

# **ASSESSING CNN ROBUSTNESS IN MEDICAL IMAGING SYSTEMS: ADVERSARIAL THREATS AND DEFENSIVE MEASURES**

D.S.C Wijesuriya

B.Sc. (Hons) Degree in Information Technology specialized in Cybersecurity

Department of Computer Systems and Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

August 2024

# **ASSESSING CNN RESILIENCE TO PENALIZING GRADIENT NORM (PGN) ADVERSARIAL ATTACKS**

Project Proposal Report

Shamal Chathuranga Wijesuriya

IT21155802

B.Sc. (Hons) Degree in Information Technology Specializing in Cyber Security


Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

August 2024

## DECLARATION OF THE CANDIDATE & SUPERVISOR

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

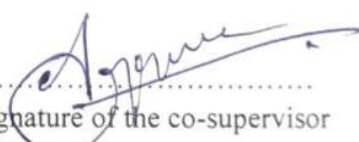
Student Name	Registration Number	Signature
D.S.C. Wijesuriya	IT21155802	

The supervisor/s should certify the proposal report with the following declaration. The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

  
.....  
Signature of the supervisor

(Dr. Harinda Fernando)

22/8/24  
.....  
Date

  
.....  
Signature of the co-supervisor

(Mr. Kavinga Yapa)

22/8/24  
.....  
Date

## Abstract

This research investigates the resilience of Convolutional Neural Networks (CNNs) against Penalizing Gradient Norm (PGN) adversarial attacks in medical imaging, specifically chest X-ray classification. PGN attacks subtly alter input data, leading to significant prediction errors. The study aims to address gaps in existing research by evaluating the impact of these attacks, testing defense strategies like data augmentation and adversarial training, and validating their effectiveness using performance metrics such as accuracy, precision, recall, and F1-score. Through a systematic methodology, the goal is to enhance the robustness of CNN models, ensuring reliable and secure medical diagnostics.

**Keywords:** *Convolutional Neural Networks (CNNs), Medical Imaging, Adversarial Attacks, Penalizing Gradient Norm (PGN), Chest X-Ray Classification*

# TABLE OF CONTENTS

Abstract .....	iv
TABLE OF CONTENTS .....	v
TABLE OF FIGURES .....	vi
LIST OF TABLES .....	vii
LIST OF ABBREVIATIONS .....	viii
1. INTRODUCTION .....	1
1.1 Background .....	1
1.2 Literature Survey .....	2
1.3 Research Gap .....	3
1.4 Research Problem .....	4
2. OBJECTIVES .....	5
2.1 Main Objective.....	5
2.2 Specific Objectives .....	5
3. Methodology .....	7
3.1 Overall System Diagram.....	7
3.2 Component System Diagram .....	9
3.3 Tools & Technologies.....	11
3.4 Requirements .....	12
3.4.1 System requirements.....	13
3.4.2 Personal requirements.....	13
3.4.3 Software requirements .....	14
3.5 Gantt Chart.....	15
3.6 WBS.....	16
4. COMMERCIALIZATION AND BUDGET .....	17
4.1 Commercialization.....	17
4.2 Budget .....	18
5. REFERENCES .....	19

## TABLE OF FIGURES

Figure 1: Overall System Diagram .....	9
Figure 2: PGN Attack System Diagram.....	11
Figure 3: Gantt Chart .....	15
Figure 4: Work Breakdown Structure.....	16

## LIST OF TABLES

Table 1: Research Gap .....	3
Table 2: Budget Allocation Table .....	18

## LIST OF ABBREVIATIONS

Abbreviation	Description
CNN	Convolutional Neural Network
PGN	Penalizing Gradient Norm
SIM	Split Image Adversarial
USSM	Uniform Scale Mix Mask
STM	Style Transfer Manipulation
IDE	Integrated Development Environment



# 1. INTRODUCTION

## 1.1 Background

Medical imaging systems, particularly those utilized for chest X-ray classification, are critical in modern healthcare for diagnosing and monitoring various conditions, including pneumonia, tuberculosis, and lung cancer. These systems rely heavily on Convolutional Neural Networks to accurately interpret medical images, providing vital support to healthcare professionals in making informed decisions. However, the integrity and reliability of these systems are increasingly challenged by adversarial attacks, which exploit vulnerabilities in machine learning models to cause misclassifications.

