Software Requirement Specification Document for Deepfake Detection

Zeina Ayman, Natalie Sherif, Mariam Mohamed and Mohamed Hazem
Supervised by: Dr. Diaa Salama and Eng. Samira Refaat

May 13, 2023

Table 1: Document version history

Version	Date	Reason for Change
1.0	5-Nov-2022	SRS First version's specifications are defined.
1.1	7-Nov-2022	functional requirements added appendices are added
1.2	11-Nov-2022	system context and use cases diagrams are added functional requirements are updated
1.3	15-Nov-2022	class diagram and use cases diagrams are updated functional requirements are updated
1.4	17-Nov-2022	class diagram is updated
1.5	2-May-2023	system overview diagram is updated

GitHub: https://github.com/zeina20/GraduationProject

Contents

Abstract

The main idea of this project is to develop a web application that can help to detect whether the input data we provide of people whether celebrities or people in general are real or fake. Recently with the evolution of technology and with advanced image editing tools, people can easily get manipulated, as deepfake algorithms can easily create fake videos and images that people can't distinguish from authentic ones, which is an emerging problem threatening the trustworthiness of online information. Deepfakes mainly affect public figures, celebrities, and politicians. Forged videos are videos that contain fake images over real ones and in this research, there are methods used with Machine and deep learning approaches that will be used with the dataset that is composed of deep fake videos and authentic ones to detect these manipulations and protect the government from criminals and there will be various techniques used to distinguish real from fake using face swapping or is there something off regarding it's behaviour, or if a voice of a person is used with another person's voice , etc. The deep fake detector can be used in courts and police stations to reduce the likelihood of crimes and frauds that may happen and detect them. This project aims to make a website to detect whether videos are fake or not.

1 Introduction

1.1 Purpose of this document

The purpose of this document is to focus on the project requirements with respect to software implementation needed. It acts as a guide for the developers to know what they need to do in order to develop the web application.

1.2 Scope of this document

The scope of this document addresses the web application's goals and possible users' characteristics. It explores similar systems to deepfake detection. It illustrates the overview, context, and objectives of the system. In addition to describing the functional and non-functional requirements, data design, and finally the operational scenarios and project plan.

1.3 Business Context

Anyone could be a victim of the deepfake especially the public figures, actresses, comedians and entertainers. The system helps people in exposing the deepfake content whether they are included in it or not to avoid them from being manipulated or blackmailed. People will use the system to check whether certain videos are authentic or not by uploading them to the website.

2 Similar Systems

2.1 Academic

1. **Deressa Wodajo, Solomon Atnafu [1]** agreed that deep fakes causes a significant threat to everyone if it was used for harmful purposes as phishing, scam and identity theft which

reduces the trustworthiness of the public data. A convolutional vision transformer is used in this project to detect the deep fake. The adds-on of this project is that a CNN module is added to the VIT architecture as CNN extracts the features as (facial features in an image) and VIT takes those features as an input to categorize them to a specific class.the datasets used are FaceForensics++ Faceswap, FaceForensics++ deepfakedetaction, FaceForensics++ deepfake, FaceForensics++ faceshifter, FaceForensics++ neuraltextures. the results are An accuracy of 91.5% is achieved, AUC value of 0.91 and a loss value of 0.32 which indicates the difference between the predictions and the actual results. CNN and RNN had accuracy of 92.6% (validation) and 91.88% (testing). CViT had 87.25 (validation) and 91.5(testing). The face recognition library had made the best results rather than MTCNN and BlazeFace libraries as it had higher accuracy than them.

2. **Yuezun Li, Xin Yang, Pu Sun[2]** discussed about the AI-synthesized face-swapping videos, commonly known as DeepFakes, is an emerging problem threatening the trustworthiness of online information. It is needed to develop and evaluate DeepFake detection algorithms calls for large-scale datasets. dataset used is Celeb-DF dataset. the results are Celeb-DF is in general the most challenging to the current detection methods, and their overall performance on Celeb-DF is lowest across all datasets with an average AUC of 56.9 %. While FF-DF had the highest results across all methods with an average AUC of 82.3 %. while the method that had the highest performance across all datasets was DSP-FWA that had AUC of 87.4 % and the lowest performance goes to HeadPose method with an AUC of 58.7 %. The celeb-DF dataset have proven that it still needs improvement.

3. **Darius Afchar, Vincent Nozick, Junichi Yamagishi, Isao Echizen [3]** agreed that the huge use of digital images has been followed by a rise of methods to change image contents, using editing software like Photoshop for example. The field of digital image forensics research is dedicated to detecting fake images in order to regulate the circulation of such fake contents. providing two possible network architectures(Meso 4 and MesoInception 4) to detect such forgeries in an efficient way with a low computational cost. Dataset used are Deepfake dataset, Face2Face dataset. Both networks have reached close scores around 90 % considering each frame independently.Higher score is not expected as some images has facial extractions made with a very low resolution. It is observed a decline of scores at the strong video compression level. The image aggregation significantly enhanced both detection rate. It even rose greater than 98 % with the MesoInception, network on the Deepfake dataset. Note that on the Face2Face dataset, the score is the same for both networks but the misclassified videos are dissimilar.

