"CYBER VICTIMIZATION: A STUDY ON DEEPFAKES AND EFFECTS OF ARTIFICIAL INTELLIGENCE"

Project Submitted to

The Department of Forensic Science

SRINIVASAN COLLEGE OF ARTS AND SCIENCE

Affiliated to

BHARATHIDASAN UNIVERSITY

TIRUCHIRAPALLI-24

In partial fulfilment of the University Regulations for the

Degree 'BACHELOR OF SCIENCE IN FORENSIC SCIENCE'

Submitted By

Mohammed Marzuk T M – CB21S610231

Vijayasarathy R – CB21S610246

Under the guidance of

Ms. Anisha M

Assistant Professor

Department of Forensic Science

THE DEPARTMENT OF FORENSIC SCIENCE

SRINIVASAN COLLEGE OF ARTS AND SCIENCE

PERAMBALUR – 621212

Batch 2021 - 2024

Ms. Anisha M

Assistant Professor

Department of forensic science

Srinivasan College of Arts and Science

Perambalur-621212

Date: 09/03/2024

CERTIFICATE

This is to certify that the project entitled **"Cyber Victimization: A Study on Deepfakes and Effects of Artificial Intelligence"** is a record of bonafide study and research certified out by **Mohammed Marzuk T M And Vijayasarathy R** under my supervision and guidance as the partial fulfilment of the university regulation for the course of B.Sc. in Forensic Science.

Date: 09/03/2024

Place: Perambalur

Ms. Anisha M

Assistant Professor

Department of forensic science

Dr. M Senthil Kumar

Head of the Department

Department of Forensic Science

Srinivasan College of Arts and Science, Perambalur

THE DEPARTMENT OF FORENSIC SCIENCE

SRINIVASAN COLLEGE OF ARTS AND SCIENCE

(AFFILIATED TO BHARATHIDASAN UNIVERSITY)

PERAMBALUR – 621212

CERTIFICATE

This is to certify that the project report entitled **"Cyber Victimization: A Study on Deepfakes and Effects of Artificial Intelligence"** to partial fulfilment for the degree of Bachelor of Science in Forensic Science submitted to Srinivasan college of Arts and Science. Affiliated to Bharathidasan University, Trichirapalli-620024. It is the bonafide record of original and independent work done by **Mohammed Marzuk T M (CB21S610231)** and **Vijayasarathy R (CB21S610246)** under my supervision and guidance. I further certify that it has not formed the basis for the award of any of my Degree/Diploma/Associateship/Fellowship or other similar titles under any University/Institutions in India or Abroad.

INTERNAL GUIDE

HEAD OF THE DEPARTMENT

Dr. M Senthil Kumar

Place: Perambalur

Date: 09/03/2024

SIGNATURE OF EXAMINERS

1.

2.

DECLARTATION

We, **Mohammed Marzuk T M** and **Vijayasarathy R** here by declare that this minor project entitled "**Cyber Victimization: A Study on Deepfakes and Effects of Artificial Intelligence"** is a bonafide and genuine research work carried out by us under the guidance of **Miss. Anisha M**. This project, in full or part, has not been submitted to this or any other university before, for the award or any degree.

Place: Perambalur

Date: 09/03/2024

Signature of the student

MOHAMMED MARZUK T M

CB21S610231

VIJAYASARATHY R

CB21S610246

ACKNOWLEDGEMENT

We would like to express my sincere gratitude to who worked for the flourishing completion of thiswork. First of all, we would like to thank almighty for the blessings he has showered upon us for the successful completion of our project.

We would like to take this opportunity to express our gratitude and appreciation to Srinivasan College of Arts & Science, affiliated to Bharathidasan University for having deputed us for this assignment. We want to express our heartfelt gratitude to the chancellor **Shri. A SRINIVASAN,**founder of Dhanalakshmi Srinivasan Group of Institution for his support and security. We wouldlike to thank our vice chancellor **Mr. S KATHIRAVAN** for giving us an opportunity to do this research.

We express my sincere gratitude to Srinivasan College of Arts and Science and our principal **Dr. N VETRIVELAN** for providing us this grateful opportunity to make this project successful and for giving support.

We would like to express heartfelt thanks to **Dr. M SENTHIL KUMAR**, Head of the Department for his advice and support to complete this project.

We would like to express my sincere gratitude to our mentor in this project, **Ms. ANISHA M**, Assistant Professor, Department of Forensic Science, SCAS for her guidance and support throughout this project which led us to complete this project successfully. We deeply thank her for valuable suggestion and encouragement.

We would also like to express our heartfelt gratitude to the staff members **Ms. MABEL CHANDRA, Ms. ARATHY RAJENDRAN,** and **Ms. SWETA BHARTI** of Forensic Science Department who gave us advice and support until the completion of this project.

We would like to thank **Mr. ANOOP ANIRUDHAN,** Assistant professor of Criminology Department for his immense contribution and support to complete this project. He guided us throughout the process of our project.

We would like to thank my friend **Mr. VISHNU SIDDHARTH** for providing ideas regarding our project and thanks to every respondent who have provided input data and their time for this project.

And finally, we would like to thank our parents, siblings and friends for sticking and cooperating with us to finalize the project successfully.

ABSTRACT

On the course of technological development, cybercrime developed with it. New features, new updates in the world of cyber space open doors and possibilities for many more cybercrimes along with initiating new ways to perform the existing one. **Artificial Intelligence** boomed from 1970 to 1980 and developed to the next level during the 2010s. With advancement in AI, the fear of AI taking control over us also grew, even though we didn't reach that level, we are already in the phase of people misusing AI. AI enables a person to edit, morph images and videos seamlessly without much effort. The only requirement for doing this is a bunch of inputs which can be photos, videos of a person. With matter of seconds a normal image of a person can be converted into **pornographi**c content and there are many free websites to do it, for example **Deepart.io.** Social media from being the hub to connect people across the world now becoming a free access point of porn. Twitter and Telegram for example don't have any restriction for sharing graphic contents which enables many to spread morphed and edited contents and thus increasing cyber victimization. Less awareness among the public on finding the difference between original and deepfake is the reason why this victimization is spreading and the lack of governmental actions towards these contemporary crimes make the situation worse. The solution to avoid these crimes and spread awareness lays within the problem itself, which is ironically Artificial Intelligence itself.

Keywords: Deepfake, Pornography, Artificial Intelligence, Cybercrime, Cyber Laws, Victimization.

TABLE OF CONTENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

INTRODUCTION

Deepfakes are hyper-realistic, superimposed, edited media content where one person's image or video is digitally edited over another media to make it seem like another person. As the algorithm has "learned" the face's features from various angles, and how it moves in different expressions, it able to replicate it in a way that follows the expressions. In earlier days of deepfakes, it was relatively easy to find the difference between the edited content and original content. But the advancement in the field of Artificial Intelligence made it difficult to find the morphed and edited contents. The term Deepfake originated on November 2nd, 2017, when a reddit user started a community on the title "Deepfake", under this community the members started to create and share the deepfake pornographic content of celebrity actresses[19]. Even though that community was banned for violation of guidelines by reddit in February 2018, about 90000 subscribers already downloaded the algorithm to create deep fakes and started spreading it[19]. The use of deepfake first started for advertising purposes by multinational companies but soon after its introduction, it was used to create illicit content. The creation and spread of pornographic deepfake content hiked in the year of 2019 when an opensource application was made available which made easier for common person with no such computer knowledge or editing skills to convert a normal content of someone into pornographic content[7]. Applications like Faceswap, deepfacelab, deepfakesweb, are some of the opensource applications that came during the year 2019 making it easy to create deepfakes for the public[10]. In 2019 Sensity.AI an open-source organization conducted research on the spread of deep fakes compared to its release in 2017, the report showed that the spread of deep fake online is rapid, and it had 100% increase in successive years which is 7,964 in the year 2018 to 14,678 in 2019 and in 2020 it became 52,000 constantly increasing[1]. As the technology advance over the years the realism of the deepfakes increased. With the ability to clone audio in the year 2020 with just 5 seconds of input audio increased the realism and nature of the deepfake as the cloned audio made it difficult to differentiate between original and fake[11]. During the same year, opensource application for mobile phones came in which increased the number of creators and deepfake pornographic contents as the population of mobile phone users is more. One main reason for the increase of deepfakes is the psychological kick that gave watching the pornographic content of their own favourite person or actors. Various psychological facts state that humans by nature like fantasies which give them the sense of accomplishment to experience things which cannot be done in real life, the sense of escapism. As deepfakes made the wildest of the fantasies come true to watch and experience it had a major impact on the

population wanting them to watch it and spread it. As the spread of deepfake increased the need to counter it and spread awareness about it also increased. In the year 2020 Microsoft first started to develop Deepfake Detection software[17]. Many open source and also advanced websites were created to detect the effect of Artificial Intelligence on a media content but only few made it free for public making it difficult for the public to access the sites and also the reliability of the sites was a question mark increasing the need of the government to step in.

