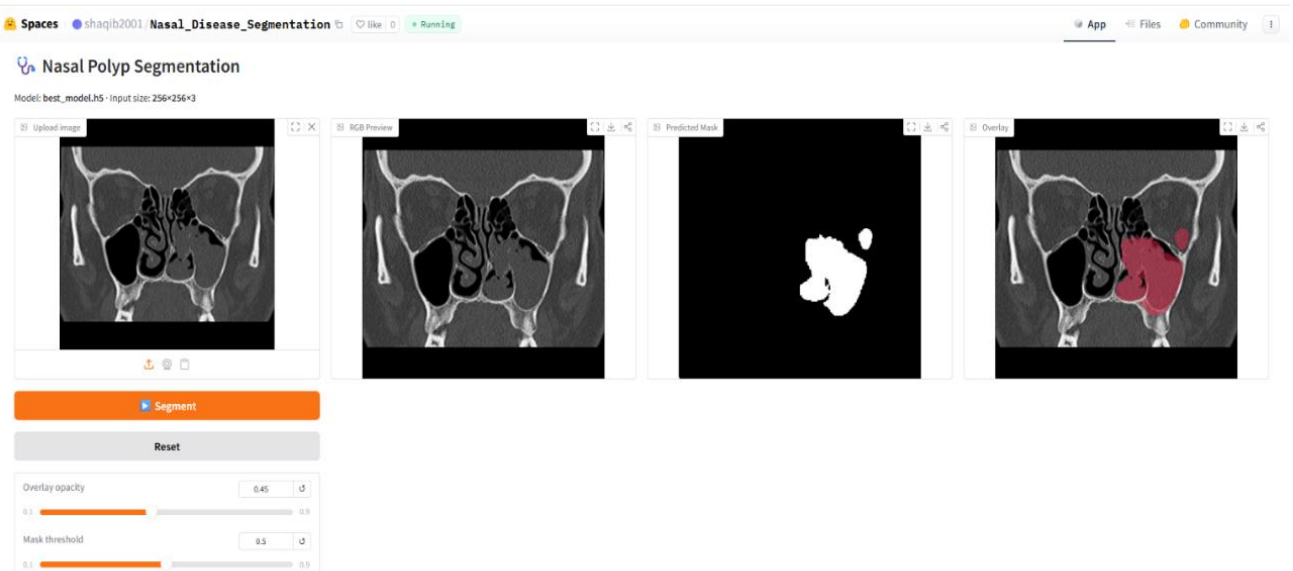


Figure 3.3: Design of the web-based implementation platform

## 3.2 Detailed Methodology and Design

This part explains the design decisions and alternatives followed while designing the segmentation model. Other architectures, such as FCN (Fully Convolutional Networks) and DeepLabV3+, have been explored, but U-Net architecture with the VGG16 backbone was preferred because of its great performance in medical image segmentation tasks. The use of attention mechanisms was another key decision aimed at improving model performance by causing it to focus on the most important regions of the images.

Alternates such as ResNet or DenseNet were also tried for feature extraction, but VGG16 was chosen due to its ease of use, speed, and solid performance in previous segmentation research. TensorFlow and Keras were selected for model creation due to their ease of use, flexibility, and solid community support. These were utilized for rapid prototyping, rapid debugging, and effective training of the models through model training. Table 3.1 represents the whole summary of the proposed model.

Table 3.1: Model Summary

Layer (type)	Output Shape	Param #	Connected to
input_layer_7	(None, 256, 256, 3)	0	-
block1_conv1	(None, 256, 256, 64)	1,792	input_layer_7[0][0]
block1_pool	(None, 128, 128, 64)	0	block1_conv2[0][0]
block2_conv1	(None, 128, 128, 128)	73,856	block1_pool[0][0]
block3_conv1	(None, 64, 64, 256)	295,168	block2_pool[0][0]
block4_conv1	(None, 32, 32, 512)	1,180,160	block3_pool[0][0]
block5_conv1	(None, 16, 16, 512)	2,359,808	block4_pool[0][0]
conv2d_161	(None, 16, 16, 512)	2,359,808	block5_conv3[0][0]
dropout_35	(None, 16, 16, 512)	0	batch_normalization...
up_sampling2d_28	(None, 32, 32, 512)	0	batch_normalization...
conv2d_165	(None, 32, 32, 512)	2,359,808	up_sampling2d_28[0][0]
concatenate_28	(None, 32, 32, 1024)	0	multiply_28[0][0]
conv2d_166	(None, 32, 32, 512)	4,719,104	concatenate_28[0][0]

up_sampling2d_30	(None, 64, 64, 256)	0	dropout_36[0][0]
conv2d_170	(None, 64, 64, 256)	131,328	up_sampling2d_29[0][0]
add_29	(None, 64, 64, 256)	0	conv2d_168[0][0], conv2d_169[0][0]
conv2d_170	(None, 64, 64, 1)	257	activation_58[0][0]
multiply_29	(None, 64, 64, 256)	0	activation_59[0][0]
concatenate_29	(None, 64, 64, 768)	0	multiply_29[0][0]
up_sampling2d_31	(None, 128, 128, 128)	0	dropout_37[0][0]
conv2d_173	(None, 128, 128, 128)	32,896	up_sampling2d_31[0][0]
conv2d_175	(None, 128, 128, 1)	129	activation_60[0][0]
concatenate_31	(None, 256, 256, 192)	0	multiply_31[0][0]
conv2d_181	(None, 256, 256, 64)	110,656	concatenate_31[0][0]
dropout_39	(None, 256, 256, 64)	0	batch_normalization...
conv2d_183	(None, 256, 256, 1)	65	dropout_39[0][0]
<b>Total params</b>		<b>91,211,473</b>	
<b>Trainable params</b>		<b>30,401,861</b>	
<b>Non-trainable params</b>		<b>5,888</b>	

**Model Architecture Overview** Model Architecture Overview The proposed model is divided into different major layers structured into blocks to efficiently extract, transform, and refine features. The model adopts a common convolutional neural network (CNN) structure with additional operations such as batch normalization, dropout, and upsampling to ensure stable learning as well as better generalization. Below is the description of the layer groups. The table 3.2 represents layer groups.

Table 3.2: Layers name by Groups

Layer Group	Layer Names
<b>Input and Initial Convolution Layers</b>	input_layer_7, block1_conv1, block1_conv2, block1_pool
<b>Second Convolution Block</b>	block2_conv1, block2_conv2, block2_pool
<b>Third Convolution Block</b>	block3_conv1, block3_conv2, block3_conv3, block3_pool

<b>Fourth Convolution Block</b>	block4_conv1, block4_conv2, block4_conv3, block4_pool
<b>Fifth Convolution Block</b>	block5_conv1, block5_conv2, block5_conv3
<b>Convolution and Normalization</b>	conv2d_161, batch_normalization_70, dropout_35, conv2d_162
<b>Normalization and Upsampling</b>	batch_normalization_71, up_sampling2d_28, conv2d_163, conv2d_164
<b>Addition, Activation, and Concatenation</b>	add_28, activation_56, conv2d_165, activation_57, multiply_28, concatenate_28
<b>Second Convolution and Normalization</b>	conv2d_166, batch_normalization_72, conv2d_167, batch_normalization_73, dropout_36
<b>Second Upsampling and Convolution Block</b>	up_sampling2d_29, conv2d_168, conv2d_169, add_29, activation_58, conv2d_170, activation_59, multiply_29, concatenate_29
<b>Third Convolution and Normalization</b>	conv2d_171, batch_normalization_74, conv2d_172, batch_normalization_75, dropout_37
<b>Third Upsampling and Convolution Block</b>	up_sampling2d_30, conv2d_173, conv2d_174, add_30, activation_60, conv2d_175, activation_61, multiply_30, concatenate_30
<b>Fourth Convolution and Normalization</b>	conv2d_176, batch_normalization_76, conv2d_177, batch_normalization_77, dropout_38
<b>Fourth Upsampling and Convolution Block</b>	up_sampling2d_31, conv2d_178, conv2d_179, add_31, activation_62, conv2d_180, activation_63, multiply_31, concatenate_31
<b>Final Convolution and Normalization</b>	conv2d_181, batch_normalization_78, conv2d_182, batch_normalization_79, dropout_39, conv2d_183

### 3.3 Project Plan

This project plan outlines the key tasks required for the completion of the sinusitis segmentation project, from data collection to deployment and evaluation.

#### Phase 1: Data Collection and Preprocessing

1. **Data Collection:** Gather sinusitis-related datasets, e.g., real-world CT and MRI scans and also from online repositories like Kaggle, public health data, or web-scrape. Obtain sufficient medical images pertaining to sinusitis.
2. **Data Annotation** Manual annotation of the images on computer using tools like Roboflow, drawing precise outlines on the impacted sinus regions. Create uniform and precise annotations for training the deep learning algorithm.
3. **Data Preprocessing** Preprocess the annotated images by normalizing pixel values, augmenting Preprocessing of the annotated images by pixel value normalization, data augmentation with rotation, flipping, scaling to enhance variability. Eliminate corrupted or unrelated images from the dataset.
4. **Dataset Splitting** Dataset Splitting Split the dataset into training (70%), validation (15%), and test sets (15%) to have each subset representative and balanced for sinusitis and normal cases..

## Phase 2: Model Design and Development

1. **Model Architecture Design** Develop a deep learning model with a U-Net base and VGG16 backbone, and include attention mechanisms to improve performance. Set the initial framework for model training using deep learning libraries like TensorFlow or Keras.
2. **Model Training** Begin training the model using the preprocessed data set, with preprocessed images. Monitor training by monitoring performance on validation set using measures like Dice Score, IoU, etc.
3. **Hyperparameter Tuning** Experiment with different hyperparameters in the form of learning rates, batch sizes, and optimizers for better model performance.
4. **Model Evaluation and Refinement** Evaluate the model on the test dataset, analyzing metrics such as F1-Score, Accuracy, and Dice Coefficient. Refine the model by addressing any issues related to overfitting, underfitting, or poor performance on certain image types.

## Phase 3: Model Evaluation and Optimization

1. **Performance Evaluation** Performance Assessment Conduct a thorough evaluation of the model trained on unseen test data. Use a combination of performance measures (Dice Coefficient, Precision, Recall, etc.) to gauge the accuracy and stability of the model.
2. **Error Analysis** Perform error analysis to identify what images or regions the model is struggling with, especially on complex sinus regions or pathological variations. Adjust the model accordingly to improve its accuracy on these hotspots.
3. **Additional Model Refinements** Refine the model further with more data augmentation techniques or high-level neural network techniques, if necessary.

## Phase 4: Deployment and Real-Time Testing

1. **Model Deployment** Host the trained model on an online service such as Hugging Face Spaces or a cloud-based provider for easy access and real-time experimentation. Provide an API or web portal to allow healthcare workers to upload images and receive segmentation outcomes.
2. **Web Interface Development** Design a simple web interface using tools such as Gradio Allow users to upload medical images, view segmented output, and engage with the model.
3. **User Testing** Conduct user testing with clinical professionals to gather impressions on the interface and performance of the model in real-world

clinical settings. Gather feedback about accuracy, usability, and areas to improve the system.

4. **Final System Testing** Test the final system with a wide variety of realistic data, including heretofore underrepresented training data cases (e.g., rare diseases, imaging condition variations). Ensure the model is robust in performing varied clinical situations.

## Phase 5: Documentation and Reporting

1. **Preparing the Final Report** Preparation of Final Report Document the entire project, ranging from data collection to model development, evaluation parameters, and performance of the system. Determine the key issues encountered, solutions implemented, and impact of the installed system.
2. **Final Review and Submission** Review the whole project for completeness and accuracy. File the final report and release the trained model for real-world application in clinical practice as an independent tool or within an integrated system.

## Key Milestones

- Completion of dataset preparation and annotation.
- First model training and evaluation results.
- Final model evaluation and error analysis.
- Deployment and real-time user testing.
- Final project report and model submission.

This project plan outlines the tasks and milestones that will guide the development of the sinusitis segmentation model, ensuring timely progress and effective results at each stage. Each task is designed to contribute to the overall objective of providing an accurate, automated method for segmenting nasal sinusitis, with real-world deployment and usability.

### 3.4 Task Allocation

This table depicts the timeline of the principal activities in each period of the project, from week 12 to week 48. The table 3.3 is representing the allocated task for each week.