4. **Xin Yang, Yuezun Li, Siwei Lyu [4]** agreed that Deep fake has a great impact on our environment nowadays, it's created by putting faces using deep neural networks into original images/videos. Together with additional forms of misinformation shared through the digital social network, digital impersonations were created by deepfake which have become a real problem with negative social influence. Accordingly, there is serious need for successful methods to detect Deep Fakes. Dataset used are UADFV and subset from the DARPA. The results are the SVM classifier reaches an AUROC of 0.89. This means that the difference between head poses evaluated from central region and whole face is a good feature to identify Deep Fake generated photos. DARPA GAN Challenge dataset, the AUROC of the SVM

classifier is 0.843. The reason is that the synthesized faces in the DARPA GAN challenges are mostly blurry, leading to a struggle to accurately predict facial landmark places, and as a result the head pose evaluation. They also calculated the performance using separate videos as unit of analysis for the UADFV dataset. This is achieved by taking the average of the classification prediction on frames over separate videos.

5. **Samay Pashine, Sagar Mandiya, Praveen Gupta, Rashid Sheikh[5]** agreed that people misused the advantage of the photoshop and advanced tools to swap w face of real person with a fake one and people get manipulated easily because they cant detect the deepfake. in this paper they contributed by comparing the 4 methods to select the best method. the dataset used were Celeb-DF, Celeb-DFv2, and DFDC. The best results were VGG-19 and Resnet-50 as they gave more accurate results than MesoNet. And by comparing VGG-19 and Resnet-50 the rate of loss and accuracy between the both of them were in the exact range and that's because the resnet-50 has more layers so it can detect much better than the VGG-19 in deepfake.

6. **Naciye Celebi, Qingzhong Liu, Muhammed Karatoprak [6]** they discussed that due to the crimes made and the numerous numbers of forged videos,for authentic evidence it is advised to use AI tools and digital forensics to detect deepfakes,for genuine evidence the court needs AI to detect the deepfakes for best results. In this paper they contributed by taking the DeepFake image of a person as an input which is called the target and another image is considered as an output which is called the source and the other person is replaced with the target's image as automatically the deepfake can map the features of the source to the target and by applying the post processing the image's output will guarantee a higher level of authenticity. The dataset used were Celeb-DF, FaceForensics++, DeepFake Detection, HOHA Dataset, UADFV, DF-TIMIT and VTD Dataset. The results were that CNN AND RNN: They gathered 300 videos of deepfake.Then, from HOHA dataset they chose randomly 300 videos. So the system's accuracy reached 95% or more.

7. **Luca Guarnera, Oliver Giudice, Cristina Nastasi, Sebastiano Battiato [7]** agreed that nowadays it became so easy to replace people's face with a fake one on images and videos and the modern techniques are not able to recognize the fakeness but with deep learning algorithms methods it could be detected. In this paper they contributed by using the preliminary idea that could analyze anomalies(exploiting noise, compression parameters, etc.) in the frequency domain. Anomalies can be detected by forensics experts and then a Fourier transform will be applied to it and that will make it clearer to detect the pattern of the deepfake creation made. The dataset used were CNN AND RNN. The results were that analyzing and tracing the anomalies could detect the deepfake images.

8. **Nicolò Bonettini, Edoardo Daniele Cannas DEIB, Politecnico di Milano, Sara Mandelli[8]** they tackled the issue of detecting the manipulated faces in videos aiming modern facial manipulation techniques. In specific, ensembling of the different trained CNN models, combining these networks (EfficientNetB4, attention layers, Siamese training) will lead to positive results of face manipulation detecting . The dataset used were FF++ and DFDC. The results were that the EfficientNetB4Att explain-ability:was an easy attention mechanism that allowed to spotlight the faces detailed parts like the facial features.On the other hand,

flat parts were too small so they were not useful for the network. however it was proven in multiple times that the deepfakes were most detected in the face facial features. Siamese features: TSNE algorithm was computed over small spaces to decide whether the produced outputs of the features were preferential for the task. And by starting from 20 videos of FF++ the resulted projection by the EfficientNetB4Att it was obviously seen how the frames of the video groups into smaller sub regions.

9. **Pratikkumar Prajapati, Chris Pollett[9]** agreed that fake media found on the internet can affect can spread propaganda and damage reputations of celebrities, therefore deep fake detection methods are produced to detect such fake content. The main contribution of this paper is that they added an MRI-GAN approach which is like an x-ray that shows the difference between the real and fake image if the image is fake. MRI-GAN training dataset, DFDC training set, Celeb DF v2 dataset were used in this paper. 50 % videos were taken from the DFDC dataset and all of the videos in the Celeb DF v2 dataset. We trained our MRI-GAN with batch sizes of 128. We had approximately 919, 590 DeepFake training and 1, 359, 717 real training samples. We had approximately 339, 930 DeepFake test and 229, 898 real test samples in our MRI-DF dataset.

10. **Chih-Chung Hsu, Chia-Yen Lee, Yi-Xiu Zhuang [10]** agreed that the nVidia-proposed gradually growing of GANs (PGGAN) has shown that realistic and high quality facial images can be easily synthesised. For example:it's used to make a new fake account on social media or to blackmail someone. This will make a lot of problems whether political problems , society or any commercial activities. Adding to that, sufficient images rigging detection way is needed. the dataset is 400198 training images and 5000 test images containing real and fake images sized of 64 × 64 and the batch size is 32. the results are made by different state of the art GANs.It's shown that the proposed DeepFD reach higher performance and is easily converged. In terms of accuracy, precision, and recall, the suggested DeepFD performed better than existing baseline techniques. The proposed DeepFD performs noticeably better than competing methods. Additionally, it has been demonstrated that the proposed DeepFD is more universal and efficient than others. contrariwise, it's found that the performance increase of the suggested strategy for the training dataset without LSGAN is noticeably superior due to the fact that the LSGAN produces images with fewer implausible details.For all types of fake photos, our DeepFD makes it simpler to extract the jointly discriminative feature, improving performance.