One such adversarial technique is the PGN attack, which subtly alters input images in ways that are almost imperceptible to the human eye but can lead to significant errors in CNN predictions. The potential consequences of these errors in a medical setting are severe, as they can result in incorrect diagnoses and inappropriate treatments, ultimately compromising patient safety.

Despite the growing use of CNNs in medical imaging, there is a noticeable gap in the literature regarding the resilience of these models to PGN attacks. While considerable research has been conducted on other types of adversarial attacks, the specific impact of PGN attacks on CNN performance, particularly within the context of medical imaging, remains underexplored. Furthermore, existing defense strategies have not been thoroughly tested against PGN attacks, leaving questions about their effectiveness and reliability.

This research aims to close these gaps by evaluating the strength of CNN models against PGN adversarial attacks and confirming the effectiveness of current defense mechanisms. This study will help improve the security and dependability of medical imaging systems, ensuring that they can still offer accurate and reliable diagnostic support even in the presence of advanced adversarial threats.

## 1.2 Literature Survey

The application of Convolutional Neural Networks in medical imaging has been transformative, providing enhanced diagnostic accuracy and efficiency. CNN models have demonstrated their effectiveness in various medical imaging tasks, such as the detection of COVID-19 from chest X-rays, where they have achieved significant success in classifying complex visual patterns in medical data [1]. Despite these advancements, CNNs are vulnerable to adversarial attacks, which involve subtle perturbations to input data that can lead to significant misclassifications [2].

Adversarial attacks on CNNs, such as the PGN attack, present a significant challenge in ensuring the robustness of these models. PGN attacks work by manipulating the gradient norms during the model's training process, making the model highly sensitive to small input changes. This results in the model misclassifying perturbed images, which poses serious risks in critical applications like medical imaging [3]. Recent studies have explored various methods to enhance the transferability and effectiveness of adversarial attacks, including those that focus on finding flat local maxima in the model's loss landscape [4].

In the context of medical imaging, the impact of PGN attacks has been less explored, particularly in terms of their effects on CNN models trained with medical datasets. Previous research has largely focused on the general aspects of adversarial attacks without delving deeply into their implications for medical diagnostics [5]. Additionally, there is a notable gap in the literature regarding the testing and validation of existing defense mechanisms against PGN attacks within medical imaging contexts. While some defense strategies, such as adversarial training and data augmentation, have been proposed, their effectiveness against PGN attacks specifically in medical imaging remains under-researched [6].

This literature survey highlights the need for comprehensive research that addresses these gaps. Specifically, it underscores the importance of evaluating the resilience of CNN models against PGN attacks in medical imaging, testing existing defense strategies, and exploring new methods to enhance model robustness. By filling these gaps, the research aims to contribute to the development of more secure and reliable diagnostic tools in healthcare.

### 1.3 Research Gap

Despite previous research on PGN adversarial attacks, significant gaps remain, particularly in the context of medical imaging. While the impact of PGN attacks on CNN performance has been explored, there is a lack of comprehensive studies that apply these findings to medical datasets, which are crucial for real-world healthcare applications.

Additionally, existing defense strategies against PGN attacks have not been adequately tested or validated, leaving uncertainty about their effectiveness in critical medical contexts. Furthermore, the use of comprehensive performance metrics to assess CNN resilience under PGN attacks is inconsistent across the literature.

Our research addresses these gaps by focusing on PGN attacks within medical imaging, validating defense strategies, and conducting a detailed evaluation of model performance, aiming to enhance the robustness and reliability of CNNs in healthcare settings.

	Research 1	Research 2	Research 3	Our Research
Focus on PGN adversarial attacks	Yes	Yes	No	Yes
Impact of PGN on ResNet CNN model performance	Yes	Yes	No	Yes
Evaluation using medical imaging datasets	No	No	Yes	Yes
Test and Validation of existing defense strategies against PGN	No	No	No	Yes
Comprehensive performance metrics under PGN attack	Yes	No	No	Yes

*Table 1: Research Gap*

## **1.4 Research Problem**

The critical issue at hand is the vulnerability of CNNs used in medical imaging to PGN adversarial attacks. These attacks exploit the model's sensitivity by introducing small, almost imperceptible perturbations in the input data, leading to significant misclassifications. In a medical context, such errors can have dire consequences, potentially resulting in incorrect diagnoses and inappropriate treatment plans that could endanger patient safety.