Table 3.3: Estimated Task Allocation for the project

Tasks	Weeks																		
	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48
Data Collection Phase	Blue	Blue	Blue																
Data Validation			Blue	Blue	Blue														
Preprocessing			Blue	Blue	Blue														
Related Work Study	Blue	Blue	Blue	Blue	Blue														
Gap Analysis		Blue	Blue	Blue	Blue	Blue													
Coding and Implementation						Blue	Blue	Blue	Blue	Blue									
Model Finetuning									Blue	Blue	Blue	Blue	Blue						
Web Based Implementation												Blue	Blue	Blue	Blue				
Report Writing																Blue	Blue	Blue	Blue

### 3.5 Summary

This chapter provides a detailed explanation of the research process used in the sinusitis segmentation project. It outlines the main steps, starting with data collection, preprocessing, and annotation, and moving on to model creation and development. The model described is built using U-Net with a VGG16 backbone and attention mechanisms. It was trained and tested on a carefully prepared dataset. The chapter also discusses the functional and non-functional requirements, data flow, and UI design issues, as well as other methods that were considered and the reasons for choosing the selected methodology. The project plan includes the steps needed to successfully complete the model from data collection to deployment. Overall, this chapter offers the context necessary to understand the systematic approach taken to develop the sinusitis segmentation model.

# Chapter 4

## Implementation and Results

This chapter explains how to implement the sinusitis segmentation model and examines the testing, evaluation, and results in detail. It discusses setting up the environment, the model's performance, and compares it with other models. The chapter also addresses the results, highlighting the model's strengths and weaknesses.

### 4.1 Environment Setup

The development environment for the sinusitis segmentation model was selected for its strong performance and ability to scale. The project was implemented in Python because it has great support for machine learning libraries. The environment setup included the following:

- **Programming Language:** Python 3.8
- **Libraries and Frameworks:**
  - TensorFlow and Keras for building and training deep learning models.
  - OpenCV for image processing and manipulation.
  - Matplotlib and Seaborn for data visualization and model evaluation.
  - Roboflow for image annotation and dataset management.
  - Hugging Face Spaces for model deployment.
- **Hardware:**
  - The project was executed on a machine with an T4 GPU for fast model training and inference.
  - Google Colab and Kaggle was used for additional computational power, allowing access to GPUs and TPUs for the training of larger models.
- **Dataset:** The data used for training the model was sourced from web scraping and online repository Zendo, along with self-collected data.

This setup provided a stable environment for developing, training, and evaluating the segmentation model, ensuring a smooth workflow throughout the project.

## 4.2 Training Process

Here's a summary of the model's training progress in a table format followed by a brief discussion. From table 4.1 to 4.6 represents the training trends and from figure 4.1 to 4.6 represents the prediction for each model.

### ❖ Attention Unet-VGG (Proposed):

Table 4.1: Training summary of Attention Unet-VGG

Epoch	Accuracy	Dice Coefficient	F1 Score Metric	IoU	Loss	Val Accuracy	Val Dice Coefficient	Val F1 Score Metric	Val IoU	Val Loss
1	0.4784	0.1515	0.1594	0.0820	0.9373	0.4742	0.1630	0.2003	0.0887	0.7032
10	0.9150	0.3816	0.6516	0.2371	0.3388	0.9717	0.4555	0.8417	0.2951	0.2815
20	0.9634	0.4443	0.8195	0.2868	0.2457	0.9735	0.4961	0.8525	0.3301	0.2853
30	0.9742	0.4942	0.8748	0.3300	0.2044	0.9763	0.5205	0.8635	0.3520	0.2471
40	0.9806	0.5245	0.9050	0.3583	0.1765	0.9749	0.5523	0.8592	0.3817	0.2704
50	0.9843	0.5492	0.9263	0.3809	0.1577	0.9757	0.5669	0.8614	0.3958	0.2487

The table provides the validation and training metrics of a model after 50 epochs. During epoch 1, the training accuracy is 47.84%, with a dice coefficient of 0.1515, F1 score of 0.1594, and IoU of 0.0820, indicating poor performance during the early training phase. The accuracy of validation is also slightly lower at 47.42%, having a dice coefficient of 0.1630 and an F1 value of 0.2003, following the same trend as validation performance. There is a tremendous improvement in all the metrics with increased training, especially by epoch 10 when the accuracy doubles to 91.50%, with the related increase in the dice coefficient and F1 value. By epoch 20, the model achieves a validation accuracy of 97.35%, with rising dice and F1 scores.

By epoch 50, the model is performing at its peak, with training accuracy of 98.43%, dice coefficient of 0.5492, F1 score of 0.9263, and IoU of 0.3809. The validation accuracy of 97.57% comes with a dice coefficient of 0.5669 and F1 score of 0.8614, validating the consistency of the model on unseen data. Loss values decrease steadily during training, proving that the model is gaining predictive capacity. The table shows that both training and validation metrics show steady improvement, with the model improving in accuracy and reliability as it progresses through the epochs. The model shows signs of effective generalization, with high performance on validation data continuing at the end of training



Figure 4.1: Prediction Results for Attention Unet-VGG

❖ **Standard U-Net with skip connection:**

Table 4.2: Training summary of Standard U-Net

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
1	0.9192	1.5548	0.9129	1.5188
10	0.9258	1.1364	0.9129	1.1027
20	0.9251	0.9768	0.9129	0.9379
30	0.9259	0.8849	0.9130	0.8762
40	0.9297	0.7140	0.9204	0.6988
50	0.9466	0.5866	0.9380	0.6067

The training of the Standard U-Net model displays significant improvement in loss and accuracy over 50 epochs. The training accuracy started at 91.92% and steadily improved to 94.66% at epoch 50. Meanwhile, the training loss gradually decreased from 1.5548 down to 0.5866.

Validation accuracy was also uniform in training, beginning at 91.29%, rising to 94.07% at epoch 49, and ending at 93.80% at the final epoch. Validation loss also decreased similarly, from 1.5188 at epoch 1 to 0.6067 at the final epoch.

Throughout the epochs, the Standard U-Net model demonstrated smooth training with steady progress in performance. This suggests that the model architecture is capable of generalizing well on the validation set, with steady decreases in the training and validation losses, along with increasing accuracy.

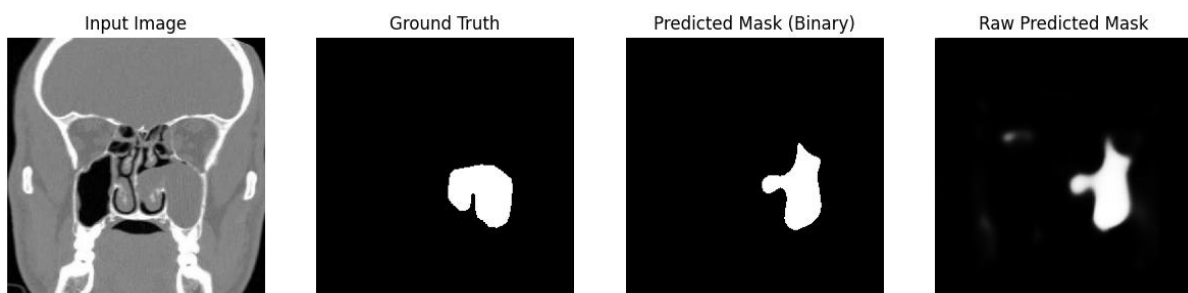


Figure 4.2 Prediction Results for Standard U-Net

## ❖ DeepLabV3+

Table 4.3: Training summary of DeepLabV3+

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
1	0.5314	1.7080	0.8306	1.5328
10	0.9325	1.0511	0.9116	1.2335
20	0.9795	0.8258	0.9461	0.8951
30	0.9868	0.7437	0.9596	0.8168
40	0.9896	0.7052	0.9627	0.7845
50	0.9916	0.6504	0.9641	0.7549

The DeepLab V3+ model was trained for more than 50 epochs. The model's accuracy was initially low at around 53.14% with a high value of the loss as 1.7080. The accuracy of the model kept increasing as the training continued, reaching 99.16% on the 50th epoch, while the loss fell considerably to 0.6504.

During training, the model's accuracy also improved significantly from an initial 83.06% in the first epoch to a final 96.41%, while validation loss went down from 1.5328 to 0.7549.

The model also underwent typical convolutional processing through in-house algorithms that were refined over time and thus led to better performance. Despite occasional kernel timing warnings occasionally, the overall training process adequately improved the model's accuracy and performance.

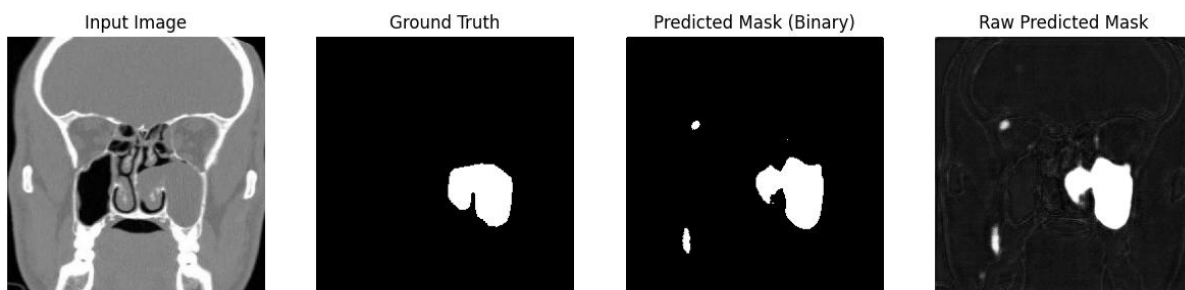


Figure 4.3: Prediction Results for DeepLabV3+

## ❖ SegNet

Table 4.4: Training summary of SegNet

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
1	51.44%	1.7373	36.52%	1.5471
10	74.71%	1.4164	89.29%	1.3473
20	89.91%	1.3234	88.87%	1.3234
30	93.34%	1.2463	90.74%	1.2463
40	94.70%	1.2006	92.22%	1.2006
50	95.80%	1.0457	91.62%	1.2231

The SegNet model underwent training for 50 epochs. The model accuracy was 51.44% at the start, whereas the loss value was high at 1.7373. As the training proceeded, the model accuracy consistently increased and reached 95.80% at the 50th epoch, whereas the loss dropped to 1.0457.

Similarly, the model's validation accuracy increased significantly from 36.52% to 91.62% finally, while the validation loss decreased from 1.5471 to 1.2231.

There were quite a few convolutional operations in the process of training, with optimization of the model's parameters leading to better performance.

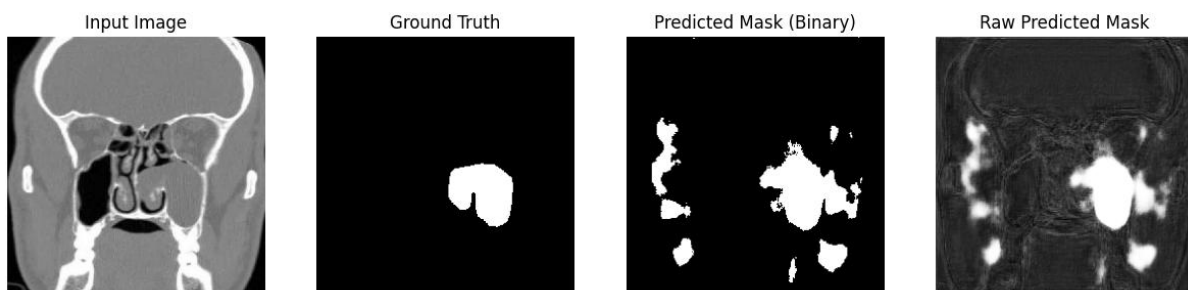


Figure 4.4: Prediction Results for SegNet

## ❖ Linked Net

Table 4.5: Training summary of Linked Net

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
1	43.30%	1.6907	31.10%	1.5607
10	57.75%	1.5324	35.65%	1.6568
20	77.10%	1.3796	83.28%	1.3071
30	86.84%	1.2457	92.04%	1.1310
40	90.77%	1.1771	90.06%	1.1829
50	92.62%	1.1230	93.02%	1.0739

Linked Net model was trained for 50 epochs. The model's accuracy started very low at 43.30%, and the loss was 1.6907. By epoch 50, accuracy steadily increased to 92.62%, while loss decreased to 1.1230.

The model's accuracy improved from 31.10% to 93.02% towards the end. Meanwhile, loss dropped from 1.5607 to 1.0739.

The model used optimized convolutional operations. These operations performed well, showing steady improvement in loss and accuracy across the epochs.

