Software Requirement Specification Document for Video Verify

Nezar Ahmed, Mohamed Amr, Mohamed Osama, Sandra Khalid
Supervised by: Dr.Khaled Hussein and Eng.Tarek Talaat

April 29, 2024

Table 1: Document version history

Version	Date	Reason for Change
1.0	20-November-2023	Proposal First version's specifications are defined.
1.1	29-Dec-2023	SRS First version's specifications are defined.

GitHub: https://github.com/Nezareltelbany/Graduation-Project

Contents

Abstract

The main idea of this project to develop an application that helps in generate and detect deep-fake . Recently with the advancement of technology and the presence of many good image manipulation tools, people can easily be deceived as deep-fake algorithms can quite effortlessly generate fake videos and images that an ordinary eye cannot distinguish; thereby recently challenging the reliability of online information. Forged videos are the video containing fake images over the real ones,there are methods used with Machine and deep learning approaches which will be used data-set is made up of deep-fake and authentic videos to detect such manipulations and there will be various techniques used to tell apart if it's real from fake nowadays using face swapping or is there something off about it behaviour, or if another person's voice is used with a voice of person, etc.This project aim to create a generation and detection application.

1 Introduction

1.1 Purpose of this document

The actual aim of SRS document is towards focusing over project requirements with respect to software implementation needed for our graduation project (Video Verify). This is mainly focused on face detection and generation . The main purpose to develop an efficient Video generation and Detection system for finding out accurate person identification from video using the latest computer vision techniques and machine learning algorithms.

1.2 Scope of this document

This document's scope covers the objectives of the online web application as well as the potential user characteristics. It involves deepfake generation detection related systems. It provides an illustration of the system's goals, background, and overview. The functional and non functional requirements, data design, operational scenarios, and project plan are all covered in addition to this.

1.3 Business Context

Deepfake can affect anyone, but it can especially target celebrities, actors, politicians and other popular people, To protect themselves from manipulation or extortion. The system assists users in revealing whether their videos is real or fake. Moreover, the revenue of the deep fake detection project can be gained from individuals that intend to utilize the web application in order to verify the authenticity of specific movies by adding them to the website.

Also, AI models may be trained more effectively and with less work by employing Deepfakes, in order to produce large amounts of training data. deepfake detection technology can help AI models in a variety of business areas.

2 Similar Systems

2.1 Academic

1. [John Doe and Jane Smith.] With the rise of misinformation and manipulated media, maintaining trust in online content becomes increasingly challenging. This study presents a novel deep fake detection system tailored for social media platforms, aiming to combat the spread of deceptive content. By leveraging advanced machine learning algorithms and user engagement metrics, the system identifies and flags suspicious media content in real-time, offering a proactive solution to mitigate the impact of misinformation.

2. [Michael Johnson and Emily Brown.] In the age of digital manipulation, ensuring the integrity of multimedia content is paramount. This research introduces a robust deep fake detection framework designed specifically for forensic analysis in legal proceedings. By integrating state-of-the-art image processing techniques and metadata analysis, the system provides forensic experts with the tools to authenticate digital evidence accurately, strengthening the credibility of multimedia content in legal contexts.

3. [David Garcia and Maria Rodriguez.] As deep fake technology evolves, the risk of its malicious use in cyberattacks escalates. This paper presents an innovative deep fake detection system tailored for cybersecurity applications, offering real-time protection against fraudulent multimedia content. By leveraging anomaly detection algorithms and network traffic analysis, the system identifies and neutralizes deep fake threats in digital communication channels, enhancing cybersecurity resilience in the face of emerging threats.

4. [Sarah Lee and Mark Wilson.] With the increasing prevalence of deep fake videos, preserving trust in digital archives becomes imperative. This study introduces a novel deep fake detection system optimized for archival preservation, aiming to safeguard the integrity of historical multimedia collections. By combining archival metadata analysis with deep learning algorithms, the system automatically identifies and filters out deep fake content, ensuring the authenticity of digital archives for future generations.

5. [Alexandra Chen and Christopher Park.] In the era of remote learning and virtual classrooms, ensuring academic integrity in online assessments is paramount. This research presents an innovative deep fake detection system tailored for e-learning platforms, offering real-time protection against fraudulent student submissions. By integrating facial recognition technology and keystroke analysis, the system detects and flags suspicious behavior during online assessments, preserving the integrity of academic evaluations in virtual learning environments.

6. [Daniel Thompson and Rachel Adams.] With the democratization of deep fake technology, preserving trust in online marketplaces becomes crucial. This paper introduces a novel deep fake detection system optimized for e-commerce platforms, aiming to combat fraudulent product reviews and counterfeit listings. By analyzing multimedia content and user engagement patterns, the system identifies and removes deep fake content from online marketplaces, enhancing consumer trust and confidence in digital commerce.

2.2 Business Applications

There are different applications for face recognition:

•Video Authenticator tool [1]: Is a tool developed by Microsoft in order to detect deepfake videos with high accuracy.

•KaiCatch [2]: KaiCatch is a mobile application developed by a Korean professor which uses neural network in order to detect deepfake images and videos. That professor agreed that this application can reach reliability of 90 percentage . This application is currently available for android devices but will soon be available for IOS devices as well.

