The Perception of Deepfake Technology Among University Students in Malaysia

Parveen Kaur[1*], Arif Jawaid[2], Steve Mathieus Anak Robert[3], Stephenie Anak Selutan[4], Tan Hoang Jenq[5], Tan Pei En[6], Veronica Nelson Nyalau[7]

[1]*Associate Professor, University Malaysia Sarawak, Malaysia; E-mail: sspkaur@unimas.my*
[2]*Professor, Lahore Garrison University, Pakistan*

[3,4,5,6,7]*University Malaysia Sarawak, Malaysia*

Abstracts: It was alarming to know that a new technology (Deepfake) is paving the way to social media by demoralising popular people (celebrities) and others. It has badly affected reputation of many people without their knowledge and consent. It was decided to know about Deepfake technology at length and its extent in harming and hurting the people. A group of university students were randomly selected and administered through a questionnaire. The questionnaire comprised questions pertaining to awareness of deepfake and its consequences. The objectives of the study were to examine the emergence of Deepfake technology, the impact of Deepfake technology utility, and understanding the strategies to mitigate irresponsible usage of Deepfake technology among social media users. The descriptive data was analyzed using numbers and percentages. The major findings included its existence, impact on social media and strategies to mitigate irresponsible usage of Deepfake technology. A wider and larger diverse community using social media extensively will provide better results.

Keywords: Deepfake, AI, Social Media, University, Malaysia

1. INTRODUCTION

The aim of this paper was to examine and develop a deep understanding of the emergence and usage of Deepfake technology among the social media community. According to Chawla (2019), recent technology developments have made it simple to produce so-called "Deepfakes," which are hyper-realistic videos created with face swaps that barely reveal the manipulation. The term "deepfake" first appeared in 2017 when a group of Reddit users produced fake celebrity pornographic movies using Artificial Intelligence (AI) (Maddocks, 2020). Artificial intelligence programs that merge, combine, swap out, and superimpose pictures and video clips to make fake videos that look real are known as "Deepfakes" (Maras & Alexandrou, 2018). The term techniques for deep learning are used in Deepfake videos with a huge sample of video pictures as input to achieve face swapping. For instance, Deepfakes frequently feature the faces of politicians, famous people, comedians, and other celebrities in pornographic videos (Hasan & Salah, 2019). It is because there are so many sources of images and videos of them on the Internet that can be used to create the substantial picture stockpiles needed to train an AI Deepfake system. The more samples used, the more realistic the result will be. Moreover, deepfake is used either to exact retribution, to generate and upload a pornographic image of a celebrity, or to extort money from a target (Harris, 2018).

2. PROBLEM STATEMENT

People can find a voice and a sense of purpose through Deepfakes. New ideas and capacities for empowerment have come from all facets of society, including business, accessibility, public safety, and the arts, expression, etc. However, Deepfake technology has advanced to a troubling degree, leading to a considerable rise in fraud and other fraudulent practices that prey on the gullible. Therefore, this study aims to describe the advantages and disadvantages of Deepfake technology with survey interviews and case studies. The researchers interviewed 50 students from University Malaysia Sarawak (UNIMAS) coming from different courses and backgrounds. This enabled the researchers to acquire a better understanding of both the advantages and disadvantages of deepfake technology. Furthermore, examining the respondents' real responses to this subject will give us important data for our study. The objectives of the study were to examine the emergence of Deepfake

technology, the impact of Deepfake technology utility, and understanding the strategies to mitigate irresponsible usage of Deepfake technology among social media users.

3. LITERATURE REVIEW

Artificial Intelligence (AI) is a powerful tool that has the potential to change the world, especially in this modern world. This is because AI is very important and has got popularity these days. Certainly, it is important to be aware of both the potential benefits and risks of artificial intelligence as it continues to evolve. Artificial Intelligence programs possess important signals that can be used to absorb and generate more new Deepfake content, leading to an increase in Deepfake videos, especially on social networking platforms (Helmus, 2022). For example, research by Helmus (2022) stated that Google released a large database of Deepfakes in 2019, with the aim of helping to improve detection and similar versions of the public domain. Thus, according to Helmus (2022), the synthesis and availability of known examples of synthetic media greatly improve the development of detection algorithms.