ACT OF DECEPTION USING DEEPFAKES AND Its EFFECTS

Another reason why people tend to easily believe deepfakes is, in the core deepfakes are a form of deception where people purposely mislead another person. According to the deception detection literature people are not good at detecting deception when consuming online content and can easily acquire false beliefs[6]. The impact of deception caused by deepfakes is far greater than verbal deception as the reach of deepfakes in this internet-prone era is greater and faster. There are two sets of users of the internet, one who has enough knowledge about these fakes and the other who can't assess the authenticity of the content they consume. Sadly, the latter percentage is higher as even small children are now users of the internet along with illiterate people who can't understand the concept of deepfakes. As deep fakes use visual deception along with audio it makes it harder to believe that it is not true as there are too many stimuli and perceptions to be done before realizing making it easier for many to believe deep fakes as human visual receptors are more dominant than others and we rely more on that and thus when encountering something visually it is harder for humans to distinguish fake and real if it is done at perfection or close to perfection. The dominance of visual signals is called "The Colavita Visual Dominance Effect"[6]. Visual misleading generates greater misconception than verbal because of "Realism Heuristic", in which people tend to believe audiovisual content over verbal content because audiovisual content has more resemblance to the real world[6]. Video media are one last media that was thoroughly believed because humans tend to believe what they see but now when even those can be fake it creates a sense of disbelief among people raising the question of what to believe and what not to create a dilemma. Professor Don Fallis refers to this as "epistemic threat of deepfakes"[6]. This means the spread of deepfakes makes people doubt every bit of news they come across as video format is one of the most reliable formats of conveying news. The effect of deepfakes is not only as a potential format of

spreading porn but also as a way to convey false information in society and deceive people easily very much concerning and fearful of this technology which not going to have a downgrade but only upgrades in upcoming years as the technology develops. Deepfakes also can change one's memory of a targeted person in a deepfake. When people consume deepfake content of targeted politicians or any person, their attitude towards that particular person changes, the best example of this would be our Prime Minister Narendra Modi. His voice was taken as input to make videos of him singing various songs which tend to create an image of comical sense to people who consume the content thus undermining once reputation in the society. This gives a group of people to target someone and defame them using deepfakes which might be something funny or even in sexual way if it involves women.

APPLICATION OF DEEPFAKES

Every technology has both a good and dark side and it can't be decided that a technology is evil by only looking at one side. Even though deepfake has been recognized as a way to fool people and a method to create illicit and fake content, still there are many industries that use deepfake. Not everyone sees deepfake that way and some people use deepfake to create positive and quality content. Certainly, deepfakes can be used in the Art and Entertainment industry but it isn't necessary to stop there. With the help of Artificial Intelligence, Deepfake can be employed in many professions, reduce the workload and provide better results. It is necessary to go through both positive and negative applications of deepfake to better understand it, fight against it and to change the thinking that deepfake is only for illicit content creation.

- **ACTING AND ART INDUSTRY**

VFX (Visual Effects) and CGI (Computer Generated Imagery) are essential in the film industry to create authentic and mesmerizing worlds, backgrounds, and fictional characters but it requires a lot of work, cost and the process itself is time consuming. In film making, the sets and locations for several shoots have to be physically constructed but using deepfake, virtual sets can be created and altered. Deepfakes can bring back deceased actors, create a younger version of them and more[8]. Digital clones of actors can be made in situations where it isn't possible for the actor to appear. Companies like "PUMA, GAMBLE, PROCTER AND NIKE" have used digital clones of actors for advertising their products[7]. Disney can create content of

1024 x 1024 resolution with deepfake tech compared to that common methods only have a resolution quality of 256 x 256[9]. Creative Artist Agency has created a method to extract the features of an actor in a single day and create a digital clone which can be accessed by the actor[7].

- **VOICE OVER**

Voice over and voice manipulation technology can grow up to its full potential, thanks to deepfakes. Voice over may seem like a simple process, but it requires correct co-ordination to convey the things with lip syncing. Voice over is commonly known by movie buffs as dubbing process. Deepfake can create cloned voices with 5 seconds of input data[8]. Voice over can be done frame by frame to provide proper lip syncing and there is not any language barrier in voice over with the use of deepfake technology. Soccer player David Beckham made a way in awareness to malaria in which he spoke nine languages, it was made by Deepfake tech with proper lip syncing to his face and mouth[8].

- **EDUCATION**

Deepfake applications can be extended to the Education system too. There are many schools, colleges and universities which seek the method of using smart classes. Smart class uses video and audio representation for the students to understand better. Deepfake can be used similar to smart classes to make learning easier for students. When taking classes like history, deepfake can be used to create clones of people of the past and make conversations between them about the events of their own history. Teachers can come up with ideas of themselves to use deepfake to make teaching simple and easy.

- **SOCIAL MEDIA AND JOURNALISM**

Deepfake applications are being used by social media users to swap their faces with celebrity images and characters from films and televisions. It fulfils their thought of being like their favourite celebrity or character. Reporters who want to keep their identity to themselves when dealing with serious and dangerous news can choose to hide their identity using deepfake. Deepfake are being used by many to share fake news which makes people of the society question the authenticity of content in both social media and News channel.

- **POLITICS**

Deepfake can create fake news which might portrait and misrepresent politicians in a bad way. It can include politicians engaging in a normless or illegal act. It might show politicians making rude comments which they didn't. These courses of actions may be done to defame, insult or to influence the election campaign. In April 2018, Barack Obama's deepfake with Jordan Peele's voice was created and circulated to increase awareness on deepfake[7]. In June 2023, Deepfake of Donald Trump was used by Ron DeSantis's Presidential campaign to misrepresent Donald[9].

- **FRAUD**

Financial scam use has increased in recent times. Fraudsters use deepfake to scam people into fake financial schemes, investments and cryptocurrencies. These scams are usually done disguising themselves as celebrities to attain their victim's trust. Celebrities like Lee Hsien, Gayle King, Elon Musk, Tom Hanks, Jim Chalmers and Taylor Swift had their faced deepfake by scammers to attract potential victims[9]. Scammers use investment schemes and free product giveaways with just delivery charges but in reality, there won't be any product and scammers will continue to receive hidden payment using the victim's payment details.

- **DEEPFAKE PORNOGRAPHY AND Its HARMFULL EFFECTS**

Deep fakes have a significant impact on the creation of pornographic content at ease. Deepfake technology is being used to superimpose images of celebrities or any person's face with pornographic content. Just with access to some input images and knowledge of how to use certain Artificial Intelligence dependent websites or applications anyone can convert a normal image or video of someone into pornographic content. The reason that drives people to create such content is simply a psychological aspect of humans' loving fantasies[8]. People love experiencing things that cannot be attained in reality. This sense of fulfilment is given by deep fakes. With the use of many online free websites to advanced paid software, telegram bots deep fake pornography can be created with a click of a button. Women having broken up with relationships in the past are being threatened by pornography deepfake content with their faces particularly made by their past partners. Among all the deepfakes, about 96% of them are pornography content[6]. In Dec 2018, about 7,964 deepfake videos were found online. Deepfake content on the internet has risen 12 times that of the content in 2018 showing 95,820 deepfake

videos online in 2023[1]. South Korean actresses are essentially being targeted with having 53% deepfake content and India is in 6th place with having 2% deepfake of actresses[1].