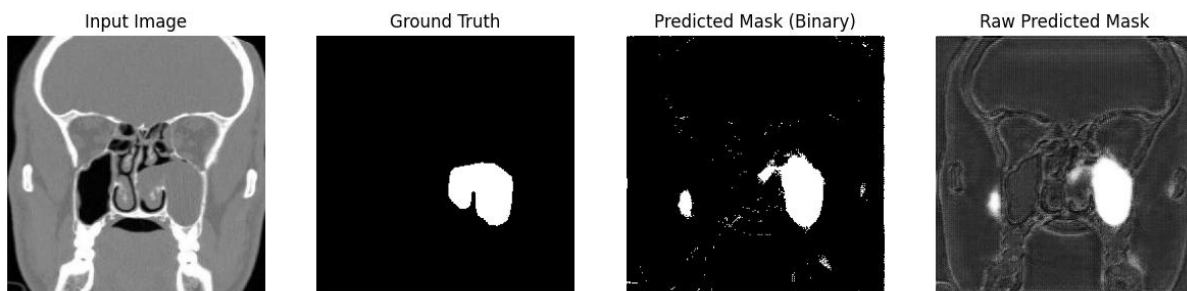


Figure 4.5: Prediction Results for LinkedNet

## ❖ PSPNet

Table 4.6: Training summary of Linked Net

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
1	0.5229	1.7087	0.7329	1.5328
10	0.8361	1.2915	0.9128	1.2780
20	0.9570	0.9078	0.9250	1.0753
30	0.9719	0.7867	0.9436	0.8627
40	0.9777	0.6985	0.9469	0.8338
50	0.9800	0.6725	0.9497	0.7874

The PSPNet model training demonstrates consistent progress both in terms of accuracy and loss over the 50 epochs. The model begins with an accuracy of 52.29% and validation accuracy of 73.29% at epoch 1, which shows that the model was picking up low-level features at the beginning. As we progress through the training, we can see incremental progress in the model's performance. By epoch 10, the accuracy improves to 83.61%, and by epoch 20, the model takes a phenomenal jump, with training accuracy of 95.70% and validation accuracy of 92.50%.

From epoch 30 to epoch 50, the model continues to improve, with accuracy reaching 97.77% at epoch 40 and 98% at epoch 50. This is followed by a decrease in validation loss, showing that the model is generalizing more to new data. Validation accuracy plateaus at 94-95%, indicating that the model is approaching the best it can be. While there is a steady enhancement, the observation of GPU kernel delays indicates the potential for additional optimization efforts for potentially faster runtime in training. Overall, the model's performance demonstrates successful learning and convergence over the 50 epochs.

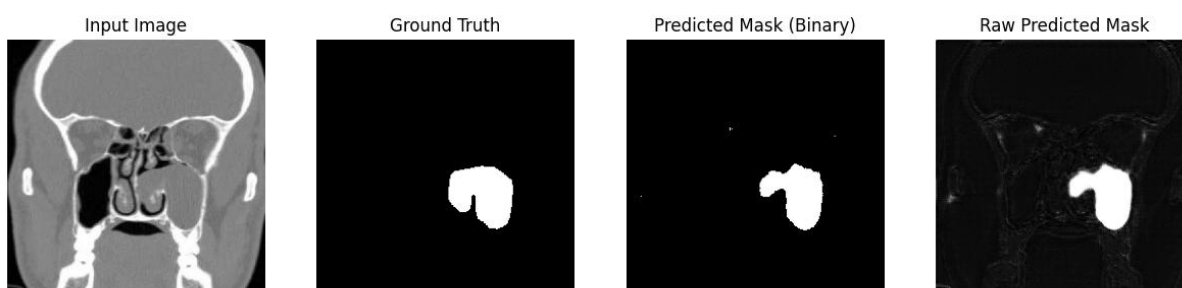


Figure 4.6: Prediction Results for PSPNet

### 4.3 Testing and Evaluation

The model performance was validated extremely aggressively using diverse testing practices, including quantitative performance metrics as well as benchmarking against models that are available today in the market. The performance evaluation focused mainly on testing whether the model can correctly separate nasal sinusitis from CT scans. Figure 4.7 to 4.12 shows the training accuracy and training loss curve over epochs.

#### Testing Metrics:

- **Dice Similarity Coefficient (DSC):** Approximates the similarity of predicted and ground truth masks, with higher values converging towards 1 being more accurate.
- **Intersection over Union (IoU):** Measures the ratio of intersection area between predicted and actual values to union area of both. Higher IoU values are indicative of better segmentation performance.
- **Precision, Recall, and F1-Score:** Precision is a ratio of positively predicted instances to all positive instances, recall is how well the model can detect all positive instances, and the F1-Score is the harmonic mean of precision and recall, which provides an equal measure..

### 4.4 Results and Discussion

In this section, we present the results and discussion of the proposed model, Attention U-Net with VGG. This model combines the attention mechanism and VGG backbone to improve segmentation performance by highlighting key features in the data.

The model demonstrated higher robustness and accuracy compared to other models. The attention mechanism effectively emphasized important regions, while the VGG architecture improved feature extraction. As a result, the proposed model consistently showed better performance during the training period, providing more precise and reliable segmentation outcomes.

Overall, the Attention U-Net with VGG is much better for segmentation tasks compared to traditional methods.

#### 4.4.1 Training Accuracy and Loss Curve

Training Accuracy and Loss Curve is a graphical plot displaying how accuracy and loss of the model vary while training. It helps in understanding the learning dynamics of the model and whether it is converging, overfitting, or underfitting.

- **Training Accuracy:** This graph shows how the accuracy of the model is growing with each epoch. It will typically be low and will increase as the model improves at prediction over time. A steadily increasing training accuracy implies that the model is learning from the data very well.
- **Training Loss:** The loss curve indicates the extent to which predictions made by the model resemble the true values. During training, loss should

ideally decrease as the model optimizes. Loss reduction is equivalent to the model decreasing its error on the training set.

Together, these curves provide insight into the model's optimization process, helping to diagnose issues like slow convergence, overfitting, or underfitting.