11. **Joel Frank, Thorsten Eisenhofer, Lea Schönherr, Asja Fischer[11]** agreed that fake news had become a real issue and Internet information campaigns can affect democratic processes developing automated detection methods is a very important task. Example for this:Ali Bongo, the president of Gabon: In late 2018, the president fell unwell and was out of the public eye for several months. As the populace got weary, the administration unveiled a video of the president, which was swiftly debunked as a complete fabrication. the dataset are from two places (Flickr-Faces-HQ (FFHQ) or from a set generated by StyleGAN. the results are In further detail, we demonstrate that frequency-representation-based classifiers provide greater accuracy while using noticeably fewer parameters. These classifiers are also more resistant to typical image disturbances. StyleGAN still has large coefficients at the upper and left sides of its spectra, but it seems to be able to mimic the spectrum of real photos better than the

other GANs. Accuracy: Ridge-Regression-Pixel 75.78 % Ridge-Regression-DCT 100.00 % better Ridge regression performed on data samples generated by StyleGAN for different upsampling techniques: Nearest Neighbor : • Ridge-Regression-Pixel: 74.77 % • Ridge-Regression-DCT: 98.24 % better Bilinear : • Ridge-Regression-Pixel: 62.13 % • Ridge-Regression-DCT: 85.96 % better Binomial : • Ridge-Regression-Pixel: 52.64 % • Ridge-Regression-DCT: % 84.20 % better The results of the source identification: Several methods are used like((KNN, KNN-DCT),(Eigenfaces, Eigenfaces-DCT),PRNU,(CNN,CNN-DCT),(CNN-PIXEL,CNN-DCT)) using LSUN and CELEBA and the best results were the CNN-DCT which had 99.64 % in LSUN , and the CNN was 99.91% in CELEBA. Results of common image disorder on LSUN bedrooms: The best results were the CNN-DCT not CNN-Pixel which were: BLUR: clean data :61.42% , perturbed data:93.61% Cropped: clean data: 83.52% , perturbed data:98.83% Compression: clean data: 71.86% , perturbed data:94.83% Noise: this was the only thing that's better in CNN-Pixel (clear data 59.51%), perturbed data: 89.56% Combined: clean data: 67.76% , perturbed data:92.17%.

12. **Vera Wesselkamp, Konrad Rieck, Daniel Arp [12]**It was proven that for spotting fake images, the frequency domain is extremely beneficial for that. Due to this action, many attacks have been able to be advanced in escaping the detection of generated images. a novel class of simple counterattacks is produced to overcome these limitations. the datasets are ProGAN, SNGAN, MMDGAN, CramerGAN. This action led to an extremely successful attack in opposition to Frank et al regression's model and CNN-based classifier. Therefore, the results supports that relying the two classifiers on the data which is stored the high-frequency bands for their decision is the solution. However, the spam provides only low success amounts compared to the detector of Joslin and Hao, proving that the method also takes into account low-frequency artefacts. The configuration, nevertheless, determines whether the peakextraction assault is successful or not; it performs almost flawlessly against ProGAN and SNGAN, which have important peaks throughout the spectrum. Even when attempting to test a regression model, the regression-weights approach seldom succeeds.

13. **Richard Durall, Margret Keuper, Franz-Josef Pfreundet[13]** fake digital contents have increased growing problem and spreading distrust in image content, leading to an urgent need for automated ways to detect these AI-generated fake images. dataset used MNIST dataset. SVM was used in order to classify the data and to produce a model. In addition to using Logistic Regression to make classification as well. K-means clustering was also used to train unlabeled data. SVM produced higher accuracy of 90 % while Logistic Regression produced a 81 % accuracy. Low resolution photos were harder to detect, but the system has efficient results with high and medium resolutions. The main method used is classical frequency domain analysis, then using a base classifier.

14. **Julian Fierrez,Aythami Morales, Javier Ortega-Garcia [14]** agreed that the free access to large-scale public databases,together with the fast progress of deep learning techniques,in particular Generative Adversarial Networks, have led to the generation of very realistic fake content with its corresponding implications towards society in this era of fake news and four types of facial manipulation are reviewed: i) entire face synthesis, ii) identity swap (DeepFakes), iii) attribute manipulation, and iv)expression swap. Methods are :Entire Face

Synthesis, this manipulation creates entire non-existent face images, usually through powerful GAN. Identity Swap: this manipulation consists of replacing the face of one person in a video with the face of another person. Dataset used was DFDC dataset. As for the classification, SVM was used to obtain high results. Different CNN architectures were used as VGG 16, VGG 19 and ResNet. Moreover attention mechanisms were used to enhance the process of facial detection. As for the future work , Face Synthesis it is intended to remove GAN fingerprints and add some noise patterns. In addition to Identity Swap, different metrics should be enhanced to obtain higher results. Finally, attribute manipulation needs to also remove GAN fingerprints and to observe the scarcity of datasets for researches. All of the mentioned methods obtained similar results.

15. **Andreas Rossler, Davide Cozzolino, Luisa Verdoliva, Christian Riess[15]** agreed that the fast progress in forged photos generation and manipulation has reached a point where obvious concerns were raised for the indications towards society Face2Face , Neural Textures and Face Swap are essential examples for facial manipulating at random level of compression and size to ensure that we can detect the fake facial images. To move the face from a source video to a target video, a graphics approach based is need which the face swap. Face2Face: is a facial reconstruction system that maintains the target person's identity while transferring the expressions of a source video to a target video. Neural-Textures demonstrates a rendering technique based on face reenactment using their Neural-Textures based technology. The neural texture of the target individual, comprising a rendering network, is learned using the original video data.