3 System Description

3.1 Problem Statement

The widespread use of deepfake technology raises serious questions regarding identity theft, false information, and the decline in public confidence in visual media. It also offers a serious danger to the legitimacy of digital material. Robust deepfake detection systems are required when traditional content verification methods become outdated due to the increasing sophistication of deepfake generating techniques. The urgent need for a comprehensive solution that can quickly and reliably identify deepfake content across many multimedia formats is addressed in this Software Requirements Specification (SRS) article. The difficulty is in developing a deepfake detection system that can identify modified images, and videos in quickly changing and dynamic online contexts. To produce convincing deepfakes, the system needs to maneuver over the complex terrain of generative adversarial networks (GANs) and other state-of-the-art tactics used by hostile actors. It also needs to deal with the volume and velocity of deepfake material dissemination, which calls for an effective and scalable algorithmic solution. This SRS paper addresses the complicated interactions of accuracy, speed, and adaptability required to counter the growing threat of misleading multimedia content in modern digital ecosystems. It attempts to outline the functional and non-functional requirements of a cutting-edge deepfake detection system.

3.2 System Overview

the following describes the system overview
.The system is separated into two parts detection and generation.
Detection side
- the dataset is used and preprocessing is done in order to extract faces from the videos/images which are then passed to machine learning classifier (cnn, vgg,etc.) and training is done on the dataset imported.
- The user uploads a video/image to the website in order to detect it and the result is shown on the user interface.
Generation side
the user uploads a source video and a target photo and the model does frames extraction on both

video (the source and the target), then frames are extracted and gathered to be trained in order to match the face of the source video to the face of the target photo,The training process consists of the identity injection network model(IIN) which contains several identity injection modules (IIM) which predicts weights of IDN and the IDN architecture is used to form the face swapping.
Finally the face is merged to the destination video the deepfake video is generated

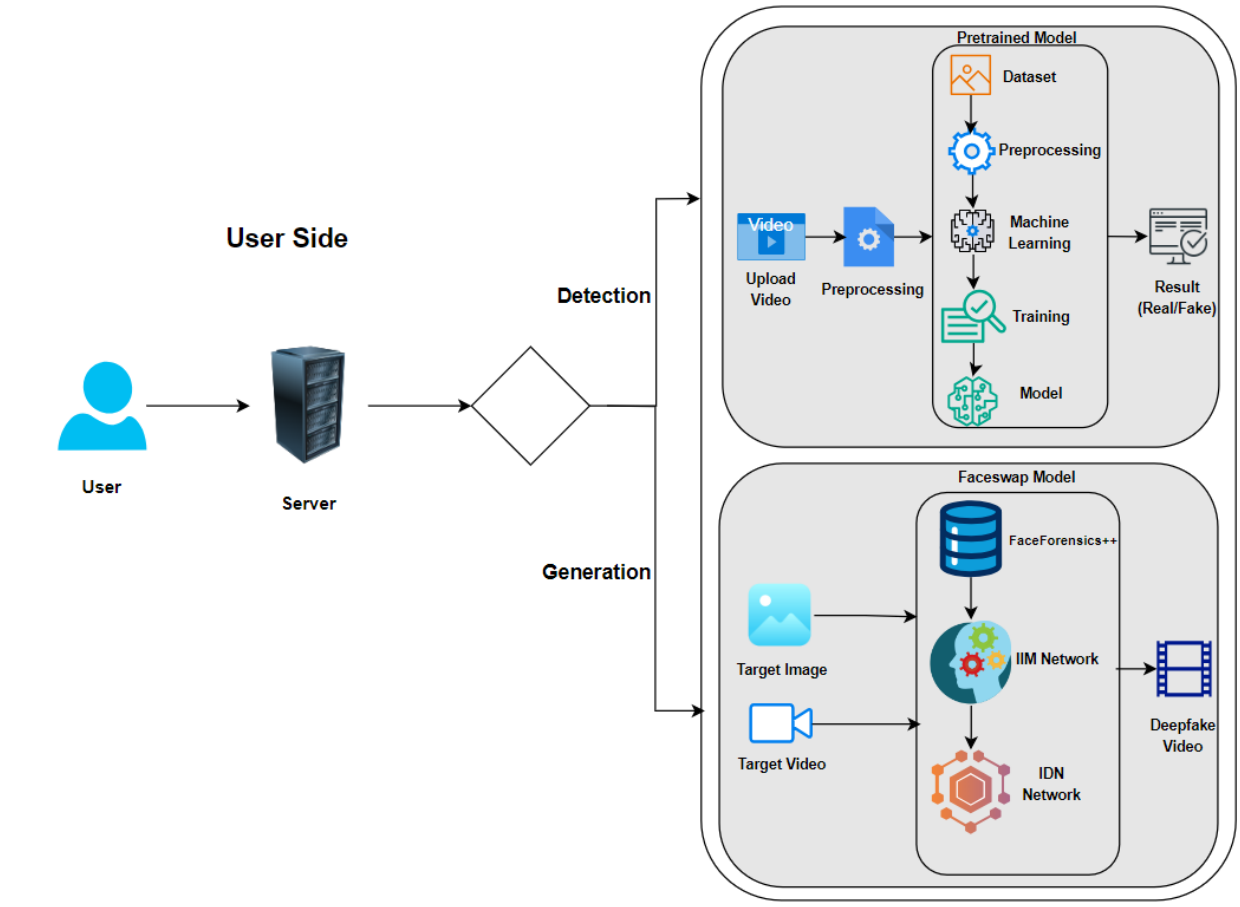

Figure 1: System Overview.