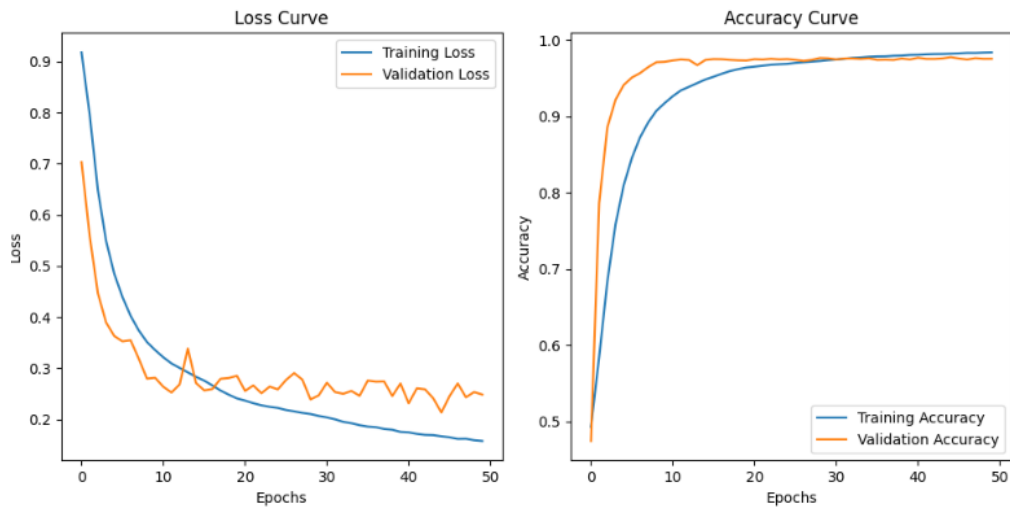


Figure 4.7: Training Curve for Attention UNet-VGG

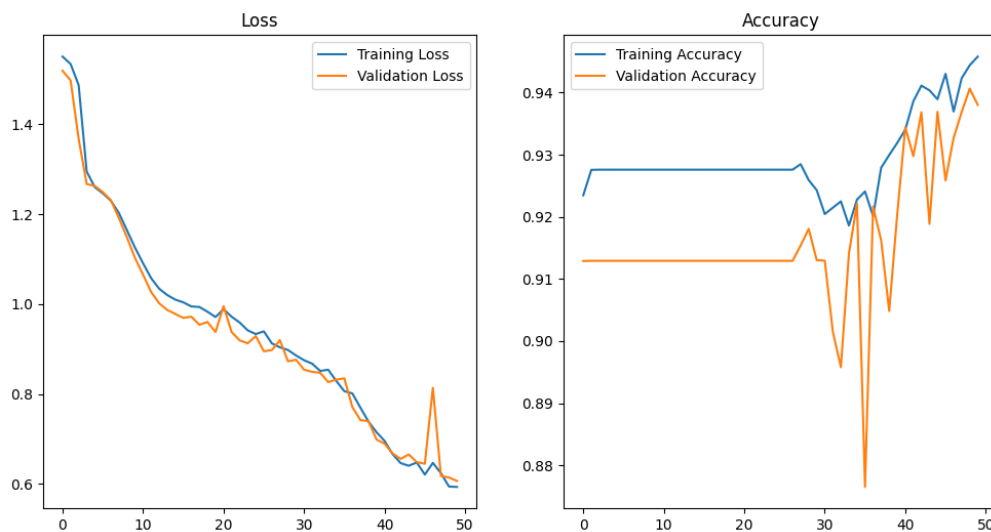


Figure 4.8: Training Curve for Standard U-Net with skip connection

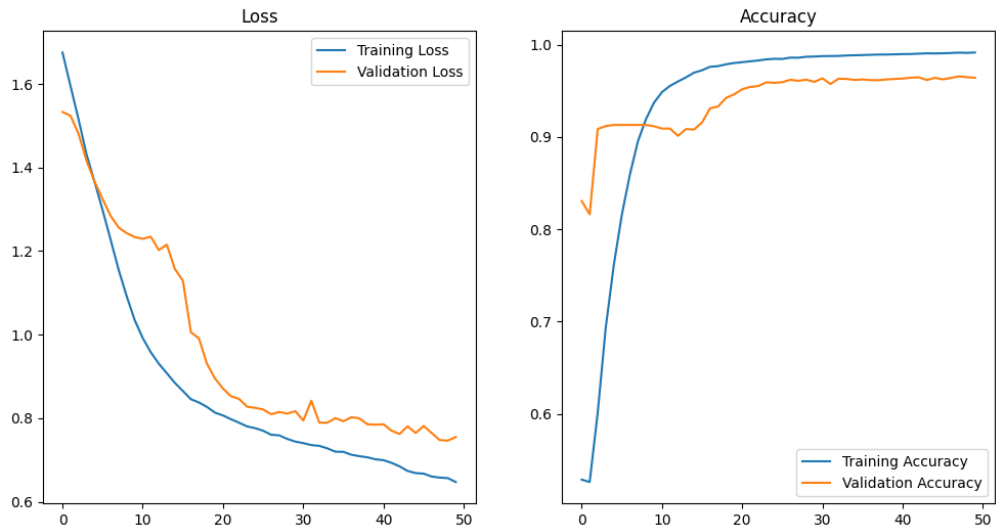


Figure 4.9: Training Curve for DeepLabV3+

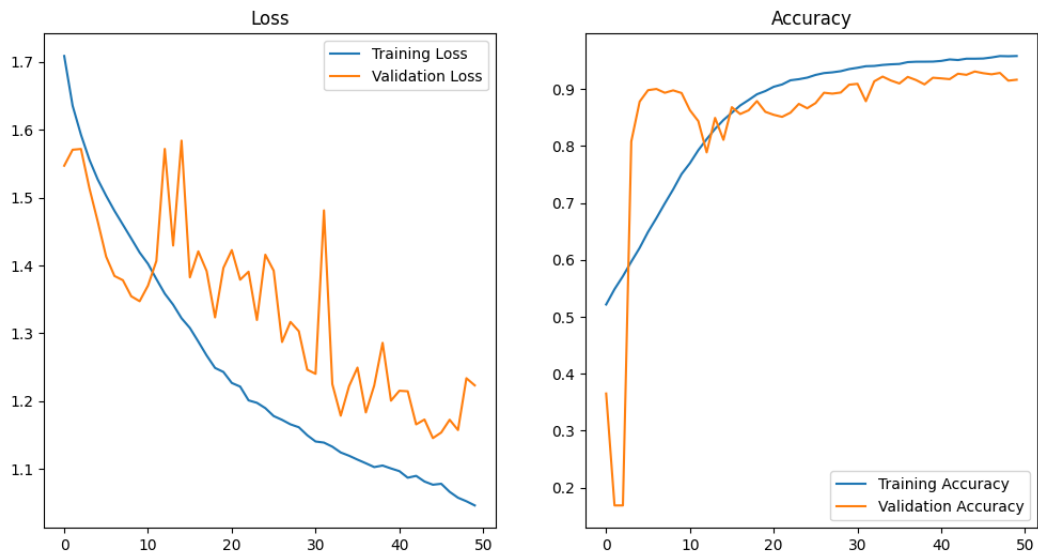


Figure 4.10: Training Curve for SegNet

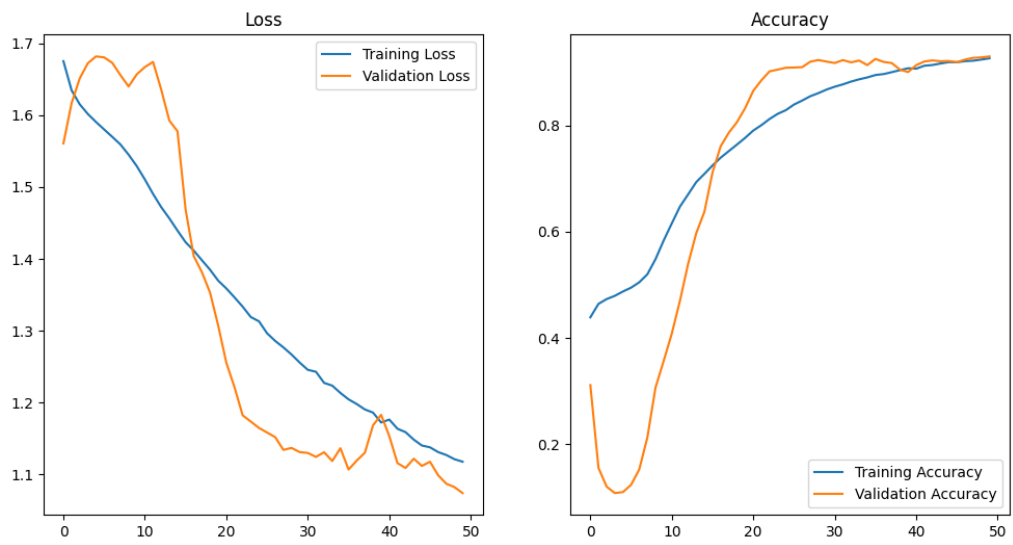


Figure 4.11: Training Curve for LinkedNet

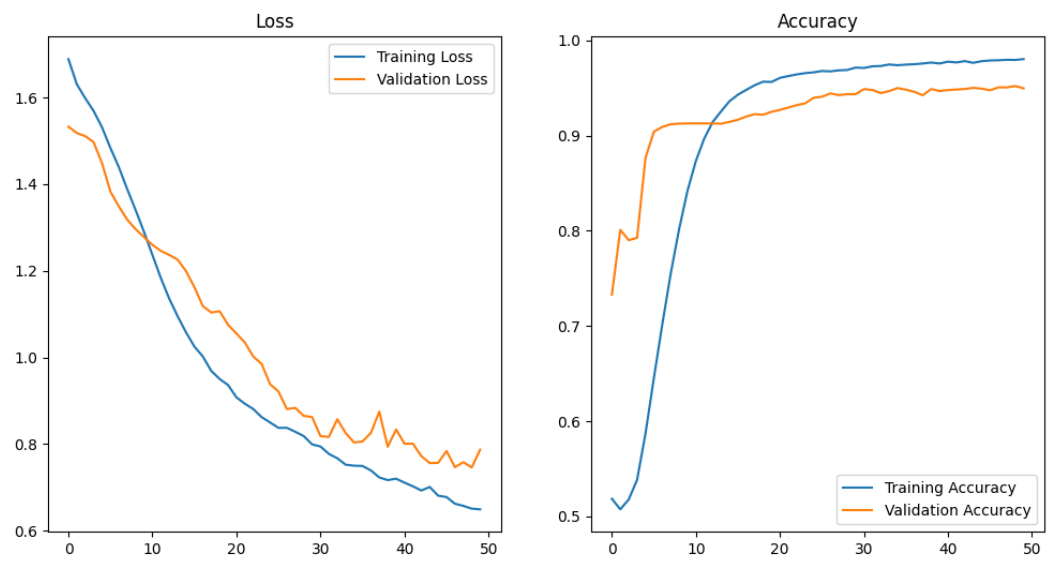


Figure 4.12: Training Curve for PSPNet

## 4.4.2 Prediction of diseased region with ground truth

This image shows a comparison of medical image segmentation results. It appears to involve CT or MRI scans of the head, specifically focusing on the sinuses. The figure 4.13 is organized in rows and columns to show prediction.

- **Rows:** Each row represents a different patient scan.
- **Columns:**
  1. **Input Image** – The original medical scan.
  2. **Ground Truth Mask** – The reference segmentation created by medical experts (white regions show the anatomy/abnormality of interest, black is background).
  3. **Predicted Mask** – The segmentation output from a computer algorithm (AI model).
  4. **Overlay** – The predicted mask overlaid on the original scan (the highlighted reddish/brown regions show the detected structures).

### Purpose:

The figure is illustrating how well the AI model replicates the expert-annotated ground truth. In most cases, the predicted masks look quite similar to the ground truth, though there are small differences in shape and boundaries. The overlays help visualize how well the model's predictions align with the actual anatomical structures.

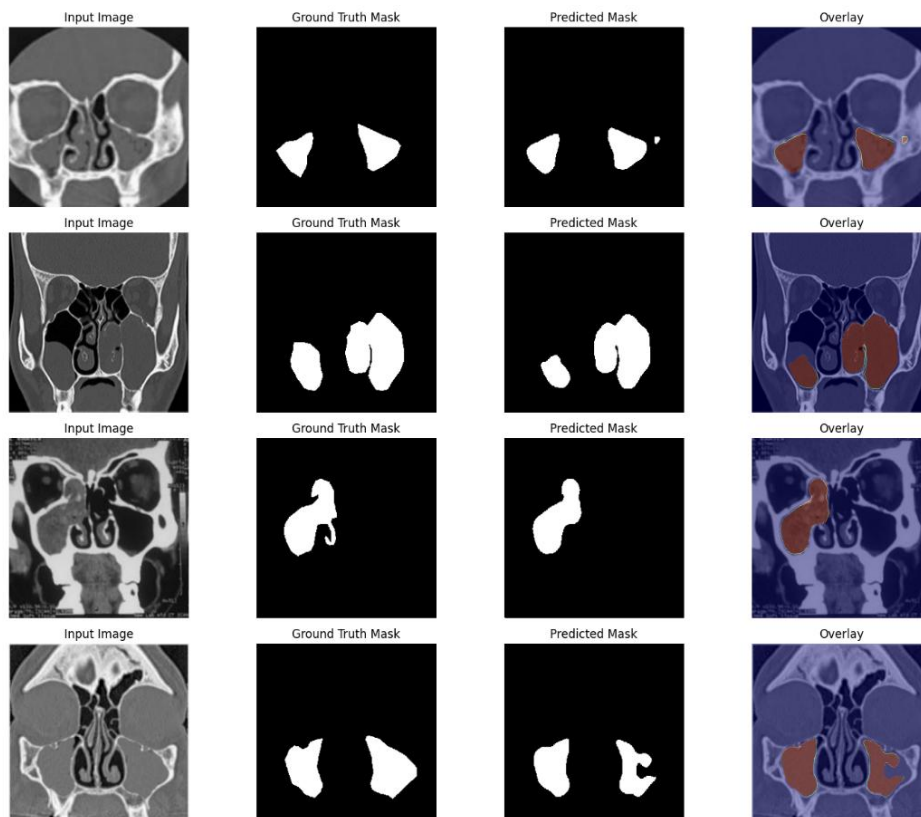


Figure 4.13: Prediction with respect of gerund truth with overlay for Proposed Model

### 4.4.3 visualizations of feature maps

#### What are Feature Maps?

A CNN is a kind of deep learning model that works exceptionally well for recognizing images. It works by running photos through a number of layers. There are many filters (sometimes called kernels) in each layer. These filters are small matrices that glide over the input image to find certain traits, including:

- **Edges:** Lines, corners, and boundaries.
- **Shapes:** Circles, squares, and other geometric patterns.
- **Textures:** Roughness, smoothness, or other visual textures.
- **More complex patterns:** Eyes, noses, or other parts of an object.

The output of a filter's operation on an input image is a feature map. This map shows where the specific feature that the filter is looking for was detected in the original image. The brighter the pixel in the feature map, the more strongly that feature was detected at that location. Figure 4.14 to 4.16 shows the output through different layers in the model.