16. **Felix Juefei-Xu, Run Wang, Yihao Huang, Qing Guo, Lei Ma Yang Liu [16]** recognised that a wide range of both beneficial and harmful uses have been pushed by the production or modification of face appearance using deep generative techniques, are called"DeepFake." Spatial based Detection Recent research have used the most widely used methods for detecting DeepFakes in the spatial domain. -Detection using image forensics Recent research for DeepFake detection look on the differences in pixel-level using the classic forensics-based approaches. -Detection based on DNN By employing current or developing new DNN-based models and extracting spatial information to increase the efficacy and generalizability of detection, these techniques are entirely data-driven. -Detection based on biological signals Real still facial photographs and films are made with cameras and are more realistic than synthetically created phoney faces. we aim to study all the techniques possible to fight this uprising problem.the dataset used is MNIST database.

2.2 Business Applications

- **Video Authenticator tool[17]:** Is a tool developed by Microsoft in order to detect deepfake videos with high accuracy.

- **KaiCatch [18]:** 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 % . 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

Deepfake has risen in the past few years as many public figures had their faces switched with other people's faces so they were accused with things they didn't do which caused them serious problems in their lives and spreading bad reputation about them. Our challenge is to solve and detect the deepfakness that people can't detect with their naked eyes in the forged videos as numerous amount of people are getting tricked and manipulated with the unreal videos, and other people are being blackmailed and accused with crimes they have nothing to do with. A website will be developed to detect deepfakes using artificial intelligence.

3.2 System Overview

3.3 System Scope

The Deep Fake Detector aims to help the users not to get deceived by the fake content and the deep fakes which look exactly like the real people so the system will target the following :

- 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 to deepfake or not.
- 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:

- **User:** will upload the video he wants to check on the website and wait for output to be displayed.
- **WebServer which Google chrome website works on:** is the Webserver that is going to run the system.
- **Deepfake detection website:** the system will take the video that the user uploaded and sends it to the preprocessing stage so that the video is converted to frames/images.
- **Pre-processing:** this stage is responsible for converting the video to images by using a cascade classifier where faces are detected and are going to be sent to the classifier.
- **Deep learning classifier:** deep learning methods are going to take the image and classify whether it's deepfake or not according to the training model.
- **Image classified:** finally the image is classified and the output is sent back to the user on the user interface.

As shown in Figure 1 the following describes the system overview .The system is separated into two parts detection and generation. As for the 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. As for the 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.