3.3 System Scope

The Video Verify project aims to help in detection of deep fake which look like the real people so the system will target thefollowing :

- The system will extract the facial features from the videos using pre-processing techniques.
- The system will use deep learning techniques in order to detect learn-able features and classify them
- We also target developing the algorithm to be able to provide us with a higher accuracy rate of classification with respect to previous researches done regarding the detection of deep fake.

3.4 System Context

As shown in Figure 2 the following describes the contents of the system:

The deep fake detection and generation website is designed with a focus on user interaction, providing a straightforward and efficient interface for video uploads. Users are guided through a simple process to submit videos for authenticity analysis, ensuring ease of use regardless of technical expertise. The system prioritizes a seamless user experience, from initial access through to the completion of the verification process.

Upon video submission, the system initiates a comprehensive processing sequence. This involves the application of a deep learning decoder, which has been extensively trained on a diverse dataset of videos to accurately identify deep fakes. The processing stage is crucial, as it determines the effectiveness of the detection algorithm, ensuring that the system can reliably differentiate between real and altered content.

The deep learning decoder operates on the principles of neural networks and machine learning, continuously improving its detection capabilities through ongoing training and testing. This iterative process enhances the decoder's precision, enabling it to stay abreast of the latest deep fake generation techniques. The system's robust architecture supports this dynamic learning environment, facilitating real-time updates and adjustments to the detection algorithms.

Finally, the results of the analysis are presented to the user in an intuitive format, clearly indicating whether the video is authentic or has been manipulated. This output is not only informative but also educational, raising awareness about the prevalence and sophistication of deep fake technology. By providing users with reliable detection and insightful information, the website serves as a valuable resource in the fight against digital misinformation.