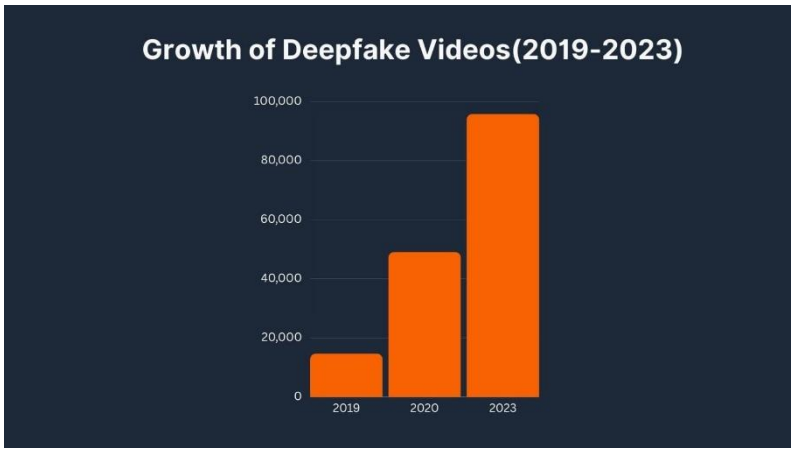

Fig 1.1: Statistics showing growth of deepfake pornography

[Source: Reddy, R. (2023, December 13). 24 Deepfake Statistics - Current trends, growth, and popularity. ContentDetector.AI. https://contentdetector.ai]

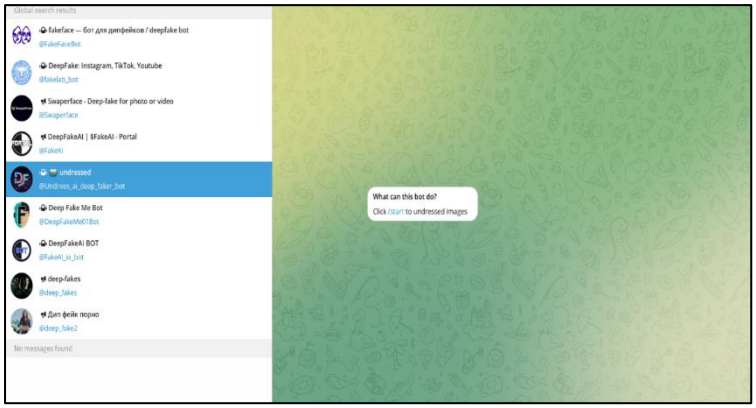

Fig 1.2: Telegram bot to create deepfake pornography

[Source: O'Rourke, D. (2023, June 26). Deepfake Telegram Bot: Disturbing exploitation. Deepfake AI - Ultimate Deepfake News. Deepfake Video, Deepfake Apps. https://deep-fake.net]

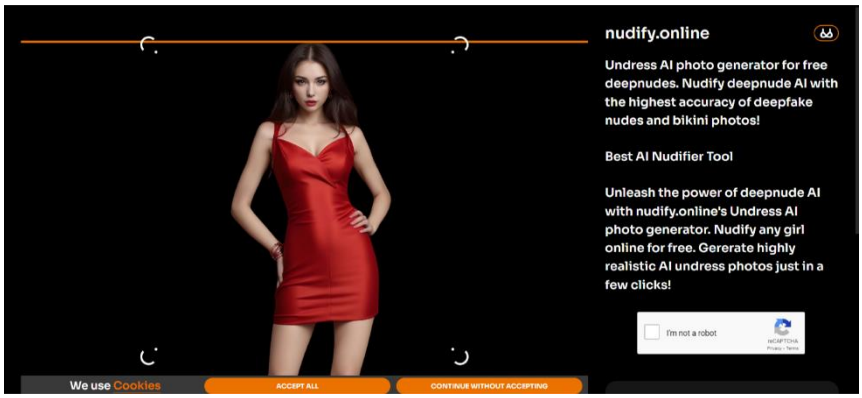

Fig 1.3: Online free site for creating Deep fake nude images

[Source: Heidi, A. (2024, February 15). Nudify online: Best AI Nudifier for free - CloudBooklet AI. Cloudbooklet. https://www.cloudbooklet.com]

There are many types of deepfake pornography which are

Table 1: TYPES OF DEEPFAKE

Type	**Description**
Face swap in image	It is when face of one person is swapped with nude image another person
Face swap in video	It is when face of one person is swapped with pornographic video.
AI Undressing	It is when the original image of a person is made nude using Artificial intelligence.
Lip-sync audiovisual manipulation	It is when both video and audio of a video is superimposed with someone else input and creating lip sync by manipulating the expressions.
Audio manipulation	It is when the audio of a person is converted as sexual audio or sounds

The scary thing about these applications and websites is, to process these images the company stores the images that are given as input and there is no security for the privacy of those images, they might be sold as data for money. The sad part is the victim in this act is not aware of any

of this and their image and video are stored and sold for various purposes and created as pornographic content online and uploaded to platforms like Pornhub and other adult video-sharing platforms for money. The one who is going to suffer in this illicit process is not the creator or owner of such websites or platforms where it is shared, the ultimate sufferer here is the innocent victim who has no idea of this happening and their face being seen by millions as porn without their consent. The psychological effect this has on such victim is very harsh, their reputation in society and among their peers is broken, and the fact many who are not able to differentiate that it is fake start to spread this content making these deep fakes viral and most seen form of porn in recent times.

METHODS ON HOW TO CREATE A DEEPFAKE

Photography and videography are two things that are considered to be a way to store people's memories and widely used by almost everyone in the 21st century. Those kinds of content can be easily altered and edited as per our needs and manipulation of photographic content isn't anything new, it has been in use since the 19th century itself[3]. Even though there are various Photoshop application and tools present, none of them can compete with Deepfake technology. Deepfake content can be created with more than one way.

Deepfake is a combination of both Machine learning and Deep learning. Machine learning is an algorithm that finds patterns and relations in a data to come up with a prediction, the user can intervene to check if the prediction is the right answer. AS for Deep learning, it is a subset of machine learning. Deep learning has its own neural network which can analyse the data and come up with its own answer, but these models need to be properly trained with different kinds of data with improve its accuracy over time.

- **ENCODER – DECODER PAIRS**

A Deepfake model has several layers for each of them has a specific task. One of the layers being the "ENCODER". Encoder is used to extract and store patterns and features of an original image to convert it into a latent image. Latent image simply refers to the patterns and features and not the whole image. Decoder's Job is to extract those latent features and reconstruct the image to the targeted image

In simple words, the encoder will extract important features of an image desired to be deep faked and will store it as a latent image and decoder will reprocess those features and paste it above the original image which will result it in a Deepfake Image.

- **FIRST ORDER MOTION MODEL**

Another approach on creating deepfake by replacing the encoder with motion model. This Artificial Intelligence focus and captures unique features like eyes, eyebrows, hair, face position and expressions. These are taken from hours of input data and analysed which can be superimposed on another face. In Films, these are widely used to create characters like monsters, angels and so. For a video, this process needs to do by each frame by frame to acquire best and realistic results of the content. One of the popular examples of this could the "AVATAR" movie in which Sam Worthington face is superimposed to his avatar character "Jake Sully"[6].

Fig 1.4: First order motion capture

[Source: Angkana. (2022, January 24). 4 Motion capture software. Terrestrial. https://goterrestrial.com]

- **GENERATIVE ADVERSARIAL NETWORK (GAN)**

Generative Adversarial Network has two neural networks which has a compete nature against each other. GAN's first neural network is called Generative neural network. This neural network processes lot of inputs and then create an image with their characteristics or entirely a new one. The second neural network is called Discriminative Classifier which analyses the image sent by the first neural network because it could be one of the real input sample or the deepfake image. Both these neural networks compete to create and identify whether the image

is real or fake. This process automatically increases the ability of the system in creating an image that is impossible to differentiate as a deepfake by people.