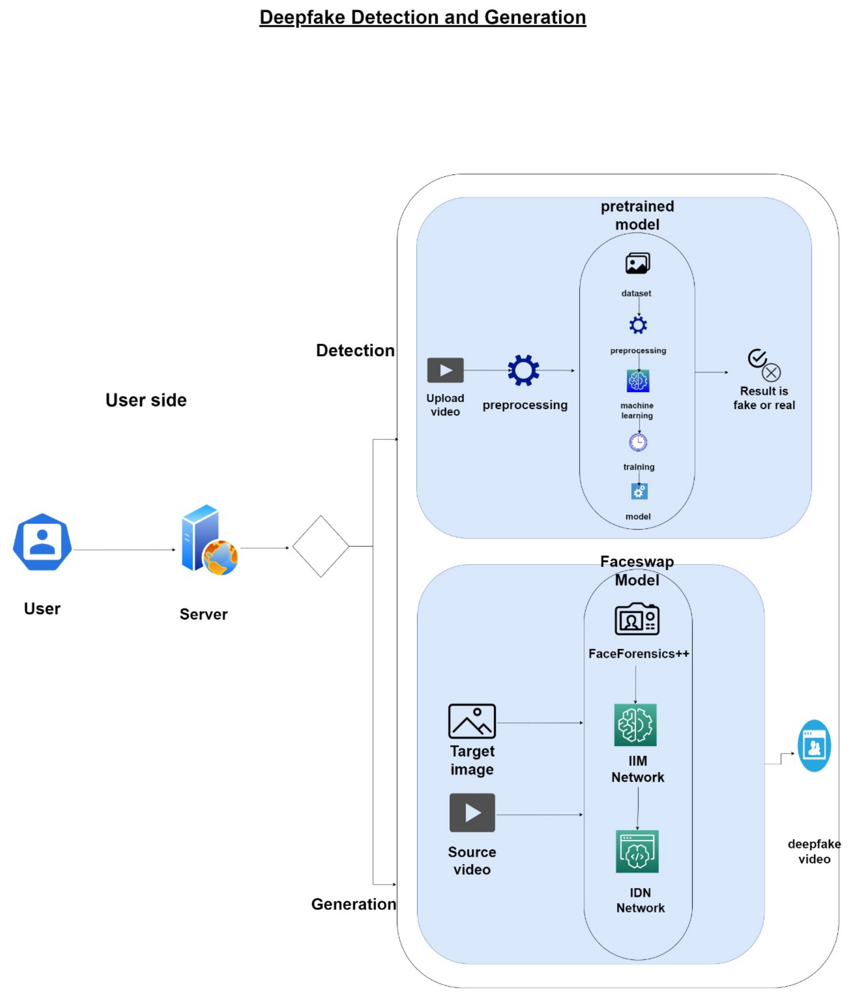


Figure 1: System Overview for deepfake detection and generation

3.5 Objectives

1. Provide a user-friendly and simple website.

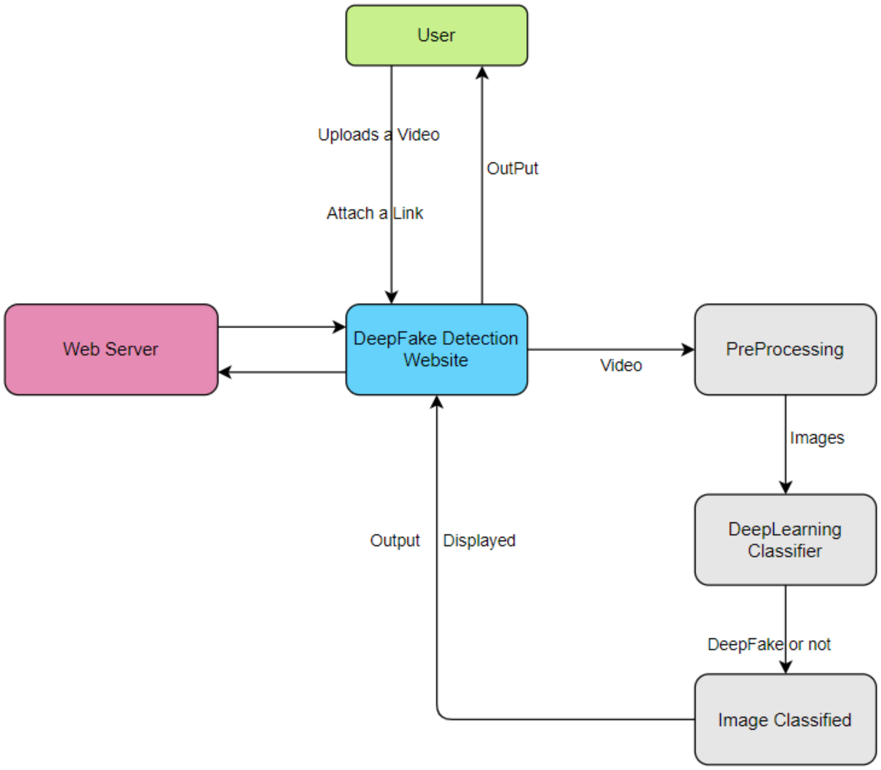


Figure 2: System Context for deepfake detection

2. Provide precise, more accurate deepfake detection
3. Result will be obtained based on the dataset entered

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

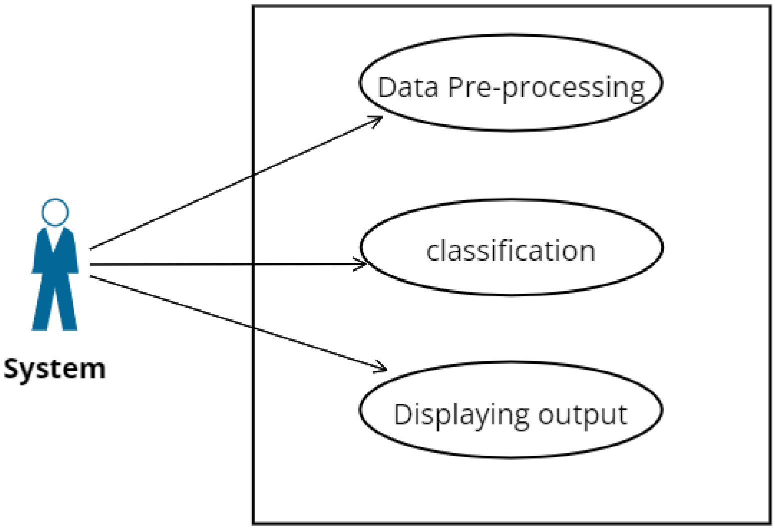


Figure 3: System use case diagram

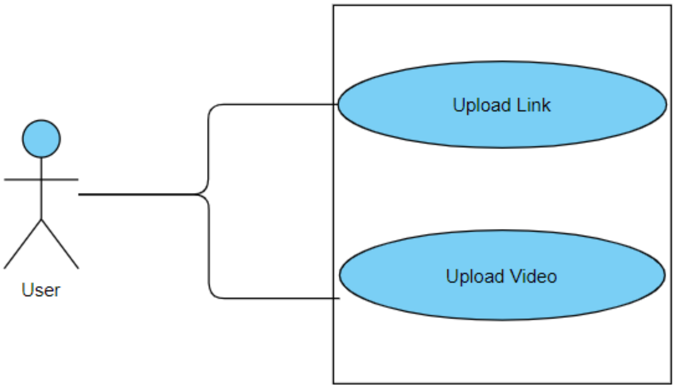


Figure 4: User use case diagram

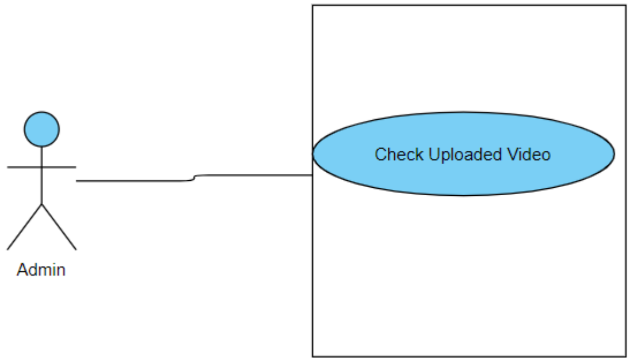


Figure 5: Admin use case diagram

1. FR01: The system shall accept a url and apply the detection process on the content of the url.

2. FR02: The system shall apply pre-processing techniques.

3. FR03: The system shall classify whether the input image is real or fake.

4. FR04: The system shall display the output to the user .

5. FR05: The user shall upload a link of the video that is going to be tested to the web application.

6. FR06: The user shall upload a video to the web application

7. FR07: The Admin shall check the uploaded videos on the web application

4.2 Detailed Functional Specification

Name	Preprocessing
Code	FR02
Priority	High
Critical	to detect faces from videos
Description	This function is responsible for extracting the faces from videos
Input	videos
Output	images
Pre-condition	reading the videos in order to extract faces
Post-condition	face images that are extracted are sent for classification
Dependency	dependant on uploading a video/ dataset being available
Risk	faces not being extracted correctly