Despite the seriousness of this danger, the specific effects of PGN attacks on CNN performance have not been thoroughly investigated, particularly in the field of medical imaging. Furthermore, we do not have a clear understanding of how effective current defense strategies are against PGN attacks, which creates a gap in ensuring the strength and dependability of these important diagnostic tools. The main focus of the research problem is, therefore, to comprehensively evaluate the resilience of CNN models to PGN attacks and confirm the effectiveness of defense mechanisms in safeguarding against these advanced threats.

## 2. OBJECTIVES

### 2.1 Main Objective

The main objective of this research is to systematically assess the robustness of CNNs used in medical imaging, particularly in chest X-ray classification, against PGN adversarial attacks. This study aims to fill the gap in current research by evaluating the impact of PGN attacks on CNN model performance and validating the effectiveness of existing defense strategies. The ultimate goal is to enhance the security and reliability of CNN models, ensuring that they can provide accurate and dependable diagnostic support even in the presence of adversarial threats, thereby safeguarding patient safety and improving clinical outcomes.

### 2.2 Specific Objectives

To achieve the main objective of enhancing the robustness of CNNs against PGN adversarial attacks in medical imaging, the following specific objectives are set:

#### 1. Implement PGN Adversarial Attacks:

- **Objective:** Develop and apply PGN attacks on CNN models using chest X-ray datasets.
- **Purpose:** This objective aims to simulate real-world adversarial conditions by introducing small, targeted perturbations to the input data. By implementing PGN attacks, the study seeks to uncover how these subtle changes can significantly impact model predictions. This will provide insights into the vulnerabilities of CNN models in a medical imaging context, where even minor misclassifications can have critical consequences for patient outcomes.

#### 2. Fine-Tune CNN Models:

- **Objective:** Fine-tune the CNN models under PGN attack conditions to assess their robustness.
- **Purpose:** Fine-tuning the CNN models involves adjusting the model parameters and architecture to enhance resilience against PGN attacks. This process will help identify weaknesses in the models that are susceptible to adversarial manipulation. By improving the model's robustness, the research aims to reduce

the likelihood of misclassifications in the presence of adversarial perturbations, thus ensuring more reliable diagnostic support.

### 3. Evaluate Model Performance:

- **Objective:** Assess the impact of PGN attacks on CNN model performance using key metrics such as accuracy, precision, recall, and F1-score.
- **Purpose:** Evaluating the model's performance under PGN attack conditions is crucial for understanding the extent of its vulnerabilities. By using a comprehensive set of metrics, the research will quantify the model's ability to maintain accuracy and reliability when faced with adversarial perturbations. This evaluation will provide a detailed understanding of how well the models can withstand such attacks and maintain their diagnostic capabilities.

### 4. Test Existing Defense Strategies:

- **Objective:** Test and validate existing defense mechanisms, such as data augmentation and adversarial training.
- **Purpose:** This objective focuses on assessing the effectiveness of current defense strategies in protecting CNN models from PGN attacks. By applying these defenses, the study will determine whether they can adequately mitigate the impact of adversarial perturbations. The goal is to identify the most effective techniques for enhancing model security, thereby contributing to the development of more robust and reliable medical imaging systems.

### 5. Validate Defense Effectiveness:

- **Objective:** Compare the performance of CNN models on clean and adversarial datasets to validate the effectiveness of the tested defense strategies.
- **Purpose:** Validation is a critical step in ensuring that the defense strategies not only work in theory but also in practice. By comparing model performance on clean and adversarial datasets, the research will verify whether the defenses can maintain model integrity and accuracy in real-world scenarios. This validation will confirm the reliability of the models and their ability to deliver accurate diagnostics even when exposed to sophisticated adversarial attacks.