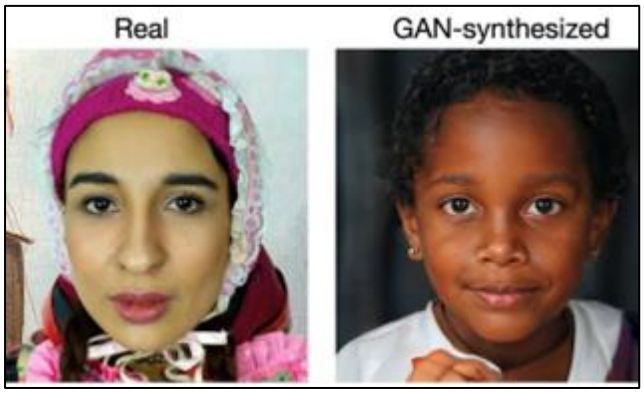

Fig 1.5: GAN-Synthesized image

[Source: Reichert, C. (2021, March 12). Deepfakes can be detected by analysing light reflections in eyes, scientists say. CNET. https://www.cnet.com]

CASE STUDY 1

PS Radhakrishnan is a 73-year-old retired person who used to work in Central Government Firm. Radhakrishnan was living in Kozhikode, Kerala. On July 9,2023, Radhakrishnan first happened to receive an anonymous call from a number he didn't recognise so he simply ignored it. After some time, He found messages in WhatsApp from a person calling himself as Radhakrishnan's former colleague who worked at Coal India Ltd. The WhatsApp profile had Radhakrishnan colleague's image and Radhakrishnan knew him for they worked for four decades together. The person sent his family pictures, talked about their other colleagues and even Radhakrishnan about his daughter. One day, the person suddenly voice called saying that he was in Dubai Airport to get to India. He had convinced Radhakrishnan that his sister law is in hospital in Mumbai, and he requires an amount of Rs 40,000 for her treatment. The person even made a video call that lasted for 25 seconds in which his face was very clear with proper lip and eyes movement. Radhakrishnan feeling pity to his friend sent the money to UPI Account referred by the person but after a few hours, the person contacted again saying he need another Rs 35,000. Radhakrishnan rejected it saying he doesn't sufficient balance. Radhakrishnan felt a doubt and he called the other number which he had of his friend and asked about this matter,

but his friend replied that he didn't make any such call or ask for money. Realizing his mistake, Radhakrishnan filed a Complaint. Police found that the video call was made using deepfake technology and the money was traced to an account in Maharashtra which had more funds probably acquired by scams.[20]

CASE STUDY 2

A Famous Indian Actress Rasmika Mandanna's face was deepfake and released in social media. It went viral during the time period of November 2023. Rashmika filed a complaint addressing that it was not her in that video. The Original video was of a British-Indian influencer named Zara Patel dressed in a Black coloured workout outfit, walking up and entering into an elevator. This video was doctored using deepfake technology to superimpose Rashmika's face onto Zara's face. This Deepfake incident raised concerns of civilians, celebrities and Indian Prime Minister has made a comment on the misuse of Artificial Intelligence. The Accused had deleted his accounts and its details, but the details provided by Social Media Application "META", they had apprehended four suspects. After investigation and analysis of data provided by META, Police has identified the culprit who created the deepfake. A Resident of Andhra Pradesh named Naveen, an Engineering graduate from a college in Chennai, Tamil Nadu had used to run fan pages for several actresses. He allegedly created the deepfake. When it went viral, he got scared and deleted all of his accounts and its contents.[16]

Fig 1.6: Rashmika's Face Deep faked with Zara

[Source: Novak, M. (2023, November 5). Viral video of actress Rashmika Mandanna actually AI deepfake. Forbes. https://www.forbes.com]

Deepfake by Google colab models

This project focuses mainly on spreading awareness regarding how easy it is to create Deepfakes and to share them in cyberspace which results in wide range spread of these fake contents. Deepfake creation models in Google Collaboratory were used to create some of the Deepfakes with basic inputs.

Furkan Gozukara's One click deepfake for free model was used to create the following deepfakes to establish how easy to make deepfakes online for free.

Inputs 1

Fig 1.7: Actor Vijay

[Source: Sahas. (2023, July 29). Thalapathy Vijay meets Vijay Makkal Iyakkam's Office-Bearers in Tamil Nadu. News18. https://www.news18.com]

Result 1

Inputs 2

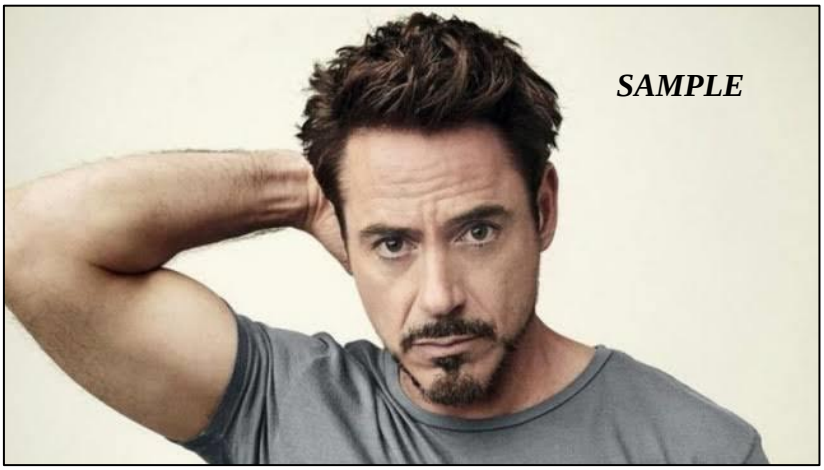

Fig1. 8: Actor Robert Downey Junior

[Source: Pti. (2016, August 16). Robert Downey Jr to star on HBO drama. The Times of India. https://timesofindia.indiatimes.com]

Result 2

CHAPTER 2

AIM AND OBJECTIVES

AIM OF THE STUDY

- To Conduct a study on Awareness of Deepfakes and Effects of Artificial intelligence

OBJECTIVES

- To study how many are aware of the term deepfake and its consequences
- To spread awareness regarding the effects of Artificial intelligence producing deepfakes
- To suggest ways to differentiate between authentic image or video and deepfake
- To suggest potential remedy or actions which can be done by government

CHAPTER 3

REVIEW OF LITERATURE

REVIEW OF LITERATURE

Ajder *et.al* (2019) studied about the "THE STATE OF DEEPFAKES: LANDSCAPE, THREATS, AND IMPACT", and found that among the 4 dedicated Deepfake pornography websites, the total views of those videos reached up to 134,364,438. Eight of the ten pornography mainstream website use Deepfake pornography content in their sites. Most of the victims of those pornography Deepfake are from the entertainment industry, primarily from western countries and south Korean celebrities.

Semwal (2020) has published a study on the name "Deepfakes, Artificial Intelligence and Eco Species". In his study, he mentions several laws regarding deepfake like Malicious Deepfake Prohibition Act 2018 of United States which will consider creating and sharing of illicit deepfakes as a criminal offence. He also mentions Texas and California passing a law to prevent sharing of deepfake to influence elections. He also mentions that India recognizes the risk after pornographic video of a journalist went viral but still hasn't come up with any laws.

Katarya and Lal (2020) conducted a Study on Combating Emerging Threat of Deepfake Weaponization. In this study, they used some of the major deepfake detection methods such as CGFace model, Eye blinking, Face-warping artifacts and more. They reached a conclusion that out of 7 model they have examined, SST Net: Spatial, Temporal and Steganalysis method has better accuracy in detecting deepfake ranging from 90% to 95%. They also suggested that organisations can create digital stamps to differentiate authentic and fake content.

Mahmud and Sharmin (2021) published a Review paper named "Deep Insights of Deepfake Technology: A Review". They mentioned that people lost their trust on online content due to deepfake. Anyone whose got access to a high-end computer system can create a deepfake without any prior knowledge on technology. Internet has made things easy to share illicit deepfakes. The study also mentions that just as deepfake are growing, the ways and methods to detecting them have also improved.

Neekhara *et.al* (2021) conducted a study "Adversarial threats to deepfake detection: A practical perspective". In this study, Adversarial attacks on Deepfake detectors are analysed in a black box to find the tolerance and ability of the Deepfake detectors to fight the attacks and provide correct results. Those adversarial attacks easily fooled the detection system which shows the need for stronger Deepfake detection tools.

Nyugen *et.al* (2022) conducted a Deepfake survey which concludes that deepfake doesn't need to research a large audience even if the deepfake manage to reach its own kind of target audience, that itself will be more than enough to sabotage a victim. The paper also predicts that a wrong deepfake can cause distress, hate, political tension and it also mentions that even though there are lot of deepfake content available, there is much benchmark content to train deepfake detection tools.

Fido *et.al* (2022) Conducted a study named "Celebrity status, sex, and variation in psychopathy predicts judgements of and Proclivity to generate and distribute deepfake pornography" and they concluded Victim blame was far less when the victim is a female compared to the victim being a male. Men blamed celebrity's actions for being a victim more than a non celebrity. Women sees deepfake pornography as greater criminal offence than men. Men sees celebrity deepfake as less criminal offence than non-celebrity deepfake but women see them as a criminal offence rather than being focused on the part whether they are a celebrity or not.