Table 2: Preprocessing Function Description

Name	Classification
Code	FR03
Priority	High
Critical	it classifies the video
Description	it detects whether this recommended video fake or not.
Input	videos
Output	fake or not
Pre-condition	preprocessing then training
Post-condition	a message to the user to let him know if this video real or not (output).
Dependency	it depends on the preprocessing
Risk	inaccurate results

Table 3: Classification Function Description

Name	Upload Link
Code	FR05
Priority	High
Critical	To link the system to the video that will be tested
Description	A link for the video is uploaded by user from the internet to the system for it to be tested
Input	Video Link
Output	Results after preprocessing and classification whether the input is real or fake
Pre-condition	The link must be connected to a video
Post-condition	face images that are extracted are sent for classification
Dependency	dependant on uploading a video/ dataset being available
Risk	The link not being connected to a video

Table 4: Upload Link Function Description

Name	Upload Video
Code	FR06
Priority	High
Critical	To Upload the video that will be tested to the system
Description	The video is uploaded by user from the his PC or Mobile to the system for it to be tested
Input	Video
Output	Results after preprocessing and classification whether the input is real or fake
Pre-condition	The video be working and not corrupted
Post-condition	face images that are extracted are sent for classification
Dependency	dependant on uploading a video/ dataset being available
Risk	The Video data is corrupted or not working

Table 5: Upload Video Function 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 deepfake detection will run using a WebServer and the webapplication 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 internet connection of at least 1 MB upload speed and 2 MB download speed. is needed in order to process the videos and classify them accurately.

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 website should be easy to use even with a non-technical user and how the system can perform its functions without any errors.
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:** the website should be able to handle the amount of data added if it is increased and perform well. [19]
6. **Portability:** The website could be accessed from any device as long as its connected to the internet connection. [20]
7. **Maintainability:** The website should be flexible for any updates or expending without crashing.

7 Data Design

The dataset used in this project is the "deepfake detection challenge(DFDC)", which is composed of two folders, training and testing. each folder has 400 videos and they consist of fake and real videos. regarding the training side, prepossessing is formed to each video using the cascade classifier converting it to images and then the images are being cropped to focus on the face dimensions. Finally, the images are sent to the proposed architecture in order to preform the classification. A sample of the dataset is shown below. For the source of the data https://www.kaggle.com/competitions/deepfake-detection-challenge/data

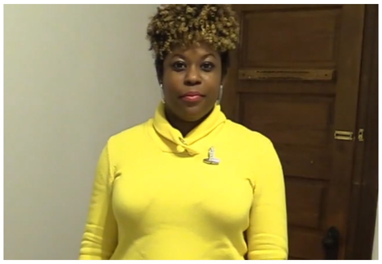


Figure 6: Real Image

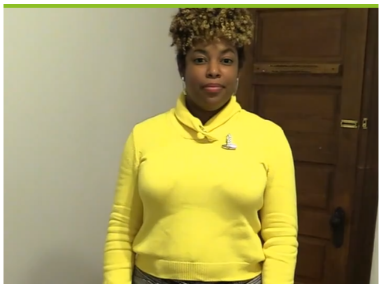


Figure 7: Deepfake Image

8 Preliminary Object-Oriented Domain Analysis

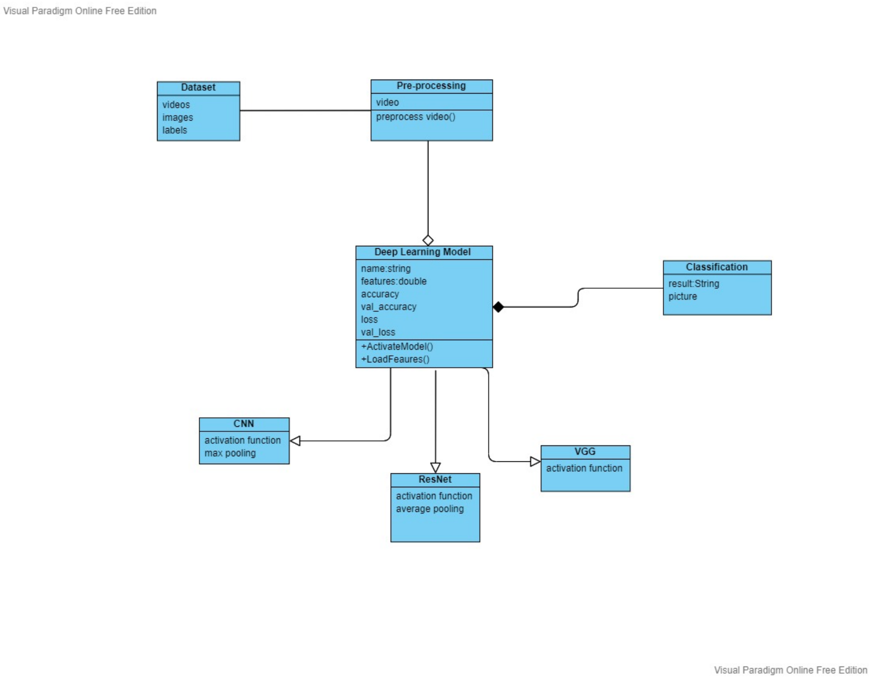


Figure 8: Class Diagram for deepfake detection system

9 Operational Scenarios

1. scenario 1: The user will access the website and upload a video to the video's website and the resulting output would be that it's a fake video or a real one.