Chesney and Citron (2019) stated that Deepfake technology is a type of artificial intelligence (AI) that can be used to create realistic and convincing videos or audio recordings of people saying or doing things they never actually said or did. This technology is also used for a variety of purposes, including entertainment, education and research. This is because Deepfakes are created by using machine learning algorithms to train a computer model on a large data set of images or audio recordings of a person. For example, both Chesney and Citron (2019), agreed that any recordings could be faked using tools that were available to almost anyone with laptops, computers, smartphones, and with access to the Internet nowadays. Furthermore, one of the common deepfake technologies that have been used by them is known as the 'Face Swap' which is also one of the segments of artificial intelligence that has been created to change and swap their face with someone else after the editing has been done. Thus, this has led to a growing concern among social media users nowadays especially about the potential for Deepfakes to be used to spread misinformation and disinformation or to damage people's reputations.

The difference between artificial intelligence (AI) and deepfake technology is that AI is a broad term that covers many types of technologies, while Deepfake technology is a specific type of AI utilised to create video or audio recordings to be more real and convincing. For example, commercial and even free Deepfake services have emerged on the open market and are even worse on the illegal black market (Chesney and Citron, 2019). Furthermore, AI is a nascent field and there are still many challenges to overcome, such as ensuring that AI systems are safe and secure and that they are not used for malicious purposes in the long run, whereas Deepfake technology is a relatively new technology that was created by artificial intelligence and there are still many ways to improve the quality of Deepfakes. Thus, the emergence of deepfake technology is a new and difficult problem, especially for social network users. For example, there are many reported cases, especially among women, where their faces were included in pornography without their consent, as these irresponsible Deepfake creators involved constantly developing new techniques to make their videos more realistic, and these developments are difficult to follow (Chesney and Citron, 2019). This is why it is so important that social media users, governments, and the tech industry work together to address the risks of Deepfakes and protect people from harm.

A Deepfake case study by Smith and Mansted (2020) sheds light on the use and impact of Deepfake technologies. They agreed that deepfake technology has become democratized because it also has other services that are accessible at a low cost, such as the fake speech-generating service Lyrebird, which offers subscriptions for its tools. Therefore, both also agreed that Deepfakes can become weapons if misused. Indeed, they can be used to smear targets, impersonate or blackmail elected officials, and used in conjunction with cybercriminal activities (Smith and Mansted, 2020). Furthermore, Helmus (2022) also agreed that the misuse of Deepfakes can increase the risk of misinformation at home and abroad. He also said that threats have realized that women are the main target of artificial intelligence (AI) on these pornographic websites. Helmus (2022) also argues that Deepfake images have been widely used as part of fake social media accounts. This is because undoubtedly, the use of fake photos allows propagandists to circumvent this defence and use unobtainable photographs. In addition, the case study performed by Sardar et al. (2019) mentioned that algorithmic social networking platforms can detect what you

like beforehand and then lock you into a group of people with similar ideologies and desires. Through their study, Sardar et al. (2019) attempted to discover that advances in artificial intelligence (AI) and the rise of robotics have raised profound ethical and moral questions that require urgent attention from the public, especially among Islamic scholars and thinkers.

4.RESEARCH METHODOLOGY

Research methodology is a set of methods used to conduct an empirical study properly. This exact section will discuss in detail the study samples, instruments, and the process of extracting data from various sources.

4.1. Sampling

This study used sampling technique to achieve its objectives. A total of 50 respondents were chosen to answer the questionnaire that was divided into three sections in line with the research objectives. The criteria for selecting the respondents were that they must be knowledgeable about technology, especially Deepfake technology for them to be valid respondents for this study. This is since they are a targeted audience that may have been exposed to this issue.