Rahman *et.al* (2022) studied about "A Qualitative Survey on Deep Learning Based Deep Fake Video Creation and Detection Method", and revealed that Attack in Deepfake detection tool makes it harder to provide accurate results. Social media and other platforms should develop detection and monitoring system to find and remove any Deepfake content. If source of the Deepfake is found, Law should be followed by the organisation which posted the Deepfake and delete those content immediately to reduce the damages to the victims.

CHAPTER 4

MATERIALS AND METHODOLOGY

MATERIALS

- Google form
- Cover letter
- Samples from the age 18 to 26
- Pie Charts were taken from Google form

METHODOLOGY

SAMPLE

Samples were taken from 150 respondents of age group between 18 – 26. All of the samples were taken from the Srinivasan College of Arts and Science, Perambalur, Tamil Nadu, India. The Pie chart data was taken from the Google form itself. The data was collected during the months of January – February

PRIMARY DATA

The data was collected from the respondents by circulating a questionnaire form with a set of questions related to the topic of this study.

The questions present in the questionnaire includes:

1. Have you heard of the term "deep fake" before?
2. Have you ever encountered a deep fake video or image on social media?
3. How concerned are you about the potential misuse of deep fake technology?
4. Do you believe deep fakes pose a threat to political elections and public figures?
5. Should there be stricter regulations on the creation and distribution of deep fake content?
6. How do you think deep fake technology can be effectively countered or detected?
7. Would you support the use of AI-driven tools to identify and flag deep fake content?
8. Do you think deep fakes have the potential to undermine trust in media and journalism?
9. How often do you verify the authenticity of the content you consume online?

10. Do you consider spreading creating and spreading deepfakes as sexual offense?
11. Would you be willing to take steps to educate yourself and others about the risks associated with deep fakes?
12. Are you aware of the procedures to be followed if you ever become a victim of deepfake?
13. In your opinion, should government create free sites to help deepfake victims?

SECONDARY DATA

Digital media (Google Form) was used to collect the secondary data.

CHAPTER 5

RESULTS AND DISCUSSION

RESULTS AND DISCUSSION

1. Have you heard of the term "deep fake" before?

Deepfake crimes have already started happening in India and many famous celebrities like journalists, actresses and even police officer's faces have been deep faked to fool. Still among the respondents of 150 who participated in our questionnaire, 26.7% have never even heard the term "DEEPFAKE" before.

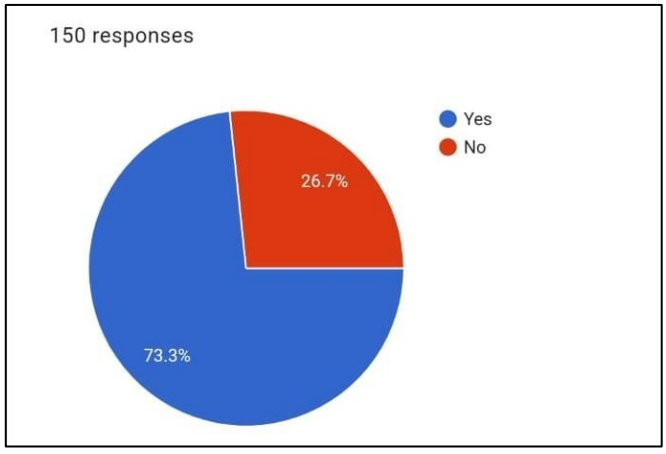

Fig 5.1: Heard of term "Deepfake"

2. Have you ever encountered a deep fake video or image on social media?

42% of the respondents has said that they have never seen a deepfake content. It is entirely possible that they have seen deepfake but was unable to differentiate between real and fake.

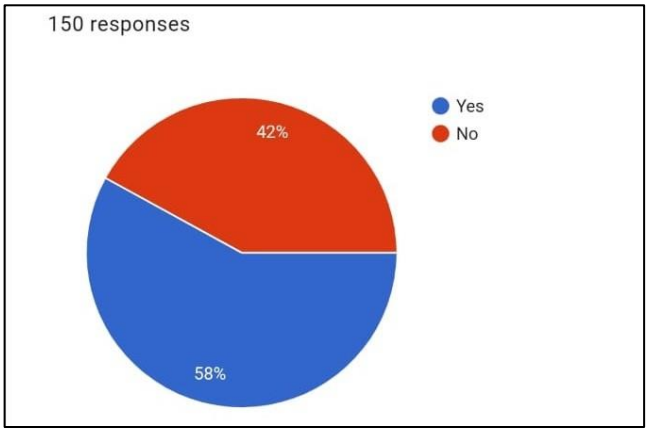

Fig 5.2: Encountered a deepfake in social media

3. How concerned are you about the potential misuse of deep fake technology?

56.7% of the respondents are said to be deeply considered that deepfake could be misused and even though 43.3% of the respondents says they are not that much concerned regarding, there is a chance that they haven't to realize it's true horror capabilities.

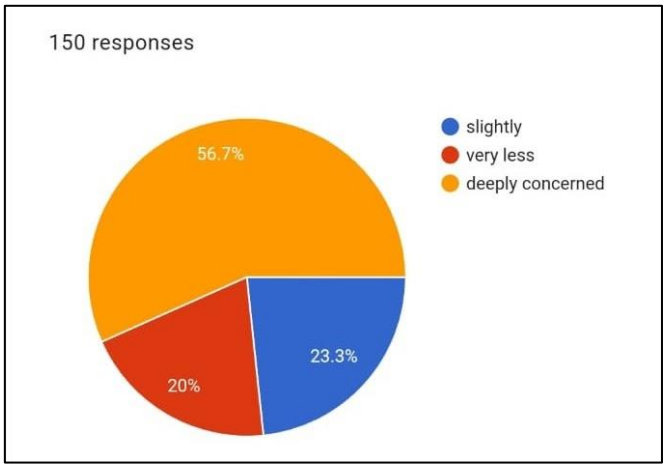

Fig 5.3: Concernment regarding misuse of deepfake

4. Do you believe deep fakes pose a threat to political elections and public figures?

Among the respondents, 48% of them thinks that deepfake content has the ability to influence elections and pose as a threat. There have been events regarding deepfake of many famous politicians like Obama.

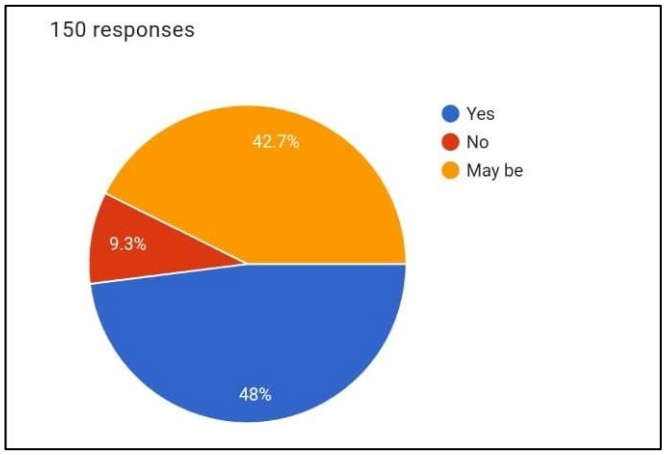

Fig 5.4: Level of threat to elections and politicians by deepfake

5. Should there be stricter regulations on the creation and distribution of deep fake content?

76% of the respondents believes that there should be strict regulations on the creation and sharing of a deepfake content

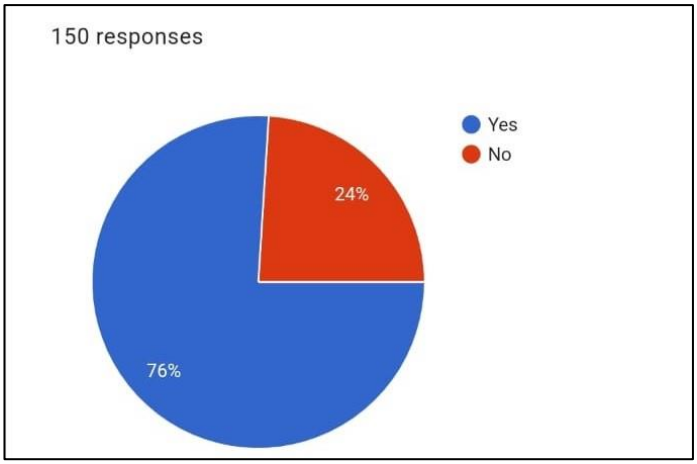

Fig 5.5: Need of regulations to create and share deepfake

6. How do you think deep fake technology can be effectively countered or detected?

After addressing a problem, the next step is probably finding a solution. Among the respondents, 56% of them thinks deepfake can be easily countered if the social media is to make guidelines that can restrict illicit deepfakes and 33.3% of the population says that it's government's responsibility to spread awareness among the people regarding such problems.