### **3. Methodology**

#### **3.1 Overall System Diagram**

The overall system diagram outlines the high-level architecture and workflow of the research project, focusing on the assessment and enhancement of CNNs against various adversarial attacks, including Penalizing Gradient Norm (PGN), Split Image Adversarial (SIA), Uniform Scale Mix Mask (USMM), and Style Transfer Manipulation (STM) attacks. This diagram illustrates the interactions between various components, from data collection to model evaluation, ensuring a comprehensive approach to improving model robustness. The diagram includes:

##### **1. Data Collection and Preprocessing:**

- Chest X-ray images are collected as the primary dataset for this study. These images undergo preprocessing, which includes tasks such as resizing and normalization, to prepare the data for effective input into the CNN models. Preprocessing ensures that the images are in a consistent format, enhancing the model's ability to learn from the data.

##### **2. Baseline Model Training:**

- The preprocessed image datasets are used to train CNN models. This stage involves loading pre-trained CNN models, which are then fine-tuned specifically on chest X-ray datasets. The purpose of this stage is to establish a baseline model that performs well on standard, non-adversarial data.

##### **3. Adversarial Attack Implementation:**

- After baseline training, various adversarial attacks, including PGN, SIA, USMM, and STM, are implemented. These attacks are applied to the validated chest X-ray images, creating adversarial examples that are designed to test the model's vulnerability. Each attack introduces unique perturbations that challenge the model's accuracy and robustness in different ways.

##### **4. Defense Strategies Implementation:**

- To counteract the effects of the adversarial attacks, various defense strategies are applied to the CNN models. These strategies include data augmentation, adversarial training, and fine-tuning the models with defenses in place. The goal

is to enhance the model's resilience to multiple types of adversarial attacks by making it more robust against such perturbations.

#### **5. Evaluation:**

- The effectiveness of the defense strategies is evaluated through rigorous testing. The models are tested on both clean and adversarial datasets, with performance metrics such as accuracy, precision, recall, and F1-score being computed. This evaluation assesses the models' overall resilience and their ability to maintain reliable performance under different adversarial conditions.

#### **6. Results Analysis and Visualization:**

- The final step involves analyzing the results and visualizing the performance metrics. This analysis provides insights into the strengths and weaknesses of the defense strategies, offering a comprehensive view of the model's robustness against various adversarial attacks. Visualizations help in clearly understanding the impact of the applied defense mechanisms and the overall performance improvements achieved.

This system diagram serves as a roadmap for the research, ensuring that each stage of the process is systematically addressed to enhance the robustness of CNN models in medical imaging applications across multiple adversarial scenarios.