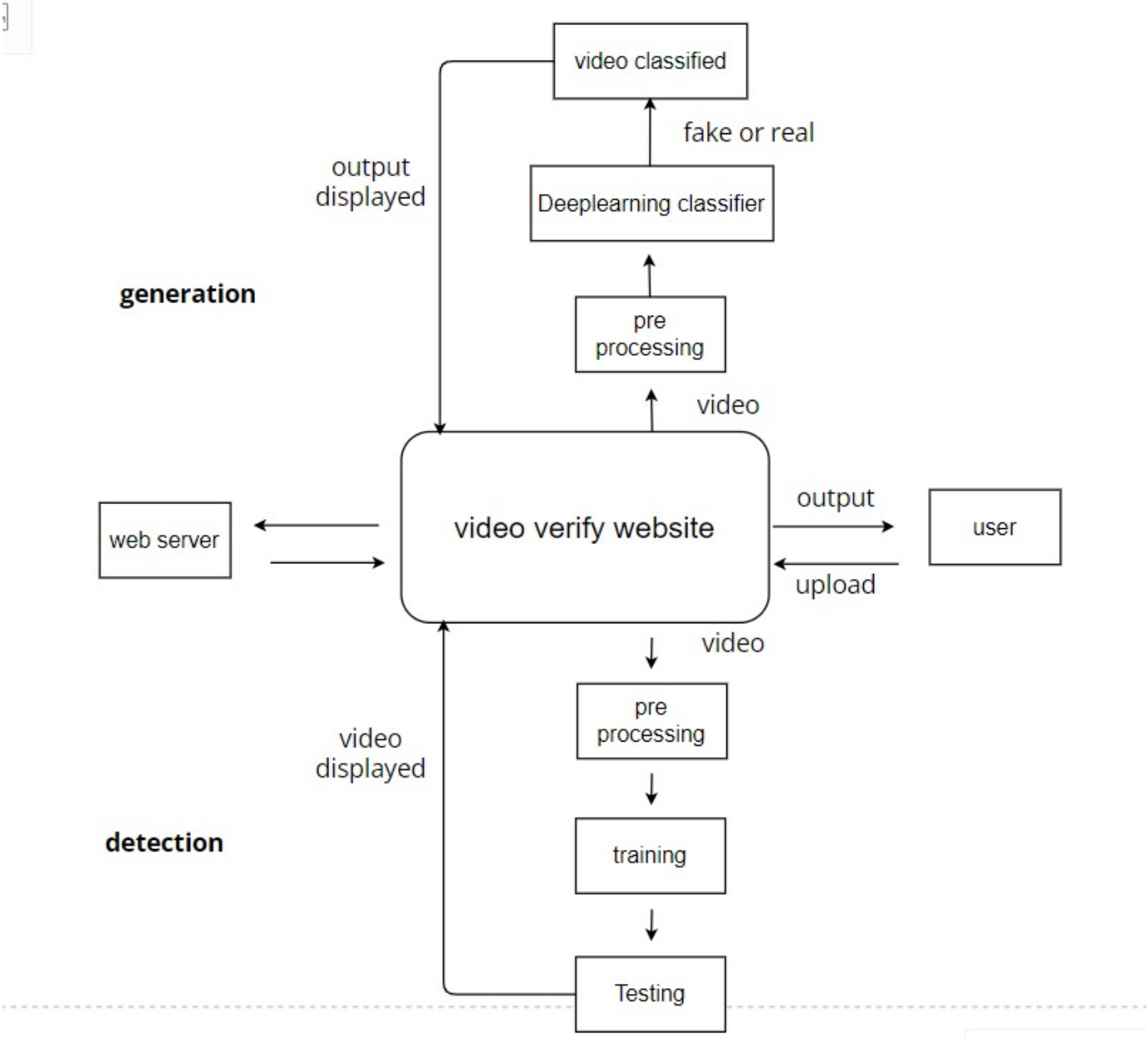

Figure 2: System Context.

3.5 Objectives

The Main Goals of the deep fake detector are:

• This project will provide a user-friendly detection

• Increase the accuracy of the deep fake detection .

• Build a platform that guards the integrity of digital content, preserve trust in online information, mitigate the potential damage caused by deceptive media

3.6 User Characteristics

1. The user can be of any age group.
2. Whether a person is the subject of the video or not, they can use website.
3. No limitations on the role of the user (student, public figure, higher authorities)

4 Functional Requirements

4.1 System Functions

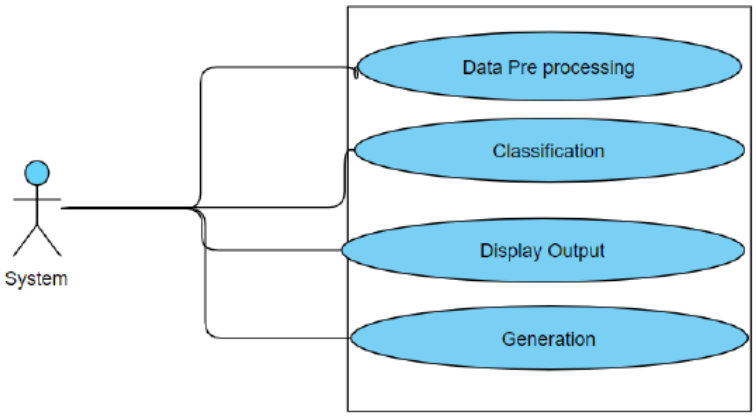

Figure 3: system usecase.

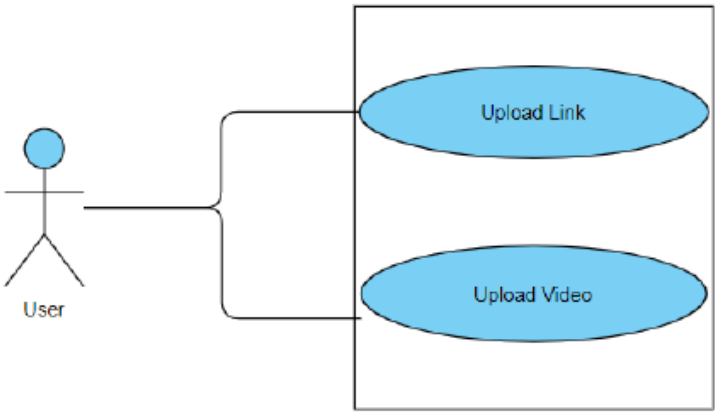

Figure 4: user usecase.

FR01: The system shall apply pre-processing techniques
FR02: The system shall classify whether the input is real or fake
FR03: The system shall display the output to the user .
FR04: The system shall generate a deepfake .
FR05: The user shall upload video/image

4.2 Detailed Functional Specification

Name	Preprocessing
Code	FR01
Priority	High
Critical	To detect faces from videos
Description	This function is responsible for extracting the faces from each video frame
Input	Video
Output	Processed images
Pre-condition	Availability of input video/image
Post-condition	Pre-processed video/image is generated
Dependency	on uploading a video/dataset being available
Risk	Faces not extracted accurately

Table 2: Preprocessing function detailed description

Name	Classification
Code	FR02
Priority	High
Critical	To classify the video as real or fake
Description	Detects whether the frames extracted from the video is real or fake
Input	Pre-processed video
Output	Classification result (real/fake)
Pre-condition	Preprocessing then training
Post-condition	An output message to the user using the system indicating the video is real or not
Dependency	On preprocessing
Risk	Misclassification leading to incorrect detection

Table 3: Classification function detailed description

Name	Display Output
Code	FR03
Priority	High
Critical	To display final output to user
Description	Displaying output (classification or generated deep fake content, to user interface
Input	Classification result/deep fake content
Output	Displayed output to user interface
Pre-condition	Availability of classification result/deep fake content
Post-condition	Output is visually presented to the user
Dependency	Depends on classification/generation
Risk	Output not displayed accurately or in a user-friendly manner

Table 4: Display Output function detailed description

Name	Deep Fake Generation
Code	FR04
Priority	High
Critical	To generate the deep fake video
Description	Generates deep fake content by synthesizing media content with altered attributes
Input	Pre-processed video
Output	Generated deep fake content
Pre-condition	Pre-processed video
Post-condition	Deep fake content is generated
Dependency	Depends on pre-processing
Risk	Generated content lacks realism or authenticity

Table 5: Deep Fake Generation function detailed description

5 Design Constraints

This section is to provide a detailed look on the system limitations and what would be an issue for us while using the system and the allowed approaches like how fast can the system work or what's the size of the uploaded videos that the system can work with for example the system works with about 10 secs videos which is around 2 to 5 Mb's so based on the given information it will be more helpful to know what can and cannot be done using this system.