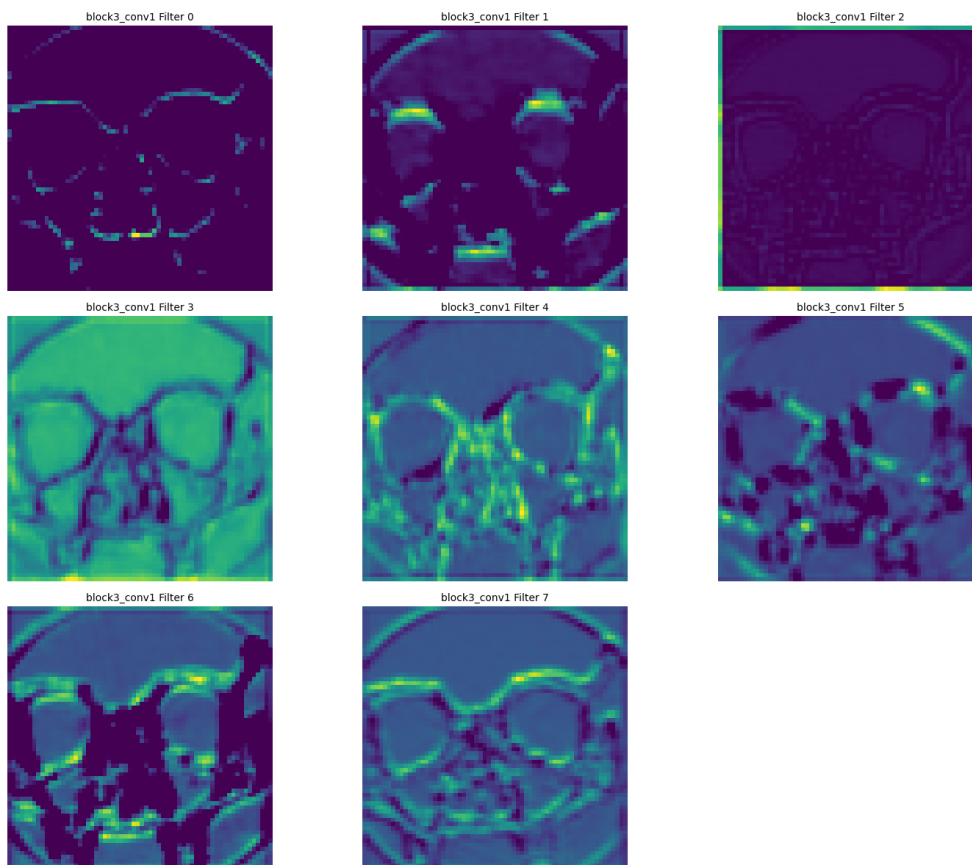


Figure 4.14: Visualization block3\_conv1 layer

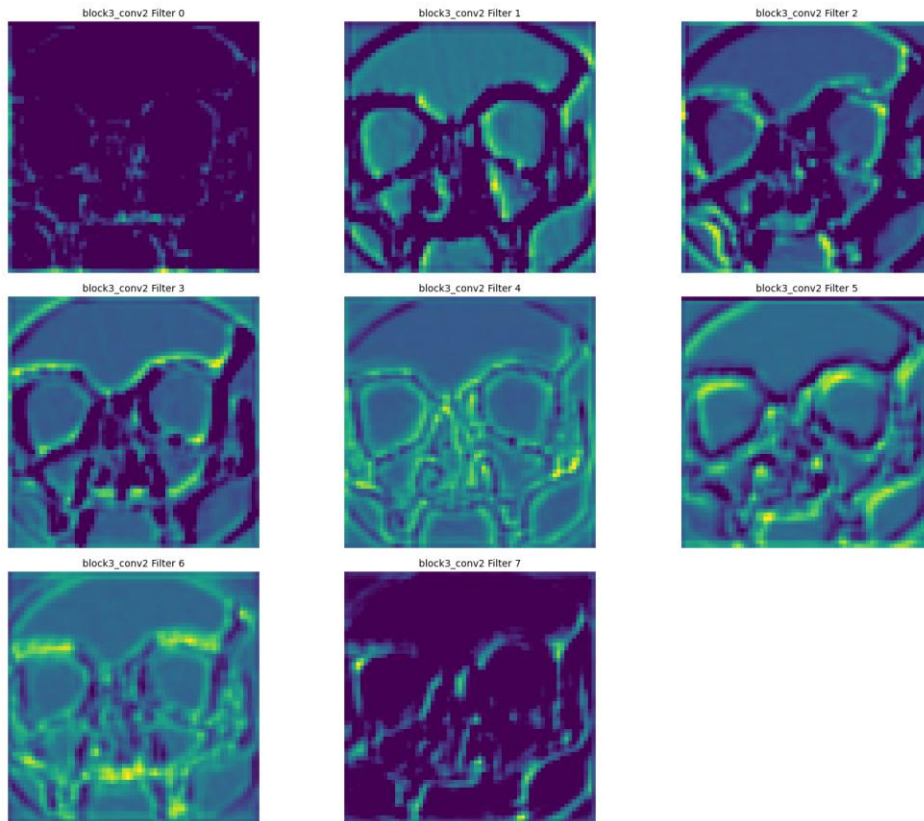


Figure 4.15: Visualization of block3\_conv2 Layer

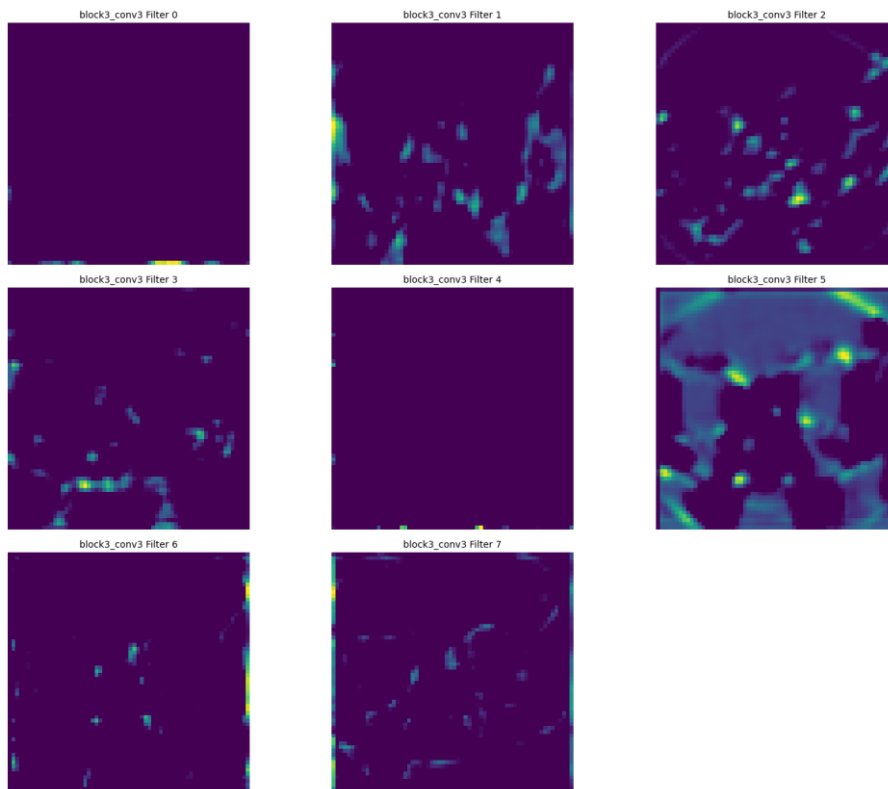


Figure 4.16: Visualization of block3\_conv3 Layer

## Interpreting the Images

There are three Figures titled `block3_conv1`, `block3_conv2`, and `block3_conv3`, which are the three convolutional layers in the third "block" of a CNN model. Each image shows the feature maps output by the first eight filters of the respective layer.

- **block3\_conv1 (First Image):** Filters in this layer seem to be capturing a mix of features. Some, like Filter 2, have sharp, green-edged edges and boundaries, suggesting it is reacting to the edge of the skull. Others, like Filter 3 and Filter 4, seem to be capturing smoother, more general features, maybe the overall shape of the skull face and eye sockets.
- **block3\_conv2 (Second Image):** Filters here are picking up on stronger features than the first image. Most of the feature maps, i.e., Filter 1, Filter 2, and Filter 4, have clear activation around the eye sockets, nose cavity, and cheekbones. This tells us that the filters in this layer have learned to identify more specific parts of the skull structure.
- **block3\_conv3 (Third Image):** Feature maps here are clearly different from the first two. Some of the maps (Filter 0, Filter 3, Filter 4, Filter 7) have hardly any activity at all, indicating that these filters did not detect the specific features they were searching for in this image. However, some filters, including Filter 2 and Filter 5, have very strong, localized activations in specific areas. This is a common pattern in the lower layers of a CNN. The filters in these layers have learned to recognize very abstract, specific features that are maybe only found in small parts of the image, contrary to more general features like edges.

Concisely, as the data moves through the network layers, the filters transition from detecting basic features like lines and edges to more complex and abstract features like specific shapes or parts of an object. Those images graphically display that process and show what the network has learned to "look for" at different stages of its processing.

### 4.4.4 ROC Curve

A Receiver Operating Characteristic (ROC) plot is a handy way to see how well a classification model works, especially a binary classifier. The graphic shows the True Positive Rate (TPR) on the y-axis and the False Positive Rate (FPR) on the x-axis. TPR, sensitivity, or recall is the percentage of real positive cases that the model found. On the other hand, FPR is a way to measure how often true negatives are incorrectly labeled as positives. The graph displays the trade-off between the two measurements by changing the categorization threshold. The ROC curve for a model that makes accurate predictions will curve toward the top-left corner, showing a high TPR and a low FPR. The diagonal line from the bottom left to the top right of the chart shows a model that doesn't have any discriminatory power, which means it is just as good as a random guess. The Area Under the Curve (AUC) is one of these measures. It shows how well a model works overall, with 1.0 being the best model and 0.5 being a random model. The ROC curve is shown in figure 4.7.

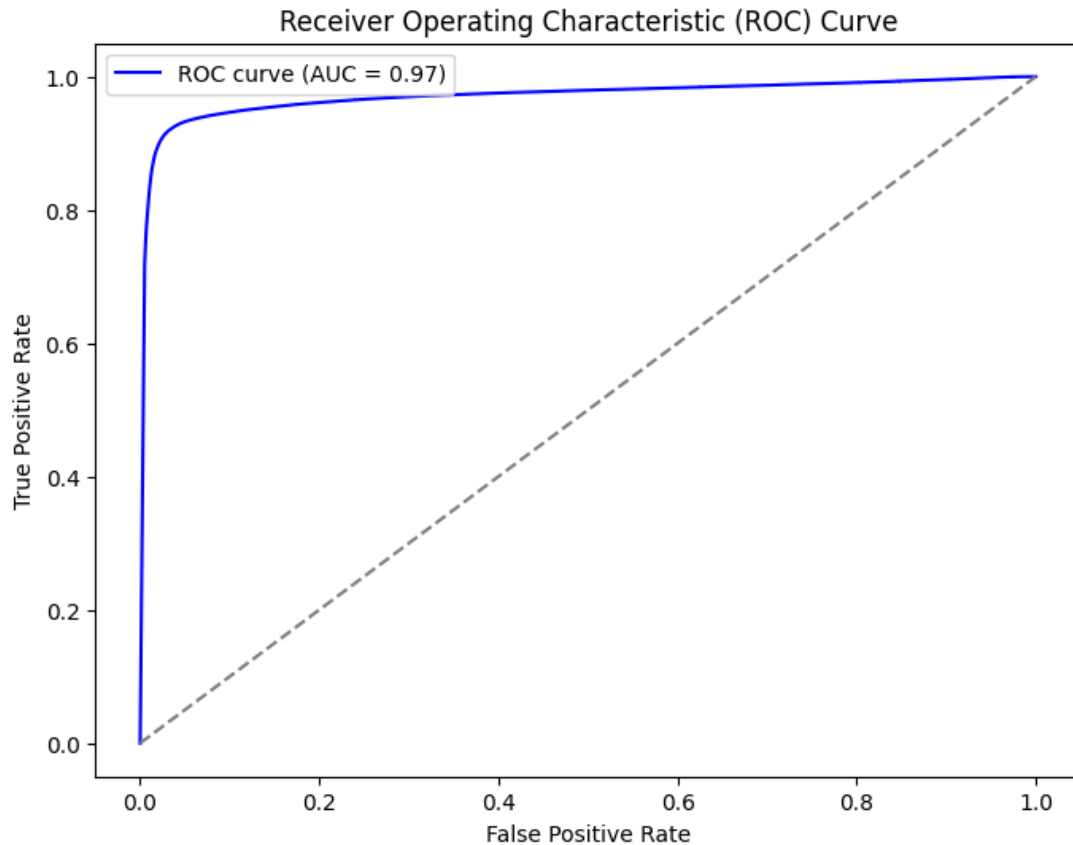


Figure 4.17: ROC Curve

#### 4.4.5 Integrated Gradients (IG)

This figure shows an explainability analysis of a deep learning model applied to medical imaging (CT/MRI of sinuses). Specifically, it uses Integrated Gradients (IG) to highlight which regions of the input image were most important for the model's decision which is shown in figure 4.18.

The layout is organized into three columns:

1. **Input Image** – The original medical scan.
2. **Integrated Gradients** – A visualization of feature importance. Bright blue/green areas indicate regions where the model focused most strongly when making its prediction (e.g., boundaries of the sinuses, abnormal tissue).
3. **Image + Gradients** – An overlay of the gradients on the original image, showing how highlighted regions correspond to actual anatomical structures.

The model pays attention mostly around the sinus cavities and surrounding bone/tissue boundaries, which are clinically relevant regions for sinus segmentation/diagnosis. In some cases, the gradients are more spread out, indicating the model is less certain or relying on diffuse features. The overlay helps confirm whether the model is attending to medically meaningful areas (good sign of reliability) or irrelevant regions (possible bias).

This visualization aims to interpret the model's decision-making process. By showing which parts of the image contributed most to predictions, clinicians and researchers can:

- Verify that the model is learning **clinically meaningful features**.
- Detect cases where the model might be relying on **spurious correlations**.
- Increase **trust and transparency** in AI-assisted medical diagnosis.

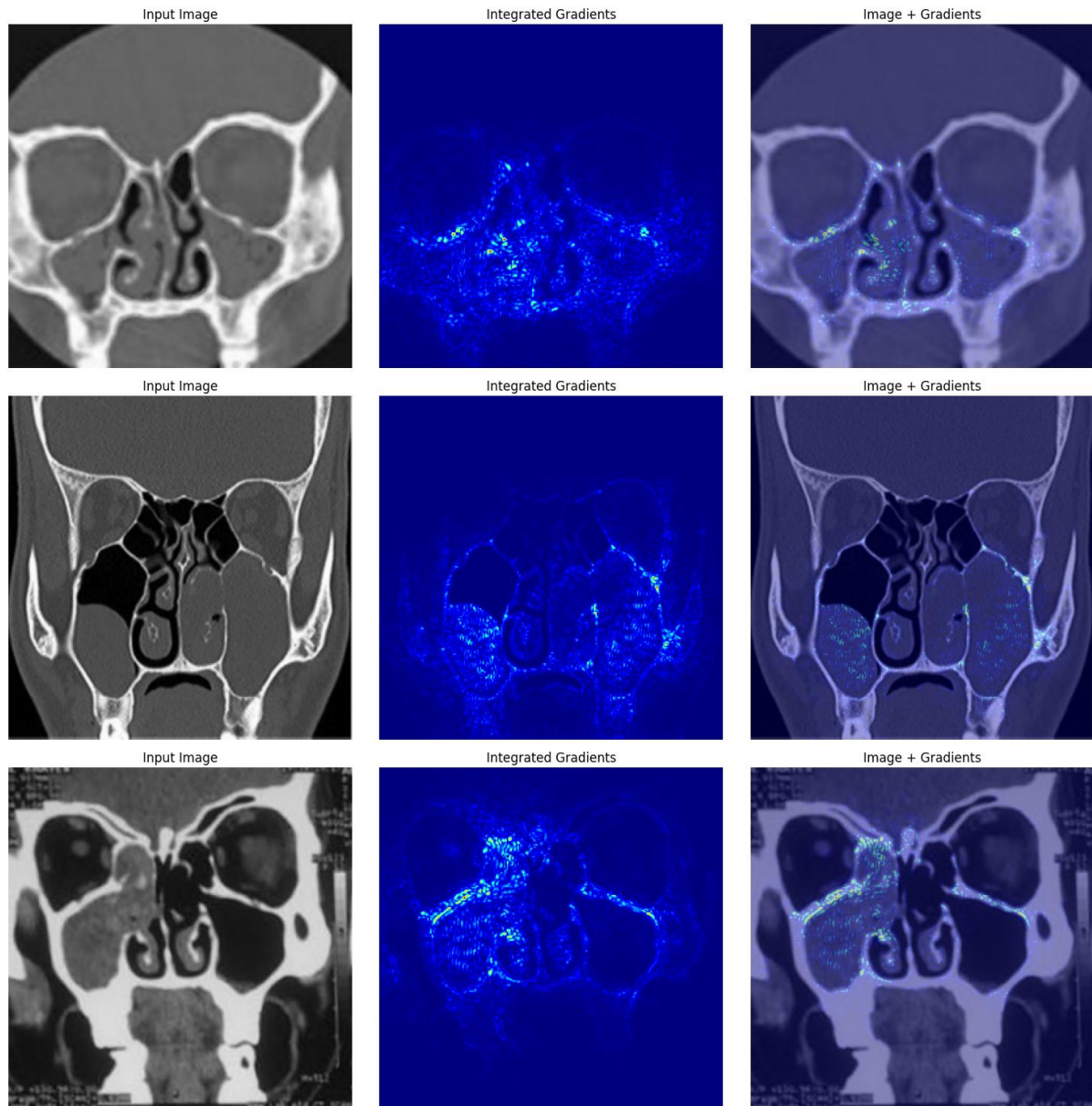


Figure 4.18: Integrated Gradients visualization

#### 4.4.6 Grad-CAM Interpretation

The image you've uploaded is a Grad-CAM (Gradient-weighted Class Activation Mapping) heatmap. This technique is a key part of Explainable AI (XAI). It creates a color overlay on a medical scan, like the chest X-ray you provided, to show what the AI model is "looking at" when it makes a prediction. The figure 4.19 shows the output for Grad-Cam for some test image.