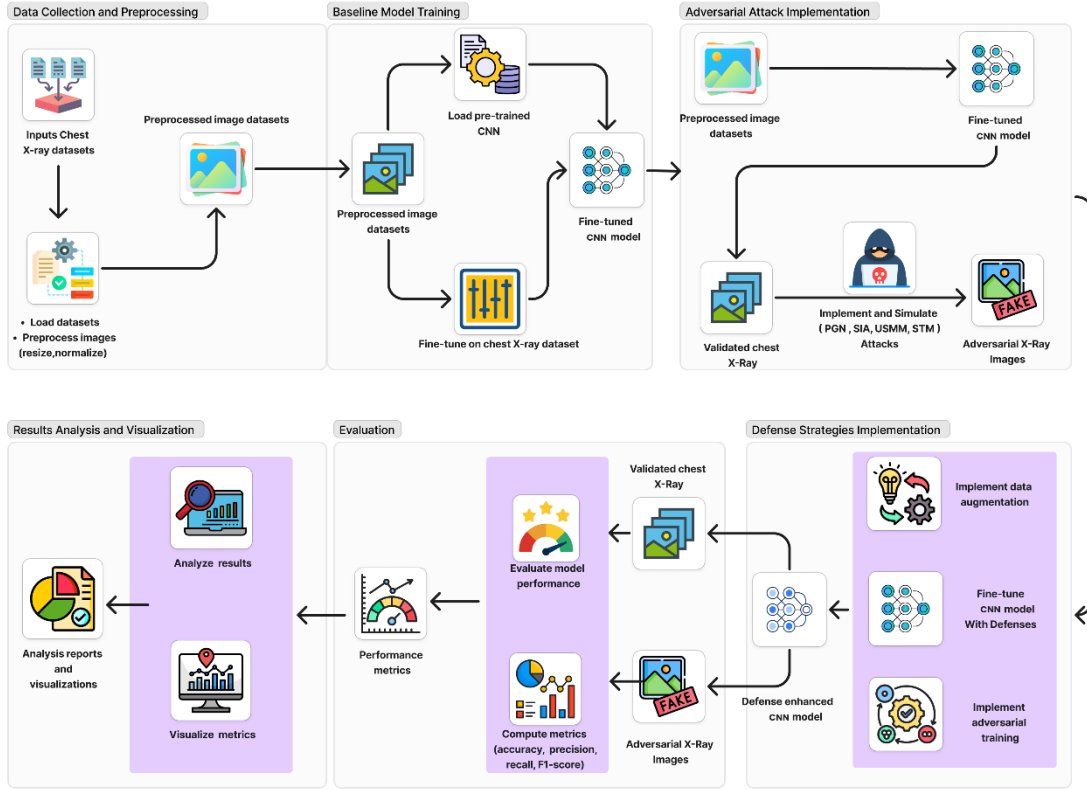


Figure 1: Overall System Diagram

### 3.2 Component System Diagram

The component system diagram for this research focuses specifically on the implementation and assessment of the PGN adversarial attack on CNNs used in medical imaging, particularly for chest X-ray classification. This diagram illustrates the detailed workflow and interactions within the PGN attack component, encompassing the following stages:

#### 1. Data Preprocessing

The process begins with the collection and preprocessing of chest X-ray datasets. Images are resized, normalized, and prepared for input into the CNN models. This ensures consistency and quality in the input data, which is crucial for the accurate training and evaluation of the models.

## 2. **Baseline Model Training**

The preprocessed chest X-ray datasets are used to train CNN models, establishing a baseline of performance under standard conditions. This stage involves fine-tuning pre-trained CNN models to accurately classify chest X-rays. The baseline model serves as a reference point for evaluating the impact of the PGN attack.

## 3. **PGN Attack Implementation**

The core of this component involves implementing the PGN adversarial attack. The PGN attack is applied to the validated chest X-ray images, introducing subtle perturbations that are specifically designed to exploit the sensitivity of the CNN model. These perturbations, while minimal and often imperceptible to the human eye, can cause significant misclassifications, revealing vulnerabilities in the model.

## 4. **Defense Strategies Application**

After the PGN attack has been implemented, existing defense strategies such as data augmentation and adversarial training are applied to the CNN models. This stage is critical in testing whether these defense mechanisms can effectively mitigate the impact of PGN attacks. The models are re-trained and fine-tuned with these defenses to enhance their robustness against the adversarial perturbations.

## 5. **Model Evaluation**

The final stage involves evaluating the CNN models under PGN attack conditions. The models' performance is assessed using key metrics such as accuracy, precision, recall, and F1-score. This evaluation provides a comprehensive understanding of how well the models can withstand PGN attacks and the effectiveness of the defense strategies in maintaining model reliability and accuracy.

## 6. **Results Analysis**

The results from the evaluation are analyzed to determine the success of the defense strategies and the overall resilience of the CNN models to PGN attacks. This analysis highlights the strengths and weaknesses of the models and identifies areas for further improvement.

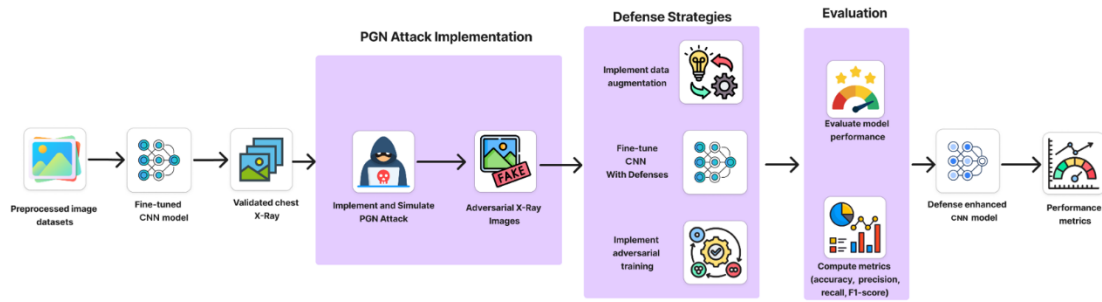


Figure 2: PGN Attack System Diagram

### 3.3 Tools & Technologies

This section outlines the key technologies and tools utilized in this research, providing a detailed overview of the resources and frameworks that underpin the methodology.

