Technical Admission Report

PhD InfoSec Lab Admission Test

Alif Wicaksana Ramadhan

December 22, 2025

Contents

Executive Summary

Provide a high-level overview of the work completed for the PhD Lab Admission Test. Summarize the key achievements in Computer Vision, LLM Security, and Adversarial Robustness.

Chapter 1

Task 1: Computer Vision (Image Task)

1.1 System Architecture

Describe the high-level architecture of the image processing pipeline here.

1.2 Methodology

1.2.1 Object Detection

Details about the object detection model (e.g., YOLO, Faster R-CNN) used, including training data and configuration.

1.2.2 Classifier

Details about the classification model used to categorize the detected objects.

1.3 Evaluation

Present the performance metrics (Precision, Recall, F1-Score, mAP) and detailed analysis of the results.

Chapter 2

Task 2: LLM Security (RAG Task)

2.1 RAG Pipeline

Explain the Retrieval-Augmented Generation (RAG) setup, including the retriever, vector database, and generator components.

2.2 Data Management

2.2.1 CVE Data

Description of how CVE data is fetched, processed, and stored.

2.2.2 Personal Data

Description of how personal/private data is handled, ensuring separation from public CVE data to prevent leakage.

2.3 Sanitization

Discuss the techniques used for input and output sanitization to mitigate risks such as prompt injection and XSS.

Chapter 3

Bonus Task: Adversarial Attacks

3.1 Attack Methodology

Describe the adversarial attack techniques employed (e.g., FGSM, PGD) and the target model.

3.2 Robustness Results

Analyze the evaluation results, discussing the model's robustness against the generated adversarial examples and any defense mechanisms tested.

Chapter 4

Conclusion

Summarize the overall findings of the admission test tasks. Discuss challenges faced, lessons learned, and potential future improvements.