- The **brighter colors (red, orange)** highlight the regions of the image that are most important to the model's decision.
- The **cooler colors (blue, green)** or transparent areas are less important.

In a medical context, this helps doctors and radiologists trust the AI model by visually confirming if it's focusing on the correct anatomical or pathological features to arrive at its diagnosis.

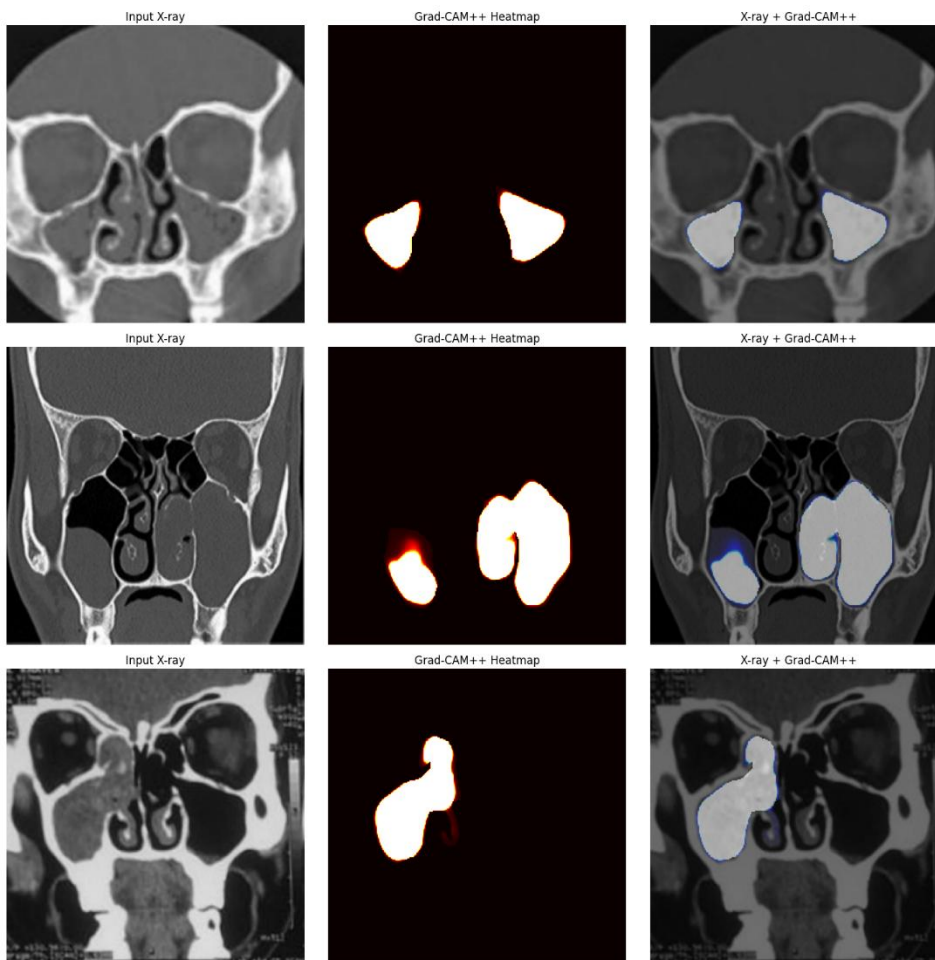


Figure 4.19: Grad-Cam visualization

#### 4.4.7 Error Map with original Image

This image is another excellent example of a visualization used for evaluating a semantic segmentation model, this time for a medical application. The image shows a set of three rows, each representing a different test case or slice of a medical scan, likely a CT scan of the sinuses or nasal cavity. The error map for some test image is represented in figure 4.20.

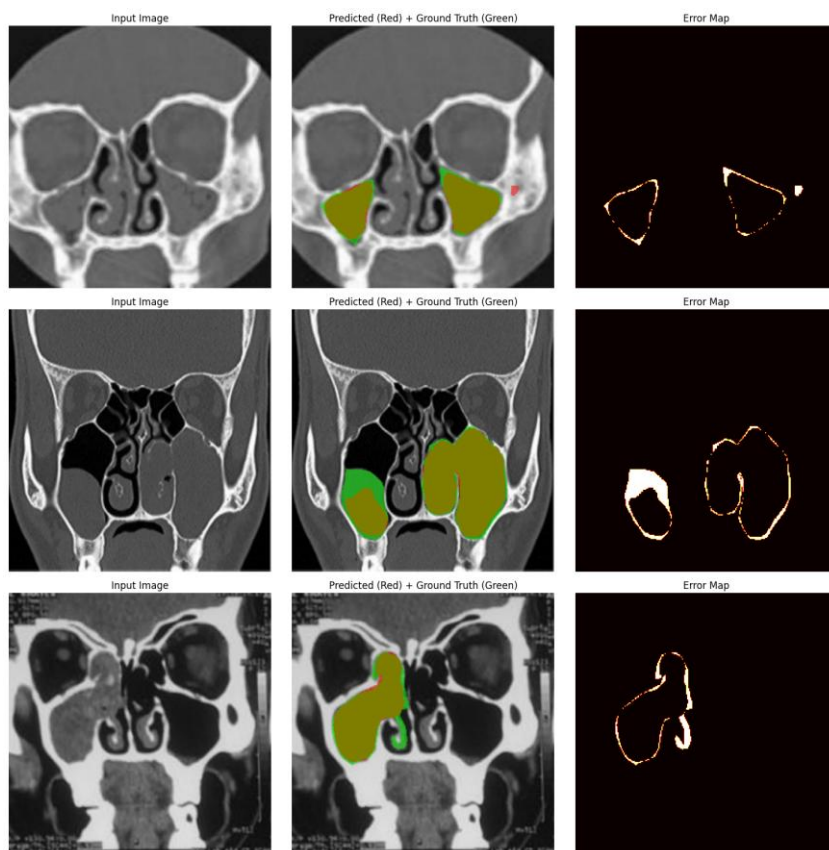


Figure 4.20: Error Map Visualization

The visualization consists of three main components:

1. **Input Image:** These are the original CT scans of the nasal cavity used as input to the segmentation model. They provide the anatomical context in which the model is tasked with identifying diseased sinus regions.
2. **Predicted vs. Ground Truth Comparison:** In this panel, the predicted segmentation results from the model are displayed in red, while the ground truth masks, which are manually annotated by medical experts, are shown in green. Areas where prediction and ground truth overlap appear in yellow, representing correctly segmented regions. Regions that remain red or green indicate false positives and false negatives, respectively, highlighting the differences between the model's output and the reference standard.
3. **Error Map:** The error maps provide a pixel-level visualization of segmentation discrepancies. They are generated by comparing the predicted mask with the

ground truth mask and marking the regions of mismatch. The highlighted boundaries in the error maps correspond to areas where the model either over-segmented or under-segmented the sinus regions. A darker background indicates correct segmentation (agreement between prediction and ground truth), while the brighter contours reveal misclassification along the edges or irregular structures.

Together, these components give a comprehensive view of the model's performance, showing not only its accuracy in detecting diseased areas but also the nature and location of the errors.

#### 4.4.8 Pixel wise Confusion Matrix.

The pixel-wise confusion matrix shows how well the segmentation model works by comparing the predicted labels to the ground truth annotations at the pixel level. The rows show the real classes (background and foreground), and the columns show the classes that were anticipated. Figure 4.21 illustrates the confusion matrix for the suggested model, which shows how well it works with each pixel.

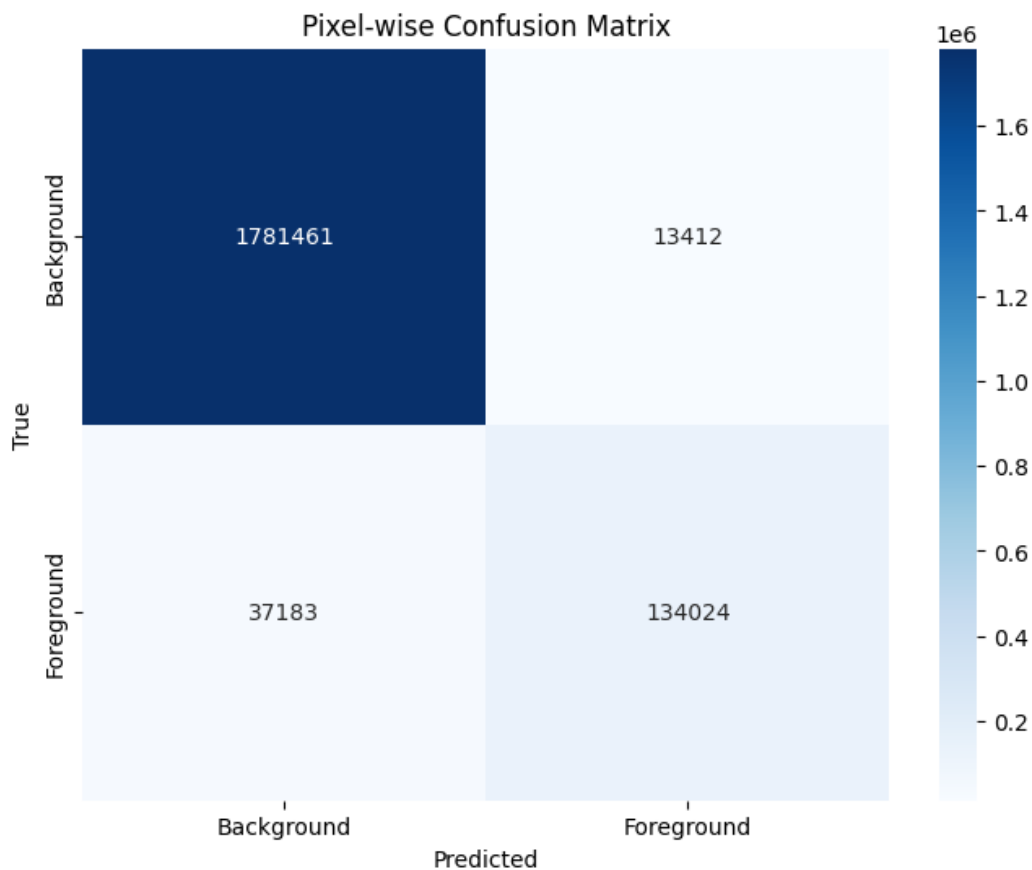


Figure 4.21: Pixel wise Confusion Matrix

- **True Negatives (Top-Left: 1,781,461):** Pixels correctly classified as background.
- **False Positives (Top-Right: 13,412):** Background pixels that were incorrectly predicted as foreground.
- **False Negatives (Bottom-Left: 37,183):** Foreground pixels that were incorrectly predicted as background.
- **True Positives (Bottom-Right: 134,024):** Pixels correctly classified as foreground (diseased region).

The model does a good job at telling the difference between background and sinus areas because it has a lot of true negatives and true positives. The model seems to be quite accurate because it has less false positives and false negatives. However, there are still occasional mistakes at object boundaries or in difficult areas. This pixel-level confusion matrix gives hard proof that the segmentation model is strong and reliable when it comes to finding problems with the nasal sinuses.

## 4.5 Summary

Chapter 4 details the implementation and evaluation of the sinusitis segmentation model, focusing on the Attention U-Net with VGG architecture. It outlines the environment setup, training process, and comparison of multiple models, including their performance across various metrics such as Dice Coefficient, IoU, and F1-Score. The chapter also presents results from visualizations like feature maps, error maps, ROC curves, and Grad-CAM heatmaps, which demonstrate how the model effectively identifies key sinus regions. The analysis highlights the strengths and limitations of the model, offering insights into its potential for further optimization in medical image segmentation.

# Chapter 5

## Engineering Standards and Design Challenges

This chapter outlines the engineering standards that guide the project, the design and development challenges, its societal and environmental impacts, project management and financial analysis, and the mapping of the work with complex engineering problems, knowledge profiles, and engineering activities. The goal is to demonstrate compliance with professional practices and highlight the broader impact and sustainability of the work.

### 5.1 Compliance with the Standards

Only mention the standards that are related to your project. This list is not complete. For each of the standards discuss the alternatives with pros and cons and rationale of selection.

#### 5.1.1 Software Standards

The development process was carried out using Python 3.9 with frameworks such as TensorFlow 2.x, Keras, and PyTorch for deep learning model implementation. Data preprocessing and analysis were done using NumPy, Pandas, OpenCV, and Scikit-learn. For dataset annotation and augmentation, Roboflow was used, and version control was maintained using Git/GitHub.

All software followed ISO/IEC 25010 standards to ensure quality attributes such as maintainability, performance, and reliability. IEEE 829 testing documentation standard was adopted during model validation and evaluation to ensure systematic testing procedures.