4.2. Questionnaire for Respondents

This study involved the respondents who had been exposed to Deepfake technology on social media platforms. The responses to the questionnaire will provide quantitative data about the study. The information gathered will be presented in various forms that consist of charts and numbers. The figures were designed to give insights into and further understanding of the objectives. The questionnaire for this study covered three sections namely sections A, B and C. Section A covered questions that involved the respondents' demographic details including age, race, and their experience with Deepfake technology. Section B covered 10 questions such as the impact of the usage of deepfake technology among social media users. And finally, section C covered 6 questions about the strategies to mitigate irresponsible usage of Deepfake technology among users. For example, how do they avoid from this grave issue and what were their opinions regarding this issue?

4.3 Data Collection

The researchers used survey methods and secondary data for this study. Primary data was gathered from respondents using Google Form platform. Survey questionnaire is a prime example of a method to collect primary data from a group of people (Roopa & Rani, 2012). This includes a set of questionnaires to be answered by the targeted audience or respondents. The questionnaire generally comprised questions regarding the understanding of the emergence of deepfake technology, its impact and the solution on how to use the technology properly. Secondary data was also used to back up the data that was collected from the respondents. A series of credible and related journal articles were thoroughly studied to support the study.

4.4 Data Analysis Method

The research data was collected mainly by utilizing questionnaire through google form. Fifty (50) respondents filled the forms voluntarily. The descriptive data was analyzed using numbers and percentages. Pie chart was used to show the percentages and numbers explicitly.

5. FINDINGS

5.1. To Examine the Emergence of Deepfake Technology Among Social Media Users.

Deepfake technology has become well-known among the social media community. This is since Deepfake technology offers unique features that are quite incomparable at this point from the other technologies. These mentioned features such as editing fake videos and images are now the center of attraction not only for good use but also being used by users that have ill intentions (Maras & Alexandrou, 2018).

Based on this study involving 50 respondents, 70% knew about this technology, and that made them well-aware of the features and capabilities the technology has to offer. According to Chawla (2019), the development of technologies has made it easy for Deepfake to be used worldwide, exposing the social media community frequently and with ease. The users do not even need to search for Deepfake content, as the creators of this technology can upload their Deepfake content on mainstream social media platforms such as TikTok, YouTube, Instagram, and many more. It can be considered true that the social media community has only been exposed to this kind of content due to the fact that the creators were responsible for spreading it. This study has revealed that 78% of the respondents have not used Deepfake technology before. There are various reasons and one of them is because this technology requires complex knowledge and editing skills where only skilled people can operate. To master this technology, one needs quite some time. Deepfake technology is operated by merely two networks that are called the ‘generator’ and ‘discriminator’. This ensured that only users that are considered professionals in Artificial Intelligence (AI) could operate this efficiently. However, the social media community has started to be well-versed in the knowledge of Deepfake technology. While it is considered basic knowledge for all, 78% of the respondents that participated in the study or survey were aware that Deepfake technology uses AI to operate. Thus, we can understand that social media users have technically been exposed to deepfake technology, especially Deepfake contents. We also anticipate that the social media community will experience an increasing trend in the number of users that will be exposed to this specific technology in the future.

5.2. To Identify the Impact of Deepfake Technology Utility Among Social Media Users.

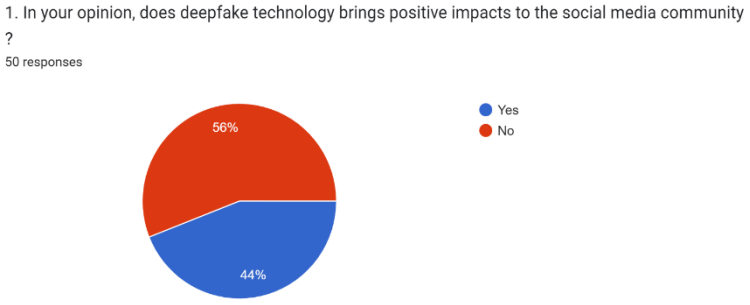

Figure 5.2. Deepfake impact

According to the survey results, 56% of the respondents agreed that deepfake technology brought positive impact on the social media community while only 44% of the respondents disagreed.