#### Deep Learning Frameworks:

- **TensorFlow:** An open-source library that offers a robust ecosystem for building and training deep learning models, particularly CNNs. TensorFlow's flexibility makes it a popular choice for complex neural network development.
- **PyTorch:** Another open-source deep learning framework known for its dynamic computation graphs and intuitive interface, enabling efficient model development and training.

#### Programming Languages:

- **Python:** The core programming language used throughout the research for implementing models, adversarial attacks, and defense strategies. Python's rich ecosystem of libraries and frameworks is essential for deep learning, data processing, and analysis.

#### Model Architecture:

- **Convolutional Neural Networks (CNNs):** The chosen architecture for medical image classification, particularly effective in analyzing visual data due to their ability to detect and capture spatial hierarchies within images.

#### Data Management Tools:

- **Pandas:** A powerful library for data manipulation and analysis, essential for handling and preprocessing large-scale datasets used in this research.
- **NumPy:** A fundamental tool in scientific computing, providing efficient support for large, multi-dimensional arrays and matrices, which are critical in deep learning workflows.

#### **Visualization Tools:**

- **Matplotlib:** A versatile plotting library used to create static, animated, and interactive visualizations, helping to illustrate data trends and model performance.
- **Seaborn:** A statistical data visualization library built on Matplotlib, offering a user-friendly interface for producing aesthetically pleasing and informative graphics.

#### **Evaluation Metrics:**

- **Scikit-learn:** A comprehensive machine learning library that provides a wide range of tools for evaluating model performance, including accuracy, precision, recall, and F1-score metrics.

#### **Integrated Development Environment (IDE):**

- **Google Colab:** A cloud-based IDE that supports Python and offers free access to GPU resources, making it ideal for developing, training, and testing deep learning models in this research.

#### **Hardware & Compute Resources:**

- **Google Cloud:** Cloud computing services that provide scalable and efficient processing power, including access to virtual machines with GPU capabilities, crucial for training CNN models and running extensive computational tasks.

### **3.4 Requirements**

This section outlines the essential requirements for the successful execution of the research project, categorized into system, personnel, and software needs.

### 3.4.1 System requirements

To conduct the research effectively, a robust computational infrastructure is necessary. The system requirements include:

- **High-Performance Computing Power:**
  - **GPUs:** Access to high-performance Graphics Processing Units (GPUs) such as NVIDIA Tesla or RTX series is crucial for training Convolutional Neural Networks (CNNs) on large chest X-ray datasets.
  - **RAM:** A minimum of 16GB of RAM is required to handle the memory-intensive tasks involved in deep learning model training and testing.
  - **Storage:** Solid State Drives (SSDs) with at least 500GB capacity are necessary to store and manage large datasets efficiently, ensuring quick access and processing speeds.
  - **Network:** A high-speed internet connection is essential for downloading datasets, accessing cloud resources, and collaborating with remote team members.
- **Cloud Computing Resources:**
  - **Google Cloud:** Scalable cloud computing services will be used to provide the necessary virtual machines and GPU support for model training and adversarial attack simulations. Cloud resources ensure flexibility and scalability, allowing the research to handle large-scale computational tasks.

### 3.4.2 Personal requirements

This section details the skill set required for the personnel involved in the project:

- **Machine Learning and Deep Learning Expertise:** The team members must possess strong expertise in Machine Learning and Deep Learning, with a focus on training, fine-tuning, and evaluating Convolutional Neural Networks (CNNs). Understanding the principles and techniques behind adversarial attacks and defense strategies is essential.
- **Python Programming Proficiency:** Proficiency in Python is crucial, as it is the primary programming language used in the project. The team should be adept at

working with Python's extensive libraries and frameworks, such as TensorFlow, PyTorch, Pandas, and NumPy, which are necessary for implementing algorithms, managing data, and performing complex computations.

- **Data Analysis Skills:** The ability to analyze and interpret model outputs is vital for understanding the impact of adversarial attacks and the effectiveness of defense strategies. Team members should be skilled in data manipulation, statistical analysis, and the use of relevant tools to derive meaningful insights from large datasets.
- **Visualization Expertise:** Strong skills in data visualization are required to present results clearly and effectively. The team should be proficient in using visualization tools like Matplotlib and Seaborn to create informative graphics that illustrate model performance, trends, and the impact of different techniques.