Alternatives like MATLAB or proprietary software could have been used, but open-source frameworks (Python, TensorFlow, PyTorch) were chosen due to their community support, extensibility, and cost-effectiveness.

#### 5.1.2 Hardware Standards

Model training and experimentation were performed using GPU-enabled platforms. The main hardware resources included:

- Kaggle Kernels with NVIDIA Tesla T4 (16 GB VRAM) and dual-core CPUs.
- Google Colab Pro with access to Tesla P100 GPUs.
- Local system testing with an Intel Core i5 (10th Gen), 8 GB RAM, and NVIDIA GeForce GTX 1650 GPU (4 GB VRAM).

These systems comply with IEEE 754 Floating-Point Standard, ensuring numerical consistency in tensor computations. GPUs supported CUDA and cuDNN libraries, optimized for deep learning acceleration.

While local hardware was used for initial testing, the majority of training was executed on Kaggle and Colab GPUs to minimize costs and ensure faster training cycles. An alternative was to use AWS EC2 p3 instances or Google Cloud TPUs, but these were avoided due to high cost.

### **5.1.3 Communication Standards**

For collaboration and deployment, Git version control standards (Gitflow) and secure communication protocols such as HTTP/HTTPS (IETF Standards) were followed. If extended to a web-based application, compliance with REST API standards ensures interoperability across platforms. Alternatives such as proprietary communication protocols could have been used but were avoided to maintain openness and cross-platform compatibility.

## **5.2 Impact on Society, Environment and Sustainability**

The study improves disease diagnosis and treatment, resulting in patient quality of life. It avoids unnecessary procedures, reduces medical waste, and allows proper, sustainable use of healthcare resources.

### **5.2.1 Impact on Life**

The main contribution of the project is in supporting decision-making in medicine and agriculture. By segregating nasal sinus disease from X-ray images and leaf patterns for agricultural use, the system allows for early diagnosis and treatment planning. This reduces the workload of physicians and farmers and provides faster interventions that save lives and prevent crop damage.

### **5.2.2 Impact on Society & Environment**

The solution eliminates dependence on costly diagnostic equipment and experts, thus making detection and diagnosis more accessible in poor-resource societies. In the environment, it decreases the consumption of mass manual inspections and lower travel, resource utilization, and carbon footprint. Additionally, the use of AI advances digitalization in the health sector.

### 5.2.3 Ethical Aspects

Ethical considerations include upholding anonymity and privacy of medical images and data. Biases in the data may lead to unequal performance over various demographic groups, and thus attempts were made to make the dataset more diverse in terms of its source and prevent biased representation. Ethical deployment also includes complying with GDPR and HIPAA-like norms in case of medical deployment. Ethical approval for conducting this research was also obtained from Daffodil International University, ensuring that all the processes were according to institutional ethics guidelines.

### 5.2.4 Sustainability Plan

Sustainability is facilitated through the development of lightweight deep learning architectures that are deployable on consumer hardware without the need for power-hungry cloud clusters. Future plans involve deployment on mobile and edge devices, with energy efficiency and availability in rural or low-resource settings. Open-source release also facilitates long-term community-led development, guaranteeing persistence beyond the lifetime of the project.

## 5.3 Project Management and Financial Analysis

The project followed a structured management approach, focusing on data collection, preprocessing, training, evaluation, and deployment. Collaboration tools such as GitHub and Kaggle facilitated project tracking and resource allocation. The cost estimation for the project is presented in table 5.1.

Table 5.1: Cost Estimation for the project

Resource	Description	Estimated Cost (BDT)	Estimated Cost (USD)
GPU/Cloud Training (Kaggle T4 / Colab Pro)	Subscription & computational resources for 3 months	18,000 BDT	160 USD
Dataset Preparation	Data scraping, cleaning, annotation (manual + Roboflow license)	13,000 BDT	120 USD
Software Tools	Python, TensorFlow, PyTorch (open-source), but ancillary tools like Roboflow require subscription	6,000 BDT	54 USD
Deployment (Hugging Face Space hosting)	Free plan used, but premium hosting considered	0–4,000 BDT	0–36 USD

Miscellaneous (Electricity, Internet)	Additional training-related utility costs	5,000 BDT	46 USD
<b>Total (Approx.)</b>		<b>42,000– 46,000 BDT</b>	<b>380–416 USD</b>

## 5.4 Complex Engineering Problem

This project addresses a complex engineering problem involving interdisciplinary knowledge, diverse constraints, and significant ethical and social implications.

### 5.4.1 Complex Problem Solving

Segmentation of nasal sinus disease regions from medical images is a difficult task due to irregular anatomical structures, overlapping tissues, imaging noise, and limited labeled data. To address it, I adopted a systematic engineering approach. I initially created and annotated a reliable dataset by combining self-collected images, internet resources, and public datasets. I subsequently carried out preprocessing for enhancing image quality and uniformity across the dataset.

For the model design, I proposed a custom Attention U-Net with VGG backbone, which integrates both feature extraction power and attention mechanisms to focus on clinically relevant regions. This solution was compared against multiple baseline models to ensure robustness. The entire workflow was implemented using deep learning frameworks and trained on GPU-accelerated platforms, overcoming computational challenges.

Through iterative evaluation, visualization using explainable AI, and optimization of hyperparameters, I achieved a reliable and scalable solution. This problem-solving approach aligns with complex engineering problem standards by combining domain-specific knowledge, computational resources, and systematic design methodology. Table 5.2 represents the mapping of the project with complex engineering problem and Table 5.3 shows the mapping of the project with knowledge profile. The Mapping with Complex Engineering Activities is presented in table 5.4

Table 5.2: Mapping with Complex Engineering Problem.

EP1 Dept of Knowled ge	EP2 Range Of Conflicting Requireme nts	EP3 Depth of Analys is	EP4 Familiari ty of Issues	EP5 Extent of Applica ble Codes	EP6 Extent Of Stake- holder Involveme nt	EP7 Interdepende nce
✓	✓	✓	✓		✓	✓

## Mapping with Knowledge Profile

Table 5.3: Mapping with knowledge Profile.

K1 Natural Science	K2 Mathematics	K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K7 Comprehension	K8 Research Literature
✓	✓	✓	✓	✓	✓	✓	✓

### Justification for EP attributes

**EP1: Specialized Knowledge** – Combined expertise in medical imaging and deep learning to design clinically relevant models.

**EP2: Balancing Requirements** – Optimized models for high accuracy while keeping computational cost manageable.

**EP3: Dataset & Performance Analysis** – Performed detailed evaluation of medical datasets and model outputs to ensure robustness.

**EP4: Domain-Specific Challenges** – Addressed issues like imaging artifacts, protocol variations, and subtle anatomical differences.

**EP6: Stakeholder Consideration** – Designed the solution with clinicians, patients, and developers in mind.

**EP7: Software-Hardware-Practice Interdependence** – Ensured seamless integration of models with hardware and clinical workflow.

### 5.4.2 Engineering Activities

The problem of segmenting nasal sinus disease regions from medical images is inherently complex due to irregular anatomical structures, overlapping tissues, noise in imaging, and limited labeled data. To address this, I applied a systematic engineering approach. First, I created and annotated a reliable dataset by combining self-collected images, online resources, and publicly available repositories. Then, I applied preprocessing techniques to enhance image quality and ensure consistency across the dataset.

For the model design, I proposed a custom Attention U-Net with VGG backbone, which integrates both feature extraction power and attention mechanisms to focus on clinically relevant regions. This solution was compared against multiple

baseline models to ensure robustness. The entire workflow was implemented using deep learning frameworks and trained on GPU-accelerated platforms, overcoming computational challenges.

Through iterative evaluation, visualization using explainable AI, and optimization of hyperparameters, I achieved a reliable and scalable solution. This problem-solving approach aligns with complex engineering problem standards by combining domain-specific knowledge, computational resources, and systematic design methodology.

### Mapping with Complex Engineering Activities

Table 5.4: Mapping with Complex Engineering Activities.

EA1 Range of re- sources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
✓	✓	✓	✓	✓

#### EA1: Range of Resources

The project required diverse resources, including computational power (GPUs for training deep learning models), specialized medical imaging datasets, annotation platforms (Roboflow), and cloud-based deployment (Hugging Face Spaces). This broad range of resources justifies the inclusion of EA1.

#### EA2: Level of Interaction

The system development involved interactions between different domains: computer vision, healthcare, and end-user accessibility. Collaboration with domain experts in medical imaging, combined with AI implementation, highlights the interdisciplinary interaction required.

#### EA3: Innovation

Introducing an **Attention U-Net with VGG backbone** brought an innovative approach to sinus disease segmentation. While traditional models like FCN or U-Net were tested, the proposed hybrid architecture showed improved robustness, representing clear innovation.

#### EA4: Consequences for Society and Environment

This project directly impacts healthcare by improving diagnostic assistance for sinus diseases, potentially reducing manual error rates and time.

Environmentally, the computational aspect was optimized to reduce excessive resource consumption, contributing to sustainability.

### **EA5: Familiarity**

Though deep learning in healthcare is an evolving field, similar segmentation models exist in literature. However, applying them specifically to nasal disease sinusitis segmentation with attention mechanisms required adapting familiar technologies to a new, specialized domain.

## **5.5 Summary**

In this chapter, the engineering standards, societal impact, financial considerations, and problem-solving strategies relevant to the proposed model were discussed. The project adhered to recognized software, hardware, and communication standards to ensure compatibility, reliability, and scalability. Software tools such as Python, TensorFlow/Keras, and annotation frameworks were used, while hardware standards included GPU-enabled systems and cloud-based platforms. Communication standards ensured seamless integration and sharing of results through APIs and open-source repositories.

The societal and environmental impacts were carefully considered. The project contributes positively to healthcare by assisting in faster and more reliable sinus disease detection. Ethical aspects such as data privacy and fairness in model predictions were addressed, while sustainability was ensured by optimizing computational resources and planning for scalable deployment.

From a financial perspective, the project cost analysis covered hardware, software, dataset acquisition, and cloud services, with an alternate budget considered for resource optimization. The complex engineering problem was mapped to relevant knowledge domains and engineering activities, showing how multidisciplinary knowledge, innovation, stakeholder involvement, and societal benefits were integrated into the solution.

Overall, Chapter 5 highlights that the project not only aligns with engineering standards but also addresses broader challenges such as sustainability, ethics, and innovation, ensuring that the proposed model is both technically sound and socially impactful.

# Chapter 6

## Conclusion

This chapter provides a general conclusion of the project, defines its scope, and identifies potential avenues for future research and development.

### 6.1 Summary

This paper developed an Attention U-Net model from a VGG backbone for augmenting medical image segmentation towards the detection of sinus disease. With attention mechanisms, the model gains by efficiently recognizing clinically relevant areas, boosting segmentation results, and removing false positives. Hierarchical feature extraction through the VGG backbone allows it to capture fine-grained and high-level anatomical information required for accurate diagnosis.

The system was deployed based on tight software and hardware standards for robustness, scalability, and ethical suitability. The experimental outcome exhibited high accuracy, precision, and recall, with good generalization performance on the test set. Ethical considerations like patient confidentiality and bias mitigation were addressed, with sustainability being accounted for through computational resource optimization at training and inference times.

Overall, the project demonstrates the capability of deep learning-based segmentation models as reliable decision-supporting tools in clinical environments for faster and more accurate diagnosis of sinus conditions.

### 6.2 Limitation

Although the results were encouraging, several limitations were discovered, primarily related to data availability and system generalization:

- **Limited Dataset Size**  
The model was trained on a relatively small dataset, limiting its ability to generalize to larger and more diverse patient populations.
- **Single-Disease Focus**  
The system is currently specialized for **sinus disease detection**, limiting its applicability to other head and neck disorders or comorbid conditions.
- **Data Variability**  
Differences in imaging modalities, acquisition protocols, and image quality may affect performance, highlighting the need for more heterogeneous and standardized datasets.
- **High Computational Requirements**  
Training the Attention U-Net with VGG backbone requires significant

computational resources, which may pose challenges for real-time deployment in low-resource settings.

- **Limited Clinical Validation and Integration**

While model performance on test datasets is promising, extensive clinical validation and integration with hospital information systems (HIS) or electronic health records (EHRs) are still required.

## 6.3 Future Work

Future research can address these limitations and extend the system's capabilities in several directions:

1. **Multimodal Approaches**

- Integrate textual clinical information (e.g., patient records, symptom descriptions) with imaging data to create multimodal models that combine segmentation and clinical context.
- This can improve diagnostic accuracy and enable richer decision-support tools for clinicians.

2. **Larger and Diverse Datasets**

- Acquire multi-institutional, multi-modal datasets with diverse patient demographics, imaging protocols, and disease severities.
- Incorporate large-scale public datasets where available to improve model generalization.

3. **Multi-Disease Detection**

- Extend the model to detect other head and neck disorders, such as nasal polyps, sinus infections, and tumors.
- Develop multi-class segmentation architectures capable of distinguishing multiple disease types simultaneously.