5.1 Standards Compliance

The VideoVerify will run using a Web Server and the web application itself will be accessed through a web browser as google chrome.The user should be connected to the internet to access the system.

5.2 Hardware Limitations

An HD webcam is required or a dedicated cam for the portable device that is at least 720p for better image pre-processing and hence better classification.

5.3 Network Limitations

Internet speed and bandwidth should be of an acceptable speed let's say 20MBps as transmitting the high quality captured images needs a sufficient bandwidth and throughout to avoid bottle neck issue.

5.4 Other Constraints as appropriate

6 Non-functional Requirements

1.Speed: The device (laptop, mobile,etc.) shall be connected to high speed internet in order to function properly for better performance.

2.Usability: The User interface is going to be user friendly.

3.Performance: The website should respond fast to the users commands.

4.Availability: The website should be always available for the user whenever he wants to access it.

5.Scalability [3]: the website should be able to handle the amount of data added if it is increased and perform well.

7 Data Design

The dataset used in this project is the "deepfake detection challenge (DFDC)", As, The overall size of the dataset is 471 GB so, we will work on 2 folders, training and testing. each folder contains 400 videos and size of each folder about 4 GB. Consists mixture of real and fake videos. Regarding the training side, pre-processing is performed to each video using the classifier converting it to images, then the images are being cropped to focus on the face dimensions. Finally, the images are sent to the proposed architecture in frames to perform the classification. A sample of the dataset is shown below. source of the data : https://www.kaggle.com/competitions/deepfake-detection-challenge/data

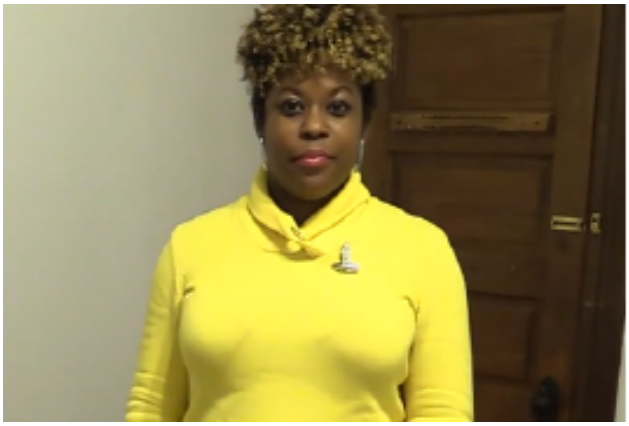

Figure 6: Real Image.

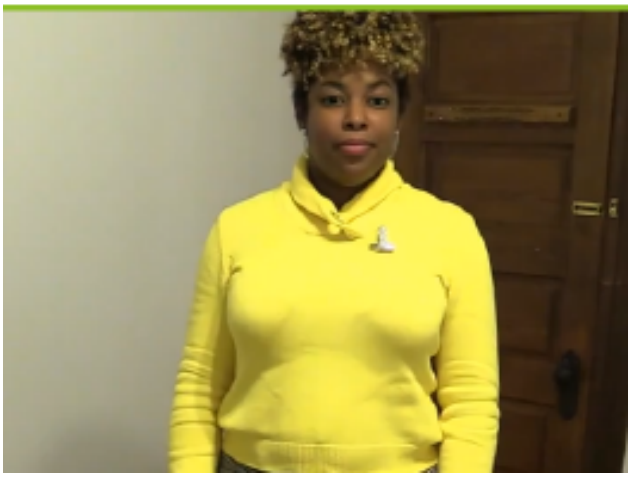

Figure 7: Deep-Fake Image.

