**AI GENERATED IMAGE DETECTION & HARM**

**PREVENTION SYSTEM**

Higher National Diploma in Software

Engineering 24.2F

Emerging Technology

VITHANA D K W COHNDSE242F-032

RODRIGO P H V V COHNDSE242F-053

FERNANDO M M S COHNDSE242F-054

WANIGASOORIYA J J M K N COHNDSE242F-062



School of Computing and Engineering

National Institute of Business Management

Colombo-07

Project title: AI GENERATED IMAGE DETECTION & HARM PREVENTION SYSTEM

Student index and name:

COHNDSE242F-032 VITHANA D K W

COHNDSE242F-053 RODRIGO P H V V

COHNDSE242F-054 FERNANDO M M S

COHNDSE242F-062 WANIGASOORIYA J J M K N

Name of the program: Higher National Diploma in Software Engineering

Institution: School of Computing and Engineering National Institute of Business Management

Date: 10/27/2025

Table of Contents

**Chapter 1: Introduction1**

Introduction application2

Problem definition3

Proposed solution………………………………………………………………………………………………………………………………..4

Chapter summary………………………………………………………………………………………………………………………………..5

**Chapter 2: Methodology4**

Introduction5

Data Collection methods6

Software Process Model………………………………………………………………………………………………………………………6

Testing Strategies…………………………………………………………………………………………………………………………………7

Implementation……………………………………………………………………………………………………………………………………7

Chapter Summary………………………………………………………………………………………………………………………………..8

**Chapter 3: Analysis9**

Introduction10

UML Diagrams11

Use Case Diagram………………………………………………………………………………………………………………………..11

Class Diagram………………………………………………………………………………………………………………………………12

Sequence Diagrams……………………………………………………………………………………………………………………..13

ER Diagram…………………………………………………………………………………………………………………………………..13

Chapter Summary………..……………………………………………………………………………………….……………………………12

**Chapter 4: Finalized Design13**

Introduction13

Interface Design6

Database Design…………..………………………………………………………………………………………………………………………6

**Chapter 5: Business Model13**

**Chapter 6: Conclusion13**

**References13**

Chapter 1: Introduction

1. Introduction Application

Generative artificial intelligence (AI) has grown so quickly that it has changed the way digital information is made and shared. With the use of tools like Midjourney, Stable Diffusion, DALLE, users can now create incredibly realistic synthetic images with a level of ease never before possible. These technologies have enormous possibilities for design, entertainment, education and innovation. But these advantages come with significant hazards and difficulties as well.

Malicious uses of AI-generated photos include disseminating false information, creating false evidence. Posing as people and swaying public opinion. Such abuse jeopardizes individual reputations, presents more general risks to societal stability, politics and journalism, and creating false evidence. These issues are exacerbated by the lack of easily available and trustworthy verification methods for regular users and which leaves the general public open to fraud.

2. Problem Definition

With the explosive advancement of generative AI technologies, it has become increasingly easy to create highly realistic fake images. Tools like Midjourney, DALL·E, and Stable Diffusion allow users to generate synthetic content that is often indistinguishable from real photography. While these innovations have powerful creative applications, they also present significant ethical and societal challenges.

One of the most pressing concerns is the misuse of these technologies to produce deceptive or harmful content. AI-generated images can be used to defame individuals, impersonate public figures, fabricate evidence, and disseminate misinformation. These capabilities pose threats to personal reputations, political stability, journalistic integrity, and public trust in visual media.

Unfortunately, there is currently no accessible, widely used solution that allows non-expert users to verify whether an image is authentic or generated by AI. While some academic and corporate tools exist, they are often not user-friendly, lack transparency, or are inaccessible to the public.

1. Proposed Solution

Our proposed solution consists of a web application with a strong backend that integrates a pre-trained AI model capable of analyzing images and predicting whether they are AI-generated or real. The model will rely on advanced deepfake, or GAN-detection techniques trained on large datasets, potentially using frameworks and weights hosted on platforms such as Hugging Face.

**Key Features**

* **Image Upload & Scan**

Users can drag and drop or upload images from their device. Once uploaded, the system will process the image and return a result indicating whether it is likely real or AI-generated.

* **Detection Result & Confidence Score**

The backend will return to a prediction along with a confidence percentage (e.g., 92% AI-generated). This helps users assess the likelihood of manipulation.