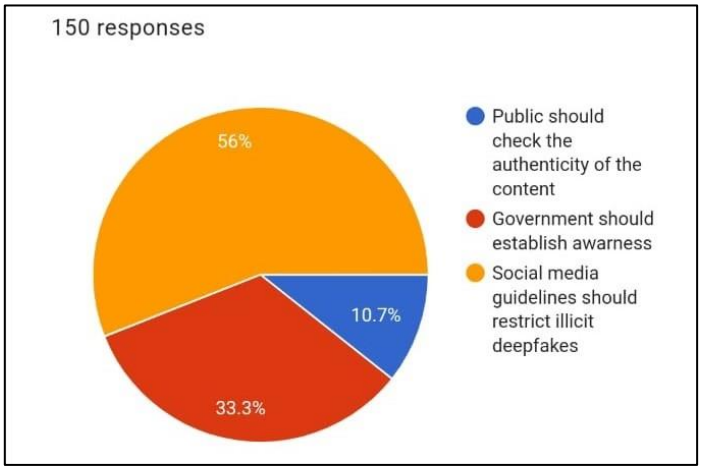

Fig 5.6: Methods to counter deepfake

7. Would you support the use of AI-driven tools to identify and flag deep fake content?

Deepfake are made by AI so it might be possible to detect them with the same, and 68% of the respondents are ready to support use of Artificial Intelligence to detect deepfake.

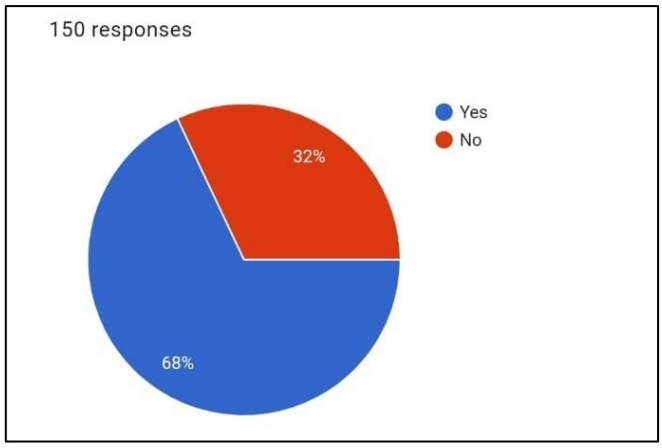

Fig 5.7: Use of AI to identify deepfake

8. Do you think deep fakes have the potential to undermine trust in media and journalism?

Media is one of the primary pieces of information sharing network and 47.3% of our study's respondents thinks that deepfake can undermine the trust of media and journalism due to spreading of fake news.

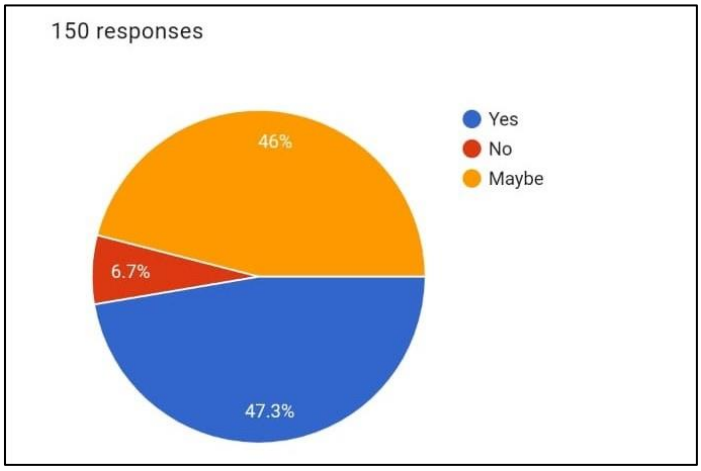

Fig 5.8: Loss of trust in media by deepfake

9. How often do you verify the authenticity of the content you consume online?

The content in our cyberspace can be both and fake but it is not possible to verify its authenticity every time. About 28% of the respondents doesn't usually verify the online content they consume.

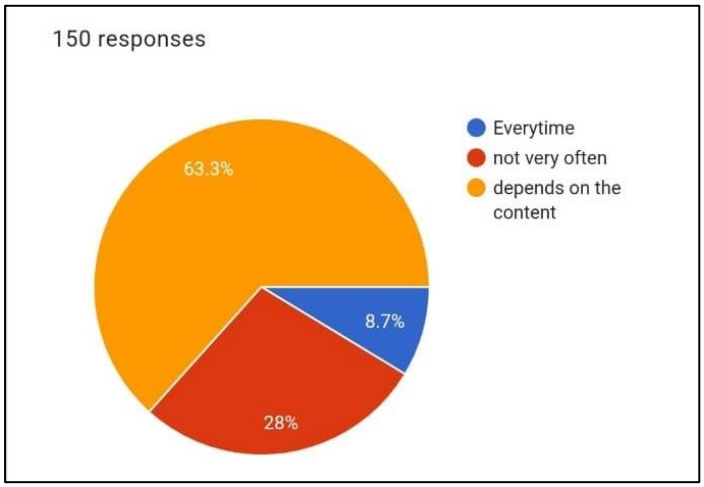

Fig 5.9: Verification of online content

10. Do you consider spreading creating and spreading deepfakes as sexual offense?

Pornography deepfake can create a terrific impact on its victims. About 56% of our respondents considers deepfake as a sexual offence.

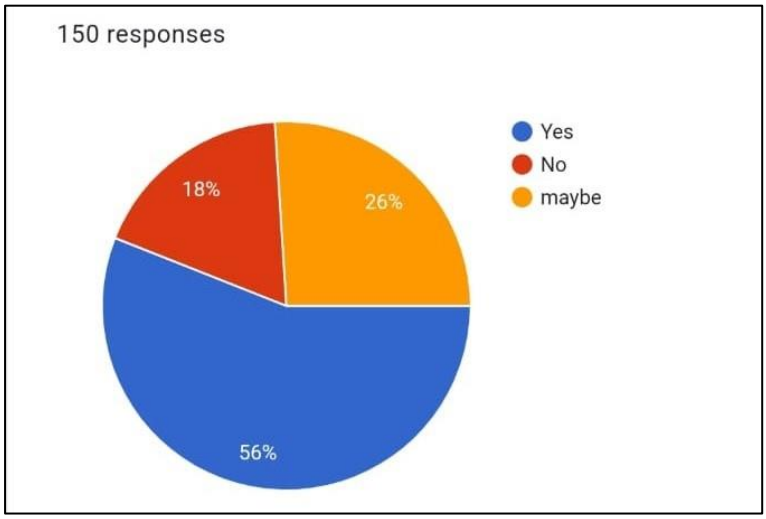

Fig 5.10: Deepfake as a sexual offence

11. Would you be willing to take steps to educate yourself and others about the risks associated with deep fakes?

Society can only tackle such problems by spreading awareness, 88.7% of the respondents are willing to learn and teach others regarding ill effects of deepfakes.