### 3.4.3 Software requirements

The software tools and environments essential for the project are outlined as follows:

- **Development Environment:** Python serves as the primary programming language, with Google Colab recommended as the integrated development environment (IDE). Google Colab's cloud-based platform supports collaborative work and provides access to powerful computational resources, making it ideal for this project.
- **Deep Learning Frameworks:** The project will leverage leading deep learning frameworks, including TensorFlow and PyTorch, which provide comprehensive tools for building, training, and fine-tuning CNN models.
- **Libraries and Tools:** Several specialized libraries are crucial for the project:
  - **Adversarial Attack Libraries:** Essential for generating adversarial examples to evaluate the robustness of CNN models.
  - **Data Management:** Pandas and NumPy will be used for efficient data manipulation and management.
  - **Visualization:** Matplotlib and Seaborn will be employed to create detailed visual representations of data and model performance.

- **Scikit-learn:** This library will be used for a variety of machine learning tasks, including data preprocessing, model evaluation, and the implementation of basic models.

### 3.5 Gantt Chart

The Gantt chart provides a visual timeline of the research project, outlining the sequence of activities and their respective durations. It serves as a project management tool, helping to ensure that the project stays on track and meets its deadlines.

PROCESS	QUARTER 1				QUARTER 2				QUARTER 3			
	Jun	Jul	Aug	Sept	Oct	Noc	Dec	Jan	Feb	Mar	Apr	May
Project Planning	■	■										
Data Preparation		■	■	■								
Model Development			■	■	■							
PGN Attack Simulation					■	■	■					
Defense Strategy Implementation							■	■	■	■		
Testing and Evaluation										■	■	
Optimization and Refinement											■	■
Documentation												■

Figure 3: Gantt Chart

### 3.6 WBS

The Work Breakdown Structure (WBS) is a hierarchical decomposition of the project into smaller, manageable tasks. It provides a detailed view of the project's scope, breaking down the work into key deliverables and tasks.

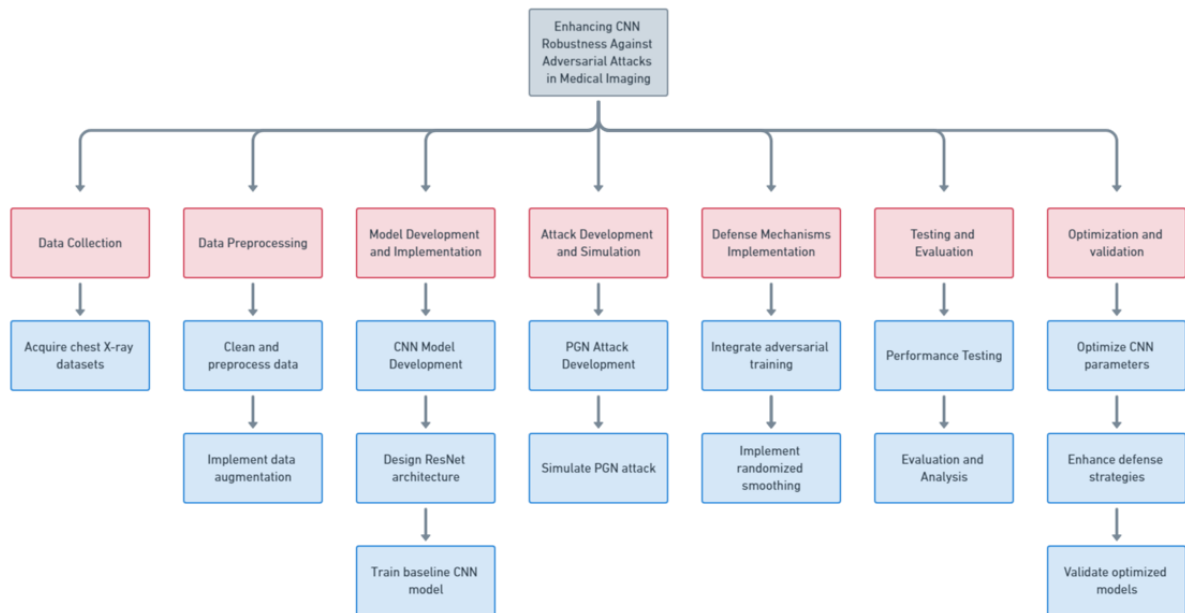


Figure 4: Work Breakdown Structure



## **4. COMMERCIALIZATION AND BUDGET**

### **4.1 Commercialization**

#### **1. Identify Target Markets**

To commercialize medical imaging technology, identify key markets like healthcare providers, software companies, academic institutions, and government health agencies. Analyze their needs and challenges, tailoring product development, marketing, and sales strategies to meet their specific needs, ensuring efficient and effective market entry.