* **Report Harmful Images**

Users can flag harmful or deceptive images, optionally providing context. Reports are logged in the database and can be reviewed by moderators or system administrators.

* **Scan History**

Logged-in users will have access to their personal scan history, allowing them to track past checks and view previous results.

* **Educational** **Resources**

To support digital literacy, the site will include links to articles, guides, and video explainers from trusted institutions such as the Electronic Frontier Foundation, MIT Media Lab, and journalism verification tools.

1. **Chapter Summary**

In this chapter we discuss the idea behind our project and explained why we decided to built this system. And discusses the problems with the growing popularity of AI generated imagery and offers a solution. Also listed the goals we want to achieve and gave a basic idea of how the solution will work. Solution describes a specialized online application that empowers people by offering picture verification, a reporting system and instructional materials.

Chapter 2: Methodology

1. **Introduction**

This chapter explains that how we planned and completed our project. The methodical process used to create the web application for AI generated picture identification and the process is intended to guarantee the development of a reliable, expandable and user focused solution. And also describe the tools and development model we used and describe how we gathered the information needed to understand what features the system should have.

1. **Data Collection Method**

* We performed interviews with journalists, social media users, and educators who often interact with digital content in order to comprehend user difficulties and skepticism.
* We sent a poll to a larger online community to get input on current verification technologies and user requirements.
* We spoke with experts in digital forensics and AI researchers to obtain technical and operational insights.
* We assessed current image verification platforms and scholarly detection techniques to assess the advantages and disadvantages of competing strategies.
* Utilizing Public Datasets: The primary data will come from sizable, publicly accessible datasets on websites such as Hugging Face and Kaggle, which comprise thousands of labeled "real" and "AI-generated" images.
* Diverse Generator Sources: To guarantee that the model performs well in general, the AI-generated portion of the data will contain images produced by a range of models, such as GANs (like StyleGAN), Diffusion Models (like Stable Diffusion, DALL·E), and other architectures.

1. Software Process Model

Agile Development Model was applied and it divides work into manageable chunks, each of which is completed and tested before going on to the next. This allowed us to get feedback early and make changes if needed. We held regular meetings with our group to check progress and solve any problem quickly. This made the project more organized and easier to manage.

1. Technologies Used

* Frontend: ReactJS / HTML-CSS-JS
* Backend: Python (Flask) or Node.js
* Database: MongoDB
* AI Model: Pre-trained deep-fake detection model (from Hugging Face)

1. **Chapter Summary**

In this chapter, we explained how we planned, developed and tested our web application. The method is supported by rigorous testing strategies for both software and AI components, an Agile Process Model for flexible and iterative development and a strong Data Collection strategy to train a gene. We used simple tools and an easy-to-follow development method. This helped us create a system that is functional and user friendly.

Chapter 3: Analysis

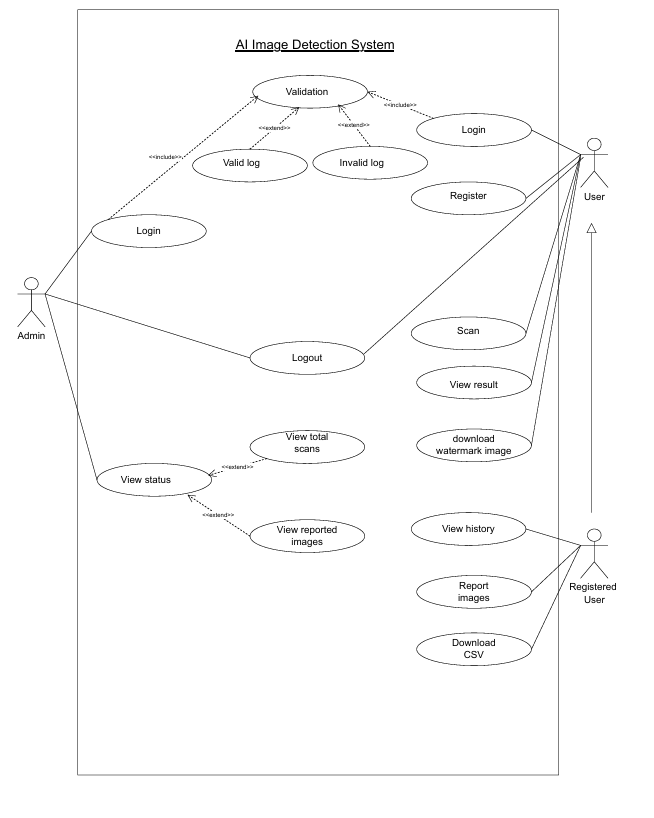
1. **Introduction**

The method of problem analysis was discussed in this part. UML diagrams are covered.