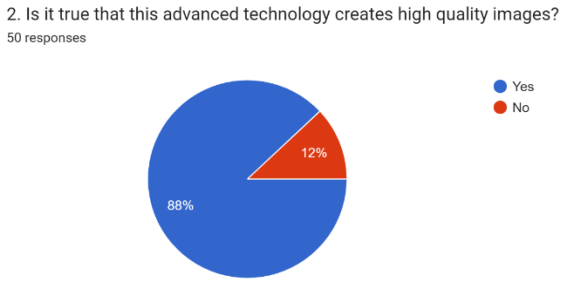

Figure 5.2 Deepfake creates high quality images

According to the survey results, 88% of the respondents agreed that it is true that this advanced technology creates high-quality images while only 12% disagreed.

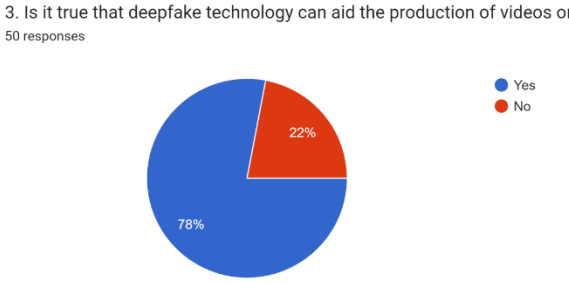

Figure 5.3 Aid in production of videos or clips

Based on the survey results, the majority of respondents (78%), indicated that it is true that deepfake technology can aid in the production of videos or clips. On the other hand, 22% of the respondents stated otherwise.

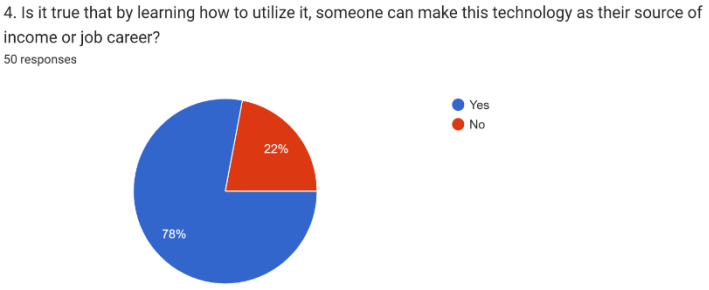

Figure 5.4 Source of income or job career

Based on the survey results, an overwhelming majority of respondents (78%) believed that by learning how to utilize it, someone can make this technology their source of income or job career while only 22% of the respondents did not agree.

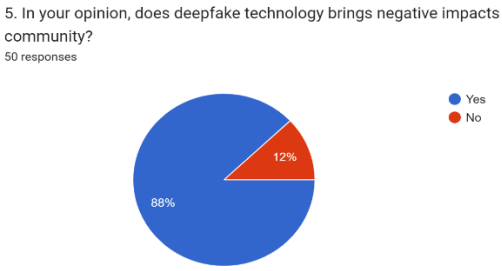

Figure 5.5 Negative Impact on the social media community

A significant majority of respondents (88%) believed that Deepfake technology brings negative impact to the social media community. This indicates that a substantial portion of the surveyed individuals recognize that Deepfake technology brings negative impact to the social media community. Only a small minority of respondents, (12%) disagreed.

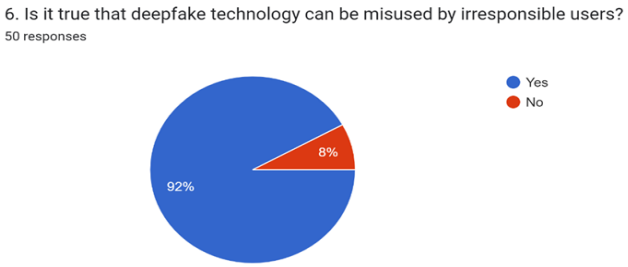

Figure 5.6 Misuse by irresponsible users

Based on the survey results, there are 92% respondents agreed to the statement that Deepfake technology can be misused by irresponsible users. According to Howard (2018), the cause of this is that irresponsible people engaged in cyberterrorism can employ Deepfake technology to make all of their acts more damaging. They have the ability to continuously adapt to the tools that hackers utilize to dominate. Cyberterrorists can use the Internet and information technology to take advantage of political, cultural, and social issues. For instance, the widespread dissemination of hoaxes and false information is intended to damage the reputations and integrity of certain people or organizations. However, only 8% of the total respondents, disagreed with the statement. It is due to certain individuals believing that Deepfake technology cannot be misused by irresponsible users.