2. scenario 2: The user will access the website and upload link to the video's website and the resulting output would be that it's a real video or a fake one.

10 Project Plan

Id	Task	Start Date	End date	Team Member
1	Reading research papers about deepfake	21/10/2022	28/10/2022	all
2	testing different codes of deepfake	1/11/2022	7/11/2022	all
3	choosing a suitable dataset for the code	10/11/2022	20/11/2022	all
4	searching for preprocessing algorithms to implement	22/11/2022	30/11/2022	all
5	working on the SRS document	3/12/2022	14/12/2022	all
6	fixing the errors in the code	16/12/2022	20/12/2022	all
7	Rehearsing for the discussion	21/12/2022	23/12/2022	all
8	SRS presentation	23/12/2022	25/12/2022	all
9	Trying different datasets	8/1/2022	11/1/2022	all
9	Starting the SDD	14/1/2022	20/1/2022	all
10	Implementing the website interface	21/1/2022	1/2/2022	all

11 Appendices

11.1 Definitions, Acronyms, Abbreviations

1. **CNN:** Convolutional Neural Network, it's a type of neural network algorithms used to extract images or objects (example:faces).

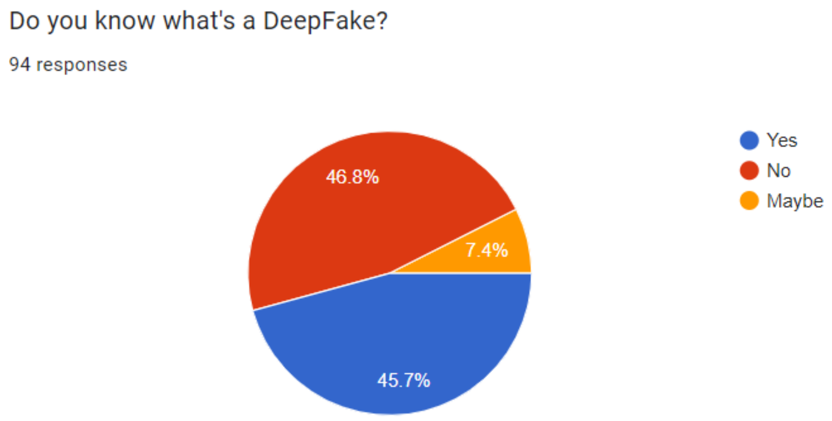
2. **MesoNet:**an algorithm used to automatically detect the editing of facial features in videos.

3. **VGG:** VGG stands for Visual Geometry Group; it is 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 of ground-breaking object recognition models. Developed as a deep neural network, the VGGNet also surpasses baselines on many tasks and datasets

4. **ResNet:**ResNet is an artificial neural network that introduced a so-called "identity shortcut connection," which allows the model to skip one or more layers. This approach makes it possible to train the network on thousands of layers without affecting performance.

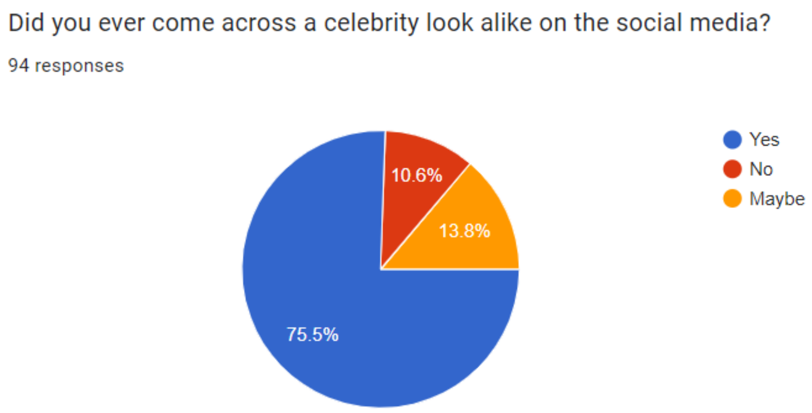
11.2 Supportive Documents

Survey:



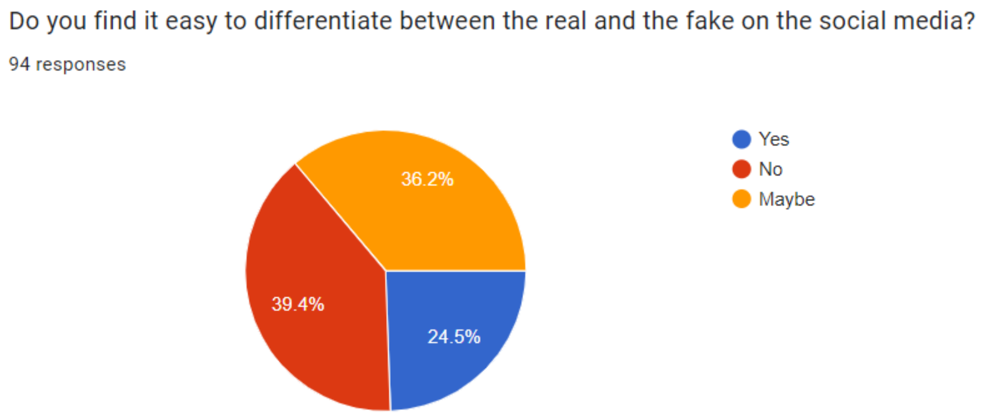
Do you know what's a DeepFake?