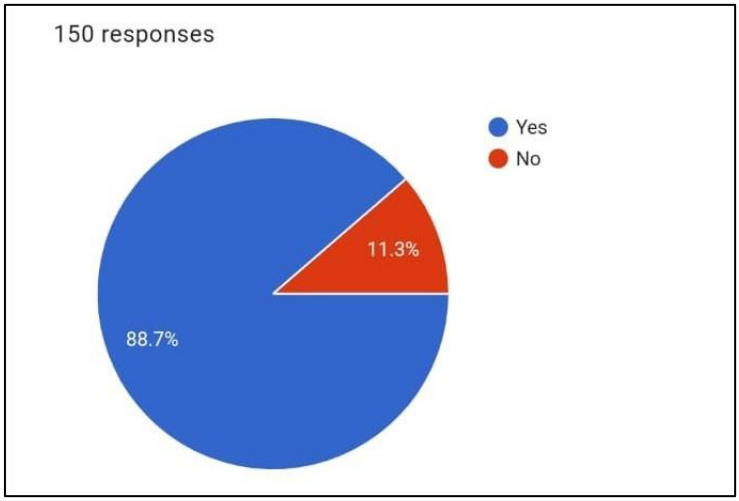

Fig 5.11: Educate yourself on deepfake

12. Are you aware of the procedures to be followed if you ever become a victim of deepfake?

Among the respondents, 58% of them doesn't the procedures that are necessary to be followed if they are to become a victim.

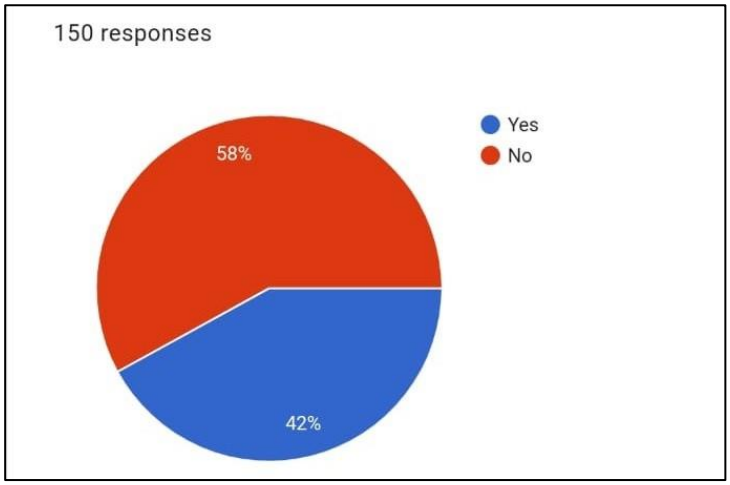

Fig 5.12: Knowledge of procedure to follow by victim

13. In your opinion, should government create free sites to help deepfake victims?

No matter how many solutions we try to muster up by ourselves, Government's help is necessary so 90.7% of the respondents says that government should come up with free websites to help deepfake victims.

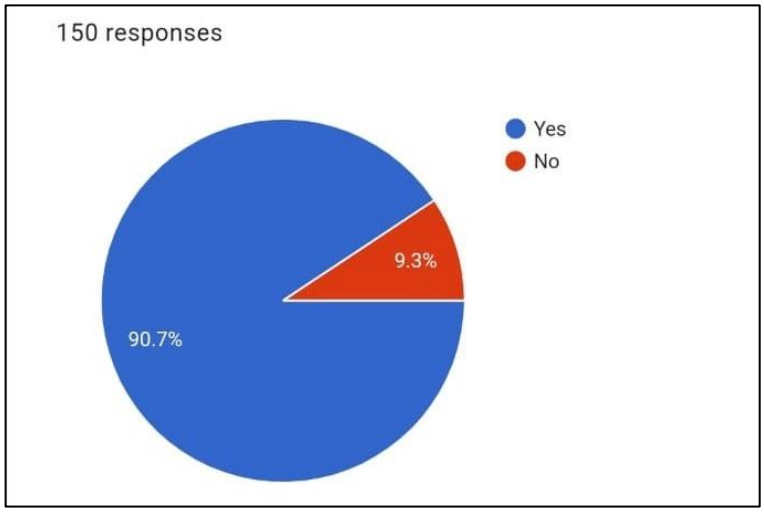

Fig 5.13: Need of Government's help

CHAPTER 6

CONCLUSION

CONCLUSION

Even though it has a lot of benefits in various sectors, deep fake's threat lies in how it is used. In the wrong hands with easy accessibility to free websites and telegram bots to create such contents deep fake becomes a threat to this society. In country like India where nearly 48.7% of the population have access to internet deep fakes have high influence in society to convey wrong information or to defame others especially women who can be subjected to be a victim of illicit creations using artificial intelligence. Steps taken by the government to address this rising problem are very less while the number of deep fakes steadily increasing since 2017. There are several ways for a common person to distinguish between fake and original media content like observing the sync between audio and video, face alignment, low quality, colour tone difference between face and body along with these factors various free sites can be used to find the influence of AI in a content. Awareness about this is very less among the public which leads to misconception of various content influenced by AI making it easy to spread fake news and content. Educating people with this new age crime and how to counter it by using their own phone will be highly effective. Government creating free site to raise complaint and finding difference between fake and real content just like in paid software encourage people to learn and counter the crime easily making difference in the modern society where even the common form of news can be manipulated to propose or spread false agenda or hatred over an individual.

CHAPTER 7

SUGGESTIONS AND RECOMMENDATIONS

SUGGESTIONS AND RECOMMENDATIONS

In the previous topics we mentioned that the root cause and mode of operation of creation of deep fakes is development in Artificial Intelligence. The irony is, to counter the spread of deepfakes and to detect what is real and what is fake, the same Artificial Intelligence can be used. As the intensity of usage of deep fakes increased in a negative manner, the need to rectify or defend us from the effect of it also increased. There are lots of ways to detect whether a content is fake or true which can be used by common people to crosscheck the content they consume on a day-to-day basis.

WAYS TO DETECT DEEP FAKE WITHOUT ANY APPLICATIONS

Even though usage of artificial intelligence to detect the influence of AI in a content is more accurate for detection of deep fake, it is not practically possible to download the content and run it on a software to verify in these modern days where people consume more than 1000 content online. So, to avoid confusion whether a content is original or fake there are various factors which can be observed to determine the originality of the content.

1. One thing which is common in the majority of deep fakes is the reduction of quality of the video or image on certain areas. Generally, when the media is being processed to create deep fakes the pixels which overlay on the top of the original pixels won't coincide perfectly causing the lack of quality which can be observed on specific areas of the media.
2. Along with the reduction of quality another common thing which can be observed is lack of originality. Deep fake image or video does not appear to be original if close attention is paid, in case of video there are various factors which can indicate that it is fake. Deep fake videos generally won't have perfect lip-sync and alignment of head and body or vice versa. The part which is being deep faked appears wavy and inconsistent. In the case of deep faked images, the colour tone variation can be seen between the original side of the content and the part where artificial intelligence is applied. The neck and face won't align perfectly.

3. Pay attention to the face, because most of the deep fakes are facial transformation. Any appearance of smooth or wrinkled skin which seems odd compared to other parts of the media can be a sign of deep fake.
4. Observe eyes and eyebrows for any sign of shadows or smudges because deep fakes often fail to replicate natural movements.
5. Pay attention to facial hairs like beard, moustache, whether the hair looks real or align which the facial design of the body because deep fakes often add or remove hair but fail to make it natural.

But these factors are not 100% accurate. These can be used to raise the benefit of the doubt so that usage of websites or software can be done which can confirm the doubt.

- **SOFTWARES/WEBSITES FOR DEEP FAKE DETECTION**

Factors such as inconsistent background, lightings and hair features can also be a result of a normal photo edit app so those factors don't necessarily prove that an image or video is manipulated with deepfake technology, in order to properly analyse and find if an online content is really made or manipulated by deepfake tech, there are several open sources and paid software. There is even some software specially made to help forensic investigators to better examine content to check if it is real or fake.

1. FaceForensic++ is a forensic database which has nearly 1.8 million manipulated content. Those dataset contents are considered to be benchmark to train any Deepfake detection tool, software and models. The content in FaceForensic++ are made mainly focusing on techniques that allows alteration of content such as Deepfake, Neural Textures and Face Swap
2. Intel Software System has developed a real time deepfake detection tool by the name "FakeCatcher". This detector has the ability to detect deepfakes in milliseconds with maximum of about 96% accuracy. This detector focuses on the flow of pixel data to detect any presence of alteration.
3. Microsoft has a tool developed in the name of "Video Authenticator Tool". This tool can detect both blending of deepfake and the underlying original image . It can

provide a confidence score for a content in real time monitoring system and detect grayscale changes.