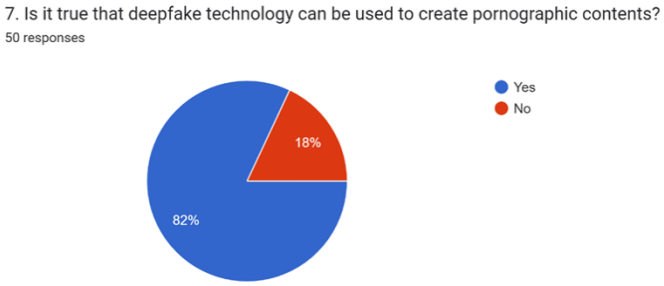

Figure 5.7 Creation of pornographic contents

According to the pie chart above, 82% agreed that Deepfake technology can be used to create pornographic contents. This is because no security measures can completely rule out the potential of deceptive Deepfake being

distributed over reliable channels (Harris, 2021). Additionally, there are no such rules to deal with pornographic contents, which make it impossible to control and negatively impact the reputation and dignity of well-known individuals. However,18%, disagreed that Deepfake technology may be used to produce sexual contents. This is due to the fact that individuals consider the pornographic Deepfake scenarios to be very bad, hurtful, and punishable (Kugler & Pace, 2021).

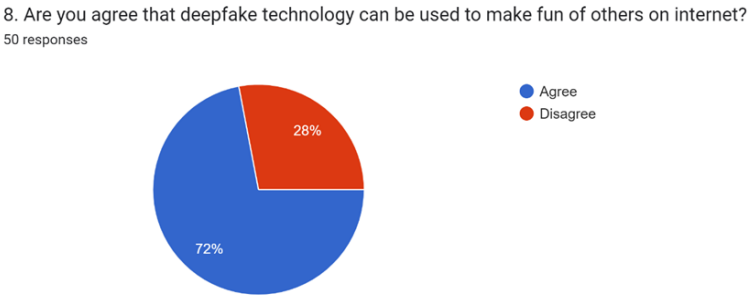

Figure 5.8 Can be used to make fun of others on Internet

Based on the pie chart above, 72% of the respondents agreed that deepfake technology can be utilised to make fun of others on the internet. Most Deepfakes, according to Westerlund (2019), are jokes, pranks, and amusing memes with funny or sarcastic effects. Such actions could bring disgrace to individuals. However, just 28% disagreed with the statement. According to the respondents, Deepfake technology cannot be used to mock others online.

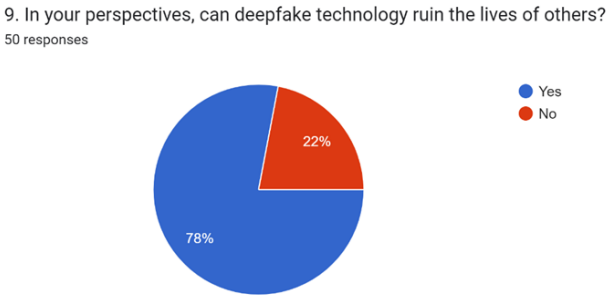

Figure 5.9 Can deepfake ruin the lives of others

According to the survey results, 78% agreed that deepfake technology may endanger the lives of others, while 22% disagreed. It appears that false news has steadily jeopardized the right to free speech and even human civilization, which is why Deepfake technology may endanger the lives of others (Qayyum, 2019). Deepfakes may be misused to harm a person's reputation and privacy. For example, to spread lies, to encourage violence, to attack institutions and leaders, and even to rig elections. Additionally, Deepfakes technology enables false information to be sent to the public, manipulating people's emotions and inciting a general distrust in society. Those who disagree with the statement are 22%. They may believe that recognising deepfake videos or contents is a strategy to stop the use of deepfake technology. Sabir (2019) claims that the authors have put forth a recurrent convolutional model (RCM) to use the time lag between frames to identify Deepfake films or other contents. Making other people's life miserable could be avoided in this way.