8 Preliminary Object-Oriented Domain Analysis

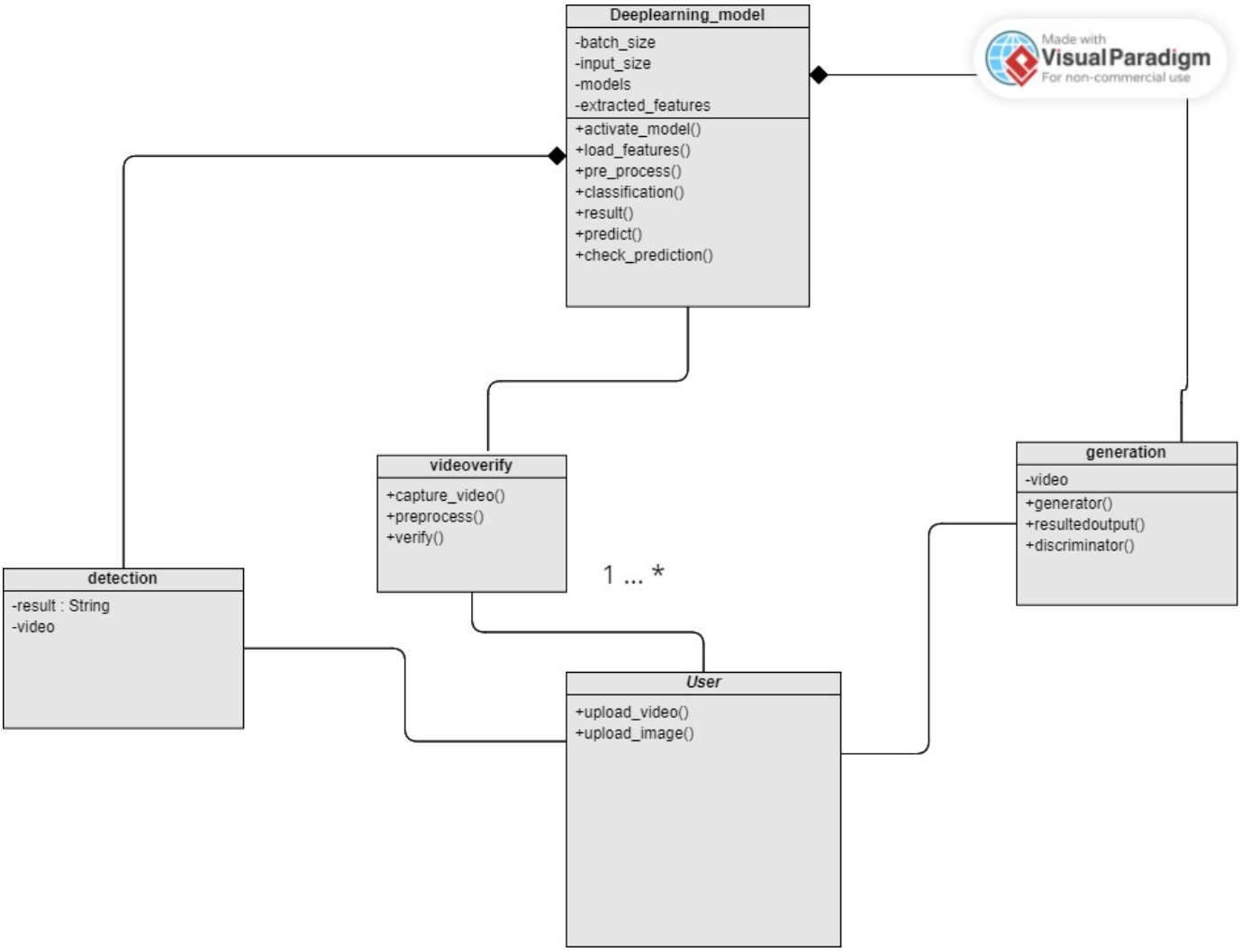

Figure 8: Class Diagram.

9 Operational Scenarios

Scenario 1:- The user will access the website .
Scenario 2:- The user should upload the video to the video's website and the resulting output would be that it's a fake video or a real one.
scenario 3:- The user can upload video and image to generate a deepfake videos .

10 Project Plan

Detailed plan from Proposal to SDD.

Task	Start date	End date	Duration	Role
Research Paper	12/11/2023	14/11/2023	2 days	All team members
Starting The Proposal	12/11/2023	15/11/2023	3 days	All team members
Survey and proposal preparation	12/11/2023	15/11/2023	9 days	All team members
10 percentage of code implementation	12/11/2023	15/11/2023	3 days	All team members
Presenting the proposal	16/11/2023	20/11/2023	4 days	All team members
SRS preparation	30/11/2023	15/01/2024	45 days	All team members
SDD preparation	20/1/2024	20/02/2024	30 days	All team members

Table 2: Time plan

11 Appendices

11.1 Definitions, Acronyms, Abbreviations

Algorithm	Description
CNN	Convolutional Neural Network, a type of neural network algorithm used to extract images or objects (e.g., faces).
MesoNet	An algorithm used to automatically detect the editing of facial features in videos.
VGG	Visual Geometry Group; a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. The "deep" refers to the number of layers, with VGG-16 or VGG-19 consisting of 16 and 19 convolutional layers. The VGG architecture is the basis for ground-breaking object recognition models, surpassing baselines on many tasks and datasets.
ResNext	A type of deep neural network architecture designed for image classification tasks.