#### **2. Product Development**

Identifying target markets is crucial for product development. Create tailored solutions that meet their unique demands. This can be integrated with existing systems or custom software tailored to specific segments. This ensures the product meets and exceeds the audience's expectations.

#### **3. Demonstrating Value**

To gain market trust, showcase the value of your technology through pilot projects, white papers, case studies, and participation in conferences and trade shows. These methods demonstrate its effectiveness, reliability, and potential benefits, attracting potential customers and partners.

#### **4. Regulatory Approval**

Regulatory approval is crucial in the medical field, ensuring technology meets industry standards and certifications from authorities like FDA or CE. It legitimizes the technology and opens doors to market adoption, ensuring reliability and safety for clinical use.

#### **5. Marketing and Sales Strategy**

A successful marketing and sales strategy is crucial for early product adoption. Educational marketing, direct sales, and tailoring strategies to target audience needs can help educate customers about technology benefits and applications. Tailoring these strategies helps build strong relationships and drive adoption.

#### **6. Pricing Strategy**

A strategic pricing strategy is crucial for ensuring the competitiveness and profitability of your technology. Cost-plus pricing, based on production costs plus a margin, is suitable for high-cost markets, while value-based pricing considers the perceived value of the product.

## 4.2 Budget and Budget Justification

Description	Cost
<b>Hardware Costs</b>	
High-Performance GPUs	LKR 50,000 – LKR 1,00,000
RAM (Minimum 16GB)	LKR 6,000 – LKR 8,000
Storage (SSD with at least 500GB)	LKR 5,000 – LKR 7,000
<b>Cloud Computing Resources</b>	
Google Colab Pro subscription	Approximately LKR 1,000 per month
Software Costs Development Environment and Libraries	Free

*Table 2: Budget Allocation Table*

This budget provides an approximate estimate of the costs associated with the hardware, cloud computing resources, and software required for the project. The Hardware Costs include high-performance GPUs, which range from LKR 50,000 to LKR 1,00,000, crucial for deep learning computations. RAM (minimum 16GB) is estimated to cost between LKR 6,000 and LKR 8,000, while SSD storage (at least 500GB) is expected to be between LKR 5,000 and LKR 7,000. The Cloud Computing Resources budget includes an approximate monthly cost of LKR 1,000 for a Google Colab Pro subscription, offering access to necessary computational power. Additionally, the sSoftware Costs for the development environment and libraries are free, utilizing open-source tools to keep costs minimal. Please note that these figures are approximate and actual costs may vary depending on specific requirements and market conditions.

## 5. REFERENCES

- [1] R. A. Al-Falluji, Z. D. Katheeth, and B. Alathari, "Automatic Detection of COVID-19 Using Chest X-Ray Images and Modified ResNet18-Based Convolution Neural Networks," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 1301–1313, 2021.
- [2] "PGN: A Perturbation Generation Network Against Deep Reinforcement Learning," *Proc. Conf. on Neural Information Processing Systems (NeurIPS)*, 2023.
- [3] X. Ge, L. Wang, and Z. Yang, "Penalizing Gradient Norm for Efficiently Improving Robustness of Neural Networks," *IEEE Trans. Neural Netw. Learn. Syst.*, 2023.
- [4] "Boosting Adversarial Transferability by Achieving Flat Local Maxima," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2023.
- [5] "Deep Learning Based Image Classification of Lungs Radiography for Detecting COVID-19 using a Deep CNN and ResNet50," *IEEE Access*, 2021.
- [6] "Automatic Detection of COVID-19 Using Chest X-Ray Images and Modified ResNet18-Based Convolution Neural Networks," *Computers, Materials & Continua*, 2021.