94 responses



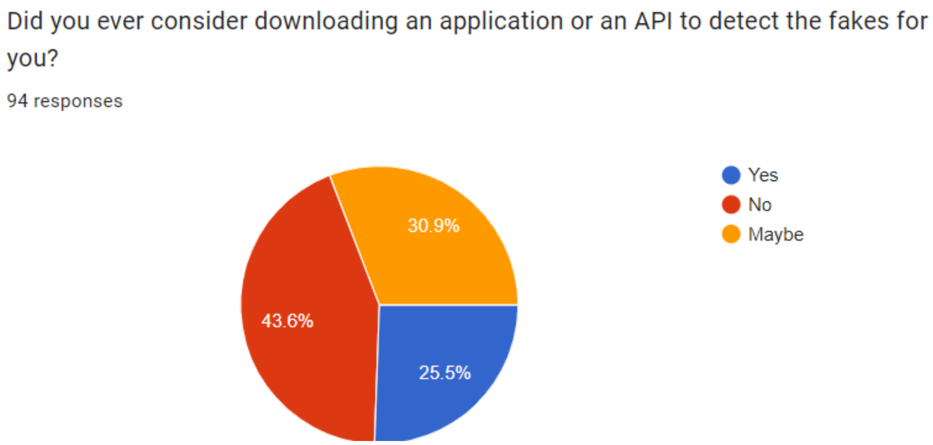
Did you ever come across a celebrity look alike on the social media?

94 responses



Do you find it easy to differentiate between the real and the fake on the social media?

94 responses



Did you ever consider downloading an application or an API to detect the fakes for you?

94 responses

References

[1] Deressa Wodajo and Solomon Atnafu. "Deepfake video detection using convolutional vision transformer". In: *arXiv preprint arXiv:2102.11126* (2021).

[2] Yuezun Li, Xin Yang, Pu Sun, et al. "Celeb-df: A large-scale challenging dataset for deepfake forensics". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, pp. 3207–3216.

[3] Darius Afchar, Vincent Nozick, Junichi Yamagishi, et al. "Mesonet: a compact facial video forgery detection network". In: *2018 IEEE international workshop on information forensics and security (WIFS)*. IEEE. 2018, pp. 1–7.

[4] Xin Yang, Yuezun Li, and Siwei Lyu. "Exposing deep fakes using inconsistent head poses". In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2019, pp. 8261–8265.

[5] Samay Pashine, Sagar Mandiya, Praveen Gupta, et al. "Deep Fake Detection: Survey of Facial Manipulation Detection Solutions". In: *arXiv preprint arXiv:2106.12605* (2021).

[6] Naciye Celebi, Qingzhong Liu, and Muhammed Karatoprak. "A Survey of Deep Fake Detection for Trial Courts". In: *arXiv preprint arXiv:2205.15792* (2022).

[7] Luca Guarnera, Oliver Giudice, Cristina Nastasi, et al. "Preliminary forensics analysis of deepfake images". In: *2020 AEIT international annual conference (AEIT)*. IEEE. 2020, pp. 1–6.

[8] Nicolo Bonettini, Edoardo Daniele Cannas, Sara Mandelli, et al. "Video face manipulation detection through ensemble of cnns". In: *2020 25th international conference on pattern recognition (ICPR)*. IEEE. 2021, pp. 5012–5019.

[9] Pratikkumar Prajapati and Chris Pollett. "MRI-GAN: A Generalized Approach to Detect DeepFakes using Perceptual Image Assessment". In: *arXiv preprint arXiv:2203.00108* (2022).

[10] Chih-Chung Hsu, Chia-Yen Lee, and Yi-Xiu Zhuang. "Learning to detect fake face images in the wild". In: *2018 international symposium on computer, consumer and control (IS3C)*. IEEE. 2018, pp. 388–391.

[11] Joel Frank, Thorsten Eisenhofer, Lea Schönherr, et al. “Leveraging frequency analysis for deep fake image recognition”. In: *International conference on machine learning*. PMLR. 2020, pp. 3247–3258.

[12] Vera Wesselkamp, Konrad Rieck, Daniel Arp, et al. “Misleading Deep-Fake Detection with GAN Fingerprints”. In: *arXiv preprint arXiv:2205.12543* (2022).

[13] Ricard Durall, Margret Keuper, Franz-Josef Pfreundt, et al. “Unmasking deepfakes with simple features”. In: *arXiv preprint arXiv:1911.00686* (2019).

[14] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, et al. “Deepfakes and beyond: A survey of face manipulation and fake detection”. In: *Information Fusion* 64 (2020), pp. 131–148.

[15] Andreas Rossler, Davide Cozzolino, Luisa Verdoliva, et al. “Faceforensics++: Learning to detect manipulated facial images”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2019, pp. 1–11.

[16] Felix Juefei-Xu, Run Wang, Yihao Huang, et al. “Countering malicious deepfakes: Survey, battleground, and horizon”. In: *International Journal of Computer Vision* (2022), pp. 1–57.

[17] 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.

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

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

[20] Lawrence Chung and Julio Cesar Sampaio do Prado Leite. “On non-functional requirements in software engineering”. In: *Conceptual modeling: Foundations and applications*. Springer, 2009, pp. 363–379.

Note that you should use 10 Reference Minimum (80% Academic)