Table 3: Overview of Neural Network Algorithms

11.2 Supportive Documents

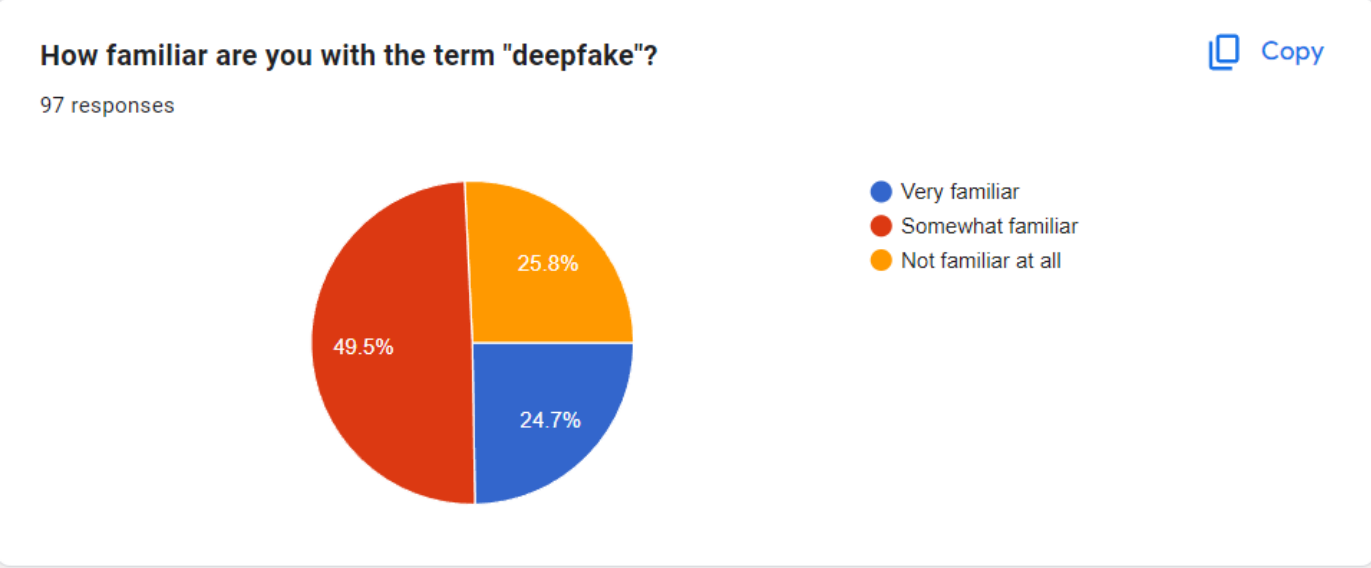

Figure 9

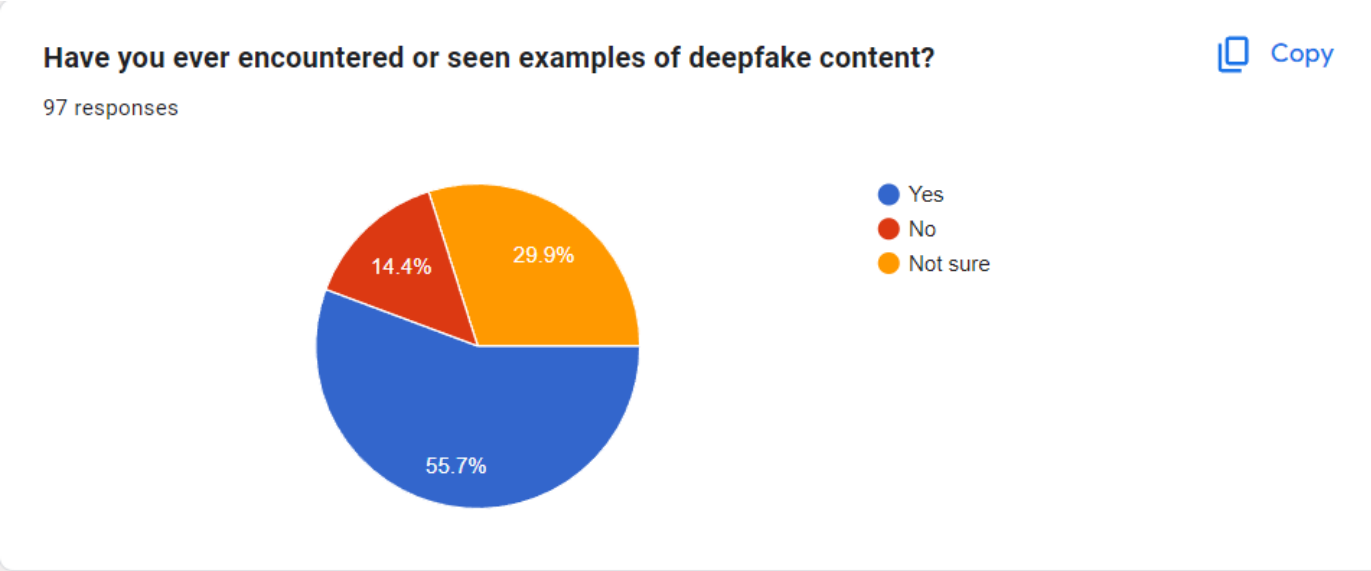

Figure 10

What potential negative impacts of deepfake technology concern you the most? (Select all that apply)