4. CyberScience Lab has introduced a DeepFake Detector for forensic analysts and researchers. This detector can allow for frame-by-frame examination. It will store the metadata of each frame along with the case details which can be entered in the Case Setting Window option of the tool. This tool can show frame-by-frame manipulation level of the data along with the result showing if it is real or fake.
5. DeepFake Image Detector is a Chrome extension introduced by CyberScience Lab. This extension allows us to examine images present on a website. The user can right click on an image and start detection process by choosing the extension. The server will run the diagnostics in the back end and send the results back to the user.
6. Sumsub has released a website model to detect deepfake. This model was trained by analysing more than 2 million deepfakes and more than 120k real images to improve its accuracy and detection levels. This is a free deepfake detection model which can be accessed and downloaded from their website.

- **META DATA TAGGING**

We could tag AI generated videos with special metadata, settings attributes that differentiate the content as AI Generated. If we can device a method to carry forward this metadata irrespective of the number of times the content is shared, we can always detect AI generated content very easily by just looking at the metadata. This will help in official settings and give rise to more awareness as AI content can be quickly revealed. Since metadata isn't visible normally, it could prove to be a lot more useful than simple watermarks, and also prove to be less inconvenient. With this method the complex use of software can be avoided and even a very common person with basic computer knowledge can easily look up metadata settings using properties to know whether the content is influenced by Artificial Intelligence or not.

- **AID OF GOVERMENT IN DEEP FAKE DETECTION**

Despite having tons of websites and applications online for detection of deepfakes, there are literally no such websites aided or created by government. India has the lowest price for 1GB

in the world, showing Indians can access internet cheaper than anyone in this world thus increasing the user base. But the government has not taken a step forward to spread awareness for this contemporary crime. Despite having many free sites online, government aided, or government owned sites increase awareness and easy accessibility for the public with less knowledge about the existence of such crime. Apart from bringing separate law and punishments for creating and misusing deep fakes, government can create free site for where victims can raise complaint safely without any fear of society or family pressure because victim blaming is very common in India especially cases involving women.

Government site can contain the following features which make complaining and forming database regarding deep fake crimes easier

1. Just like other online complaining sites hosted by the government, this site can collect personal details and other required details.
2. The feature of posting original and allegedly deep faked image can be given with space for mentioning the details about the original image, when and where the deep fake image is found.
3. Site should be able to detect whether there is any manipulation done to the image just like any other paid software.
4. The feature to track down the original source of the deep fake and user of the device can be added to increasing fear among such creators who hide behind anonymity provided by the internet.
5. After the verification of image government can take down the image/video from the internet and arrest the creator and punish accordingly.
6. Apart from complaining, this site can also act for detection of artificial intelligence in any content enabling public to cross check the data they consume.

Whenever the government steps in there is always a sense of relief and spread of awareness among society. Sites like this can easily be so much helpful for many victims who fear talking about this to their family and make them come forward seeking justice on their own. Upon increase of awareness among society on how to differentiate between fake and original victim blaming reduce.

REFERENCES

REFERENCES

1) Ajder, H., Cavalli, F., Patrini, G., & Cullen, L. (2019). The State of Deepfakes: Landscape, Threats, and Impact. Deeptracelabs.
2) Chadha, A., Kumar, V., Kashyap, S., & Gupta, M. (2021). Deepfake: an overview. In *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security: IC4S 2020* (pp. 557-566). Springer Singapore.
3) Fido, D., Rao, J., & Harper, C. A. (2022). Celebrity status, sex, and variation in psychopathy predicts judgements of and proclivity to generate and distribute deepfake pornography. Computers in Human Behavior, 129, 107141.
4) Finger, L. (2022, September 8). Overview of how to create deepfakes - it's scarily simple. Forbes. https://www.forbes.com
5) Gosse, C., & Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communication, 37*(5), 497–511.
6) Hancock, J. T., & Bailenson, J. N. (2021). The social impact of deepfakes. Cyberpsychology, behavior, and social networking, 24(3), 149-152.
7) Jha, P., & Jain, S. (2021). Detecting and Regulating Deepfakes in India: A Legal and Technological Conundrum. *Available at SSRN 4411227*.
8) Kalmykov, M. (2023, November 28). Deepfake technology in video industry. (n.d.). https://www.dataart.com
9) Karasavva, V., & Noorbhai, A. (2021). The real threat of deepfake pornography: A review of Canadian policy. *Cyberpsychology, Behavior, and Social Networking*, *24*(3), 203–209.
10) Katarya, R., & Lal, A. (2020, October). A study on combating emerging threat of deepfake weaponization. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) (pp. 485-490). IEEE.
11) Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *Iscience*, *24*(11).
12) Mahmud, B. U., & Sharmin, A. (2021). Deep insights of deepfake technology: A review. arXiv preprint arXiv:2105.00192.

13) Mania, K. (2024). Legal protection of revenge and deepfake porn victims in the European Union: Findings from a comparative legal study. *Trauma, Violence, & Abuse, 25*(1), 117-129.
14) Neekhara, P., Dolhansky, B., Bitton, J., & Ferrer, C. C. (2021). Adversarial threats to deepfake detection: A practical perspective. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 923-932).
15) Nguyen, T. T., Nguyen, Q. V. H., Nguyen, D. T., Nguyen, D. T., Huynh-The, T., Nahavandi, S., ... & Nguyen, C. M. (2022). Deep learning for deepfakes creation and detection: A survey. Computer Vision and Image Understanding, 223, 103525.
16) Ojha, A. (2024, January 20). Rashmika Mandanna deepfake: How cops traced accused, techie from Andhra's Guntur. India Today. https://www.indiatoday.in
17) Rahman, A., Islam, M. M., Moon, M. J., Tasnim, T., Siddique, N., Shahiduzzaman, M., & Ahmed, S. (2022). A qualitative survey on deep learning based deep fake video creation and detection method. Aust. J. Eng. Innov. Technol, 4(1), 13-26.
18) Semwal, A. (2020). Deep Fakes, Artificial Intelligence & Eco Species.
19) Sen, D. (2021, April 3). Explained: Why is it becoming more difficult to detect deepfake videos, and what are the implications? The Indian Express. https://indianexpress.com
20) Varma, V. (2023, July 17). Kerala man loses ₹40k to AI-enabled deep-fake fraud. Hindustan Times. https://www.hindustantimes.com

ANNEXURE

Cyber Victimization: A Study on Deepfakes and Effects of Artificial Intelligence

Vijayasarathy R and Mohammed Marzuk T M, we are currently conducting a study on the above-mentioned topic in hopes of providing awareness on Deepfake and any ill effects as a result of Artificial intelligence. This survey will help us to get sufficient data on how much people are aware of cybercrime related to Deepfake and Artificial intelligence.

Deepfakes: Deepfakes morphed videos or images in which a normal image or video is converted as a pornographic or any other form of fake one using super imposition over other inputs

We thank you for attending this survey.

---

* Indicates required question

- ***Email ****
- ***Name ****
- ***Age ****

18-20

21-26

- ***Sex ****

Male

Female

- ***Occupation ****

UG

PG

Employed

Unemployed

- ***Marital status ****

Married

Single

Divorced

SPECIMEN

- ***Have you heard of the term "deep fake" before? ****

Yes

No

- ***Have you ever encountered a deep fake video or image on social media? ****

Yes

No

- ***How concerned are you about the potential misuse of deep fake technology? ****

slightly

very less

deeply concerned

- ***Do you believe deep fakes pose a threat to political elections and public figures? ****

Yes

No

May be

- ***Should there be stricter regulations on the creation and distribution of deep fake content? ****

Yes

No

- ***How do you think deep fake technology can be effectively countered or detected? ****

Public should check the authenticity of the content

Government should establish awareness

Social media guidelines should restrict illicit deepfakes

- ***Would you support the use of AI-driven tools to identify and flag deep fake content? ****

Yes

No

- ***Do you think deep fakes have the potential to undermine trust in media and journalism? ****

Yes

No

Maybe

SPECIMEN

- ***How often do you verify the authenticity of the content you consume online? ****

Every time

not very often

depends on the content

- ***Do you consider spreading creating and spreading deepfakes as sexual offense ****

Yes

No

maybe

- ***Would you be willing to take steps to educate yourself and others about the risks associated with deep fakes? ****

Yes

No

- ***Are you aware of the procedures to be followed if you ever become a victim of deepfake ****

Yes

No

- ***In your opinion, should government create free sites to help deepfake victims. ****

Yes

No

SPECIMEN