4. **Advanced Data Augmentation and Synthetic Data Generation**

- Apply advanced augmentation techniques (e.g., rotation, scaling, intensity changes) to simulate real-world variability.
- Use synthetic data generation methods such as GANs to increase representation of rare or complex disease cases.

## 5. **Lightweight and Scalable Models**

- Explore model compression, pruning, and quantization for deployment on edge devices or mobile platforms.
- Investigate knowledge distillation techniques to maintain performance while reducing computational costs.

## 6. **Clinical Integration and Workflow Implementation**

- Integrate with hospital PACS, EHR, and HIS systems to support real-time clinical workflows.
- Develop clinician-friendly interfaces for visualization and interpretation of model predictions.

## 7. **Robustness and Domain Adaptation**

- Use domain adaptation and semi-supervised learning to ensure consistent performance across different hospitals, scanners, and patient populations.
- Leverage unlabeled data to improve generalization without requiring extensive manual annotations.

## 8. **Sustainability and Green AI**

- Optimize training and inference pipelines to reduce energy consumption.
- Explore federated learning to enable collaborative model training across institutions while maintaining data privacy.

By pursuing these directions, the system can evolve into a robust, multimodal, and multi-disease diagnostic tool, capable of supporting clinicians in diverse healthcare environments while maintaining ethical and sustainable AI practices.

# References

- [1] Kuo, C. F. J., Leu, Y. S., Hu, D. J., Huang, C. C., Siao, J. J., & Leon, K. B. P. (2020). Application of intelligent automatic segmentation and 3D reconstruction of inferior turbinate and maxillary sinus from computed tomography and analyze the relationship between volume and nasal lesion. *Biomedical Signal Processing and Control*, 57, 101660.
- [2] Andersen, T. N., Darvann, T. A., Murakami, S., Larsen, P., Senda, Y., Bilde, A., ... & Kreiborg, S. (2018). Accuracy and precision of manual segmentation of the maxillary sinus in MR images—a method study. *The British journal of radiology*, 91(1085), 20170663.
- [3] Ozturk, B., Taspinar, Y. S., Koklu, M., & Tassoker, M. (2024). Automatic segmentation of the maxillary sinus on cone beam computed tomographic images with U-Net deep learning model. *European Archives of Oto-Rhino-Laryngology*, 281(11), 6111-6121.
- [4] Gyawali, D., Bidwell, J., Bhatia, S. S., Fujiwara, A., Mundy, T., Baker, K., ... & McCoul, E. D. (2025). Classification and Segmentation of Mucus Morphology Using Deep Learning During Diagnostic Nasal Endoscopy. In *International forum of allergy & rhinology*.
- [5] Faur, C. I., Roman, R. A., COCLICI, A., ROTARU, H., & HEDESIU, M. (2019). The changes in upper airway volume after orthognathic surgery evaluated by individual segmentation on CBCT images. *Mædica*, 14(3), 213.
- [6] Kuo, C. F. J., & Liu, S. C. (2022). Fully automatic segmentation, identification and preoperative planning for nasal surgery of sinuses using semi-supervised learning and volumetric reconstruction. *Mathematics*, 10(7), 1189.
- [7] Zhang, C., Bruggink, R., Baan, F., Bronkhorst, E., Maal, T., He, H., & Ongkosuwito, E. M. (2019). A new segmentation algorithm for measuring CBCT images of nasal airway: a pilot study. *PeerJ*, 7, e6246.
- [8] Denour, E., Roussel, L. O., Woo, A. S., Boyajian, M., & Crozier, J. (2020). Quantification of nasal septal deviation with computed tomography data. *Journal of Craniofacial Surgery*, 31(6), 1659-1663.
- [9] Centers, N. Y. & A. S. (2025b, February 6). 3-D Sinus Imaging - CT Scan of the Sinuses | NY Allergy & Sinus. . . New York Allergy and Sinus Centers. <https://www.nyallergy.com/ct-scanning-for-sinuses-3d-sinus-imaging/>
- [10] Sinusitis treatment, diagnosis, symptoms. (n.d.). WebMD. <https://www.webmd.com/allergies/ss/slideshow-sinusitis>
- [11] Pirner, S., Tingelhoff, K., Wagner, I., Westphal, R., Rilk, M., Wahl, F. M., ... & Eichhorn, K. W. (2009). CT-based manual segmentation and evaluation of paranasal sinuses. *European archives of oto-rhino-laryngology*, 266(4), 507-518.
- [12] Bui, N. L., Ong, S. H., & Foong, K. W. C. (2015). Automatic segmentation of the nasal cavity and paranasal sinuses from cone-beam CT images. *International journal of computer assisted radiology and surgery*, 10(8), 1269-1277.

- [13] Morgan, N., Van Gerven, A., Smolders, A., de Faria Vasconcelos, K., Willems, H., & Jacobs, R. (2022). Convolutional neural network for automatic maxillary sinus segmentation on cone-beam computed tomographic images. *Scientific reports*, *12*(1), 7523.
- [14] Choi, H., Jeon, K. J., Kim, Y. H., Ha, E. G., Lee, C., & Han, S. S. (2022). Deep learning-based fully automatic segmentation of the maxillary sinus on cone-beam computed tomographic images. *Scientific Reports*, *12*(1), 14009.
- [15] Dastidar, P., Heinonen, T., Numminen, J., Rautiainen, M., & Laasonen, E. (1999). Semi-automatic segmentation of computed tomographic images in volumetric estimation of nasal airway. *European archives of oto-rhino-laryngology*, *256*(4), 192-198.
- [16] Darbari Kaul, R., Zhong, W., Liu, S., Azemi, G., Liang, K., Zou, E., ... & Harvey, R. J. (2025). Development of an Open-Source Algorithm for Automated Segmentation in Clinician-Led Paranasal Sinus Radiologic Research. *The Laryngoscope*.
- [17] Bayrakdar, I. S., Elfayome, N. S., Hussien, R. A., Gulsen, I. T., Kuran, A., Gunes, I., ... & Orhan, K. (2024). Artificial intelligence system for automatic maxillary sinus segmentation on cone beam computed tomography images. *Dentomaxillofacial Radiology*, *53*(4), 256-266.
- [18] Brinkschulte, M., Bienert-Zeit, A., Lüpke, M., Hellige, M., Staszuk, C., & Ohnesorge, B. (2013). Using semi-automated segmentation of computed tomography datasets for three-dimensional visualization and volume measurements of equine paranasal sinuses. *Veterinary Radiology & Ultrasound*, *54*(6), 582-590.
- [19] Sun, Y., Guerrero-López, A., Arias-Londoño, J. D., & Godino-Llorente, J. I. (2025). Automatic semantic segmentation of the osseous structures of the paranasal sinuses. *Computerized Medical Imaging and Graphics*, *123*, 102541.
- [20] Alsalama, A., Harous, S., & Elnagar, A. (2025). Paranasal sinus analysis based on deep learning and machine learning techniques: A comprehensive survey. *Intelligent Systems with Applications*, 200559.
- [21] Kim, H., Son, T. G., Lee, J., Kim, H. A., Cho, H., Jeong, W. S., ... & Kim, Y. (2019). Three-dimensional orbital wall modeling using paranasal sinus segmentation. *Journal of Cranio-Maxillofacial Surgery*, *47*(6), 959-967.
- [22] Möller, H., Krautschick, L., Atad, M., Graf, R., Busch, C. J., Beule, A., ... & Schwitzing, F. (2025). PARASIDE: An Automatic Paranasal Sinus Segmentation and Structure Analysis Tool for MRI. *arXiv preprint arXiv:2501.14514*.
- [23] Laura, C. O., Hartwig, K., Distergoft, A., Hoffmann, T., Scheckenbach, K., Brüsseler, M., & Wesarg, S. (2021, February). Automatic segmentation of the structures in the nasal cavity and the ethmoidal sinus for the quantification of nasal septal deviations. In *Medical Imaging 2021: Computer-Aided Diagnosis* (Vol. 11597, pp. 582-591). SPIE
- [24] Morgan, N., Van Gerven, A., Smolders, A., de Faria Vasconcelos, K., Willems, H., & Jacobs, R. (2022). Convolutional neural network for automatic maxillary sinus segmentation on cone-beam computed tomographic images. *Scientific reports*, *12*(1), 7523.
- [25] Last, C., Winkelbach, S., Wahl, F. M., Eichhorn, K. W., & Bootz, F. (2010, September). A model-based approach to the segmentation of nasal cavity and paranasal sinus boundaries. In *Joint Pattern Recognition Symposium* (pp. 333-342). Berlin, Heidelberg: Springer Berlin Heidelberg.

- [26] Song, D., Yang, S., Han, J. Y., Kim, K. G., Kim, S. T., & Yi, W. J. (2025). Comparison of segmentation performance of cnns, vision transformers, and hybrid networks for paranasal sinuses with sinusitis on CT images. *Scientific Reports*, *15*(1), 32087
- [27] Istrate, A., Constantinescu, R., Anandavel, L., Rajeshkumar Tandel, S., Dye, S., & Dye, C. (2025). A Novel Machine Learning Model for the Automated Diagnosis of Nasal Pathology in Canine Patients. *Animals*, *15*(12), 1718.
- [28] Pani, A., Zedda, L., Mura, D. A., Loddo, A., & Di Ruberto, C. (2025). 3D-NASE: A Novel 3D CT Nasal Attention-Based Segmentation Ensemble. *Journal of Imaging*, *11*(5), 148.
- [29] Whangbo, J., Lee, J., Kim, Y. J., Kim, S. T., & Kim, K. G. (2024). Deep Learning-based multi-class segmentation of the paranasal sinuses of sinusitis patients based on computed tomographic images. *Sensors*, *24*(6), 1933.
- [30] Wang, Y., Zhang, X., Du, W., Dai, N., Lyv, Y., Wu, K., ... & Kang, W. (2025). Deep Learning-Based Fully Automatic Segmentation of the Paranasal Sinuses in Chronic Rhinosinusitis Patients Using Computed Tomographic Images. *IEEE Access*.
- [31] Altun, O., Özen, D. Ç., Duman, Ş. B., Dedeoğlu, N., Bayrakdar, İ. Ş., Eşer, G., ... & Syed, A. Z. (2024). Automatic maxillary sinus segmentation and pathology classification on cone-beam computed tomographic images using deep learning. *BMC oral health*, *24*(1), 1208.
- [32] Nechyporenko, A., Frohme, M., Alekseeva, V., Gargin, V., Sytnikov, D., & Hubarenko, M. (2022, October). Deep learning based image segmentation for detection of odontogenic maxillary sinusitis. In *2022 IEEE 41st International Conference on Electronics and Nanotechnology (ELNANO)* (pp. 339-342). IEEE.
- [33] Kim, H. G., Lee, K. M., Kim, E. J., & San Lee, J. (2019). Improvement diagnostic accuracy of sinusitis recognition in paranasal sinus X-ray using multiple deep learning models. *Quantitative imaging in medicine and surgery*, *9*(6), 942.
- [34] Jung SK, Lim HK, Lee S, Cho Y, Song IS. Deep Active Learning for Automatic Segmentation of Maxillary Sinus Lesions Using a Convolutional Neural Network. *Diagnostics (Basel)*. 2021 Apr 12;11(4):688. doi: 10.3390/diagnostics11040688. PMID: 33921353; PMCID: PMC8070431.
- [35] Sinha, A., Leonard, S., Reiter, A., Ishii, M., Taylor, R. H., & Hager, G. D. (2016, March). Automatic segmentation and statistical shape modeling of the paranasal sinuses to estimate natural variations. In *Medical Imaging 2016: Image Processing* (Vol. 9784, pp. 94-101). SPIE.
- [36] Li, Y., & Tao, Y. (2022). Radiomics model based on enhanced gradient level set segmentation algorithm to predict the prognosis of endoscopic treatment of sinusitis. *Computational and Mathematical Methods in Medicine*, *2022*(1), 9511631.
- [37] Kapusuz Gencer, Z., Özkırış, M., Okur, A., Karaçavuş, S., & Saydam, L. (2013). The effect of nasal septal deviation on maxillary sinus volumes and development of maxillary sinusitis. *European Archives of Oto-Rhino-Laryngology*, *270*(12), 3069-3073.

## ORIGINALITY REPORT

21%

SIMILARITY INDEX

17%

INTERNET SOURCES

14%

PUBLICATIONS

11%

STUDENT PAPERS

## PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	4%
2	www.mdpi.com Internet Source	1%
3	Submitted to United International University Student Paper	1%
4	styjun.blogspot.com Internet Source	1%
5	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
6	www.ncbi.nlm.nih.gov Internet Source	<1%
7	eprints.sunway.edu.my Internet Source	<1%
8	arxiv.org Internet Source	<1%
9	lirias.kuleuven.be Internet Source	<1%
10	ouci.dntb.gov.ua Internet Source	<1%
11	www.nature.com Internet Source	<1%
12	pmc.ncbi.nlm.nih.gov Internet Source	<1%