Copy

97 responses

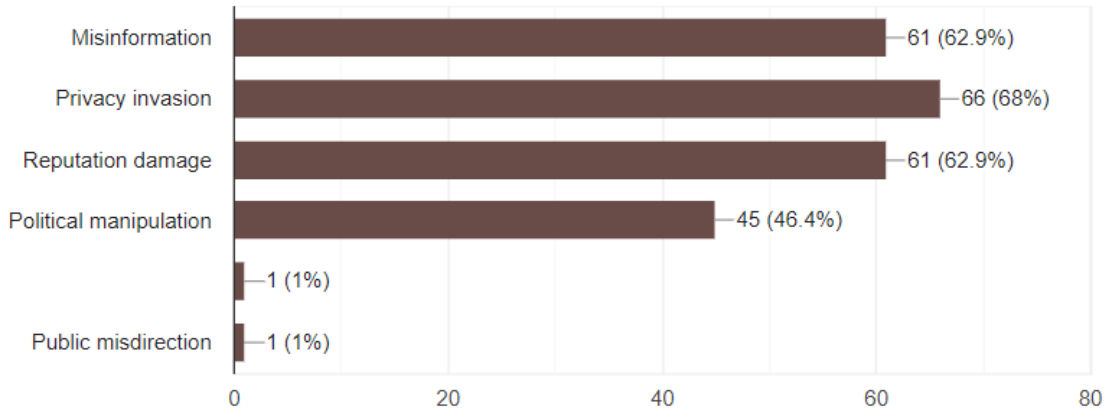

Figure 11

Do you believe it is important to have reliable systems for detecting deepfake content?

Copy

97 responses

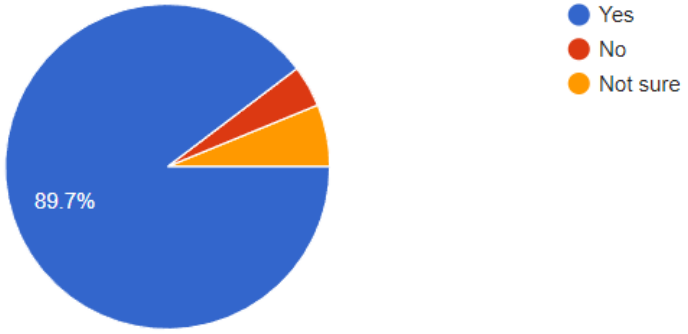

Figure 12

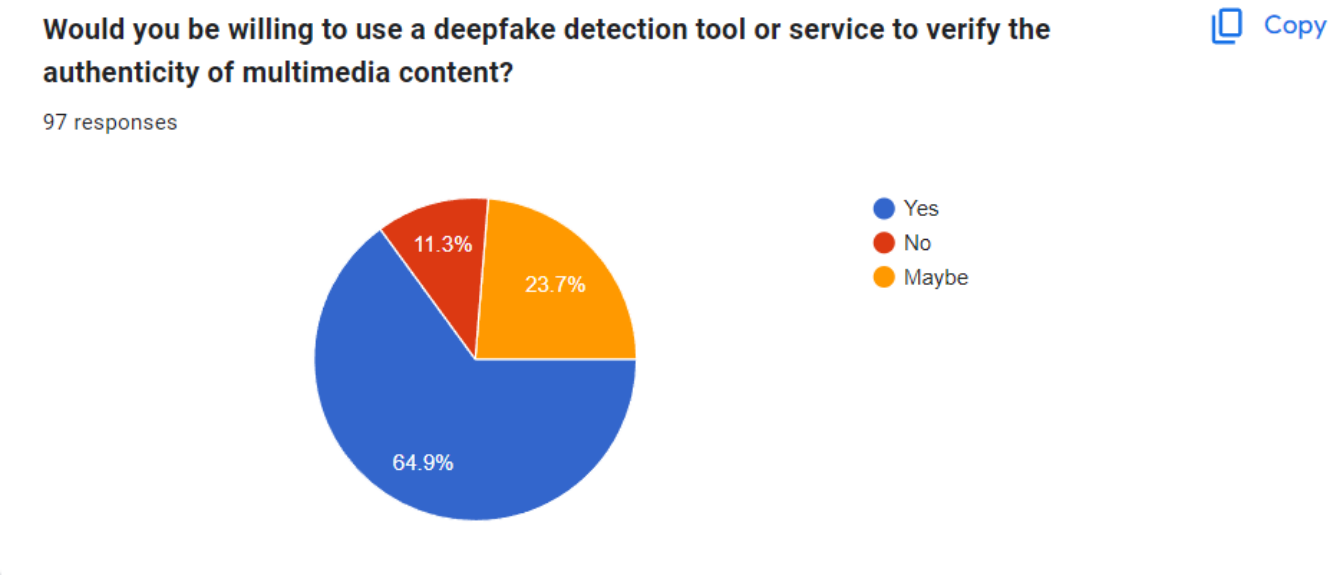

Figure 13

References

[1] Vineet Mehta, Parul Gupta, Ramanathan Subramanian, et al. “Fakebuster: a deepfakes detection tool for video conferencing scenarios”. In: *26th International Conference on Intelligent User Interfaces-Companion*. 2021, pp. 61–63.

[2] Dymples Leong Suying. “Deep fakes and Disinformation in Asia”. In: *Deep Fakes*. Routledge, 2022, pp. 23–49.

[3] Martin Glinz. “On non-functional requirements”. In: *15th IEEE international requirements engineering conference (RE 2007)*. IEEE. 2007, pp. 21–26.