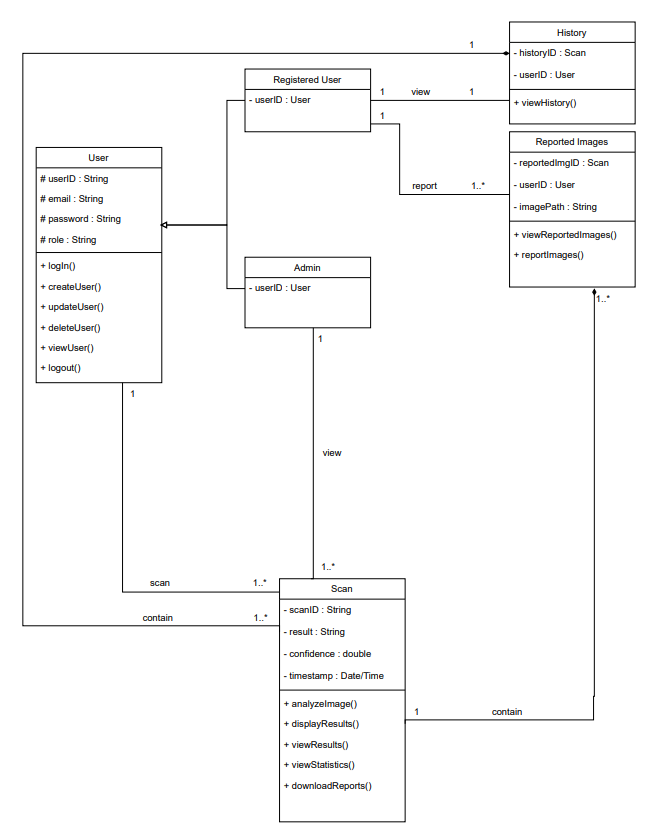
The use case diagram for this application is detailed in the UML Diagrams section. It also discusses sequence diagrams (each use case), class diagrams and the application's ER Diagram.

1. UML Diagram

Use Case Diagram

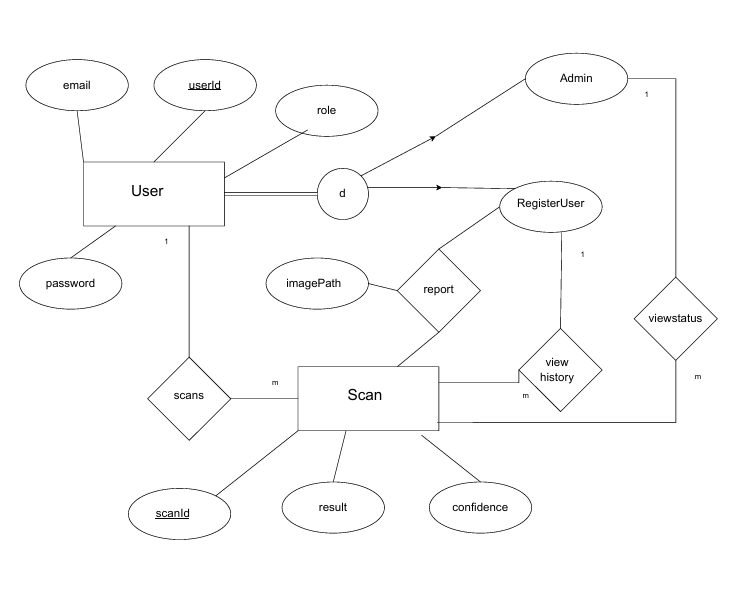


Class Diagram



Sequence Diagram

ER Diagram



1. Summary

Using key diagrams that illustrate the system's behavior, structure, and data relationships, this chapter offers a thorough picture of the system design. The Use Case Diagram shows how the system's actors and Users—interact with its features, which include handling service data, uploading photos, and managing user profiles. Through connections like include and extend, these interactions specify the essential characteristics and dependencies.

Chapter 4: Finalized Design

1. Introduction

Chapter 5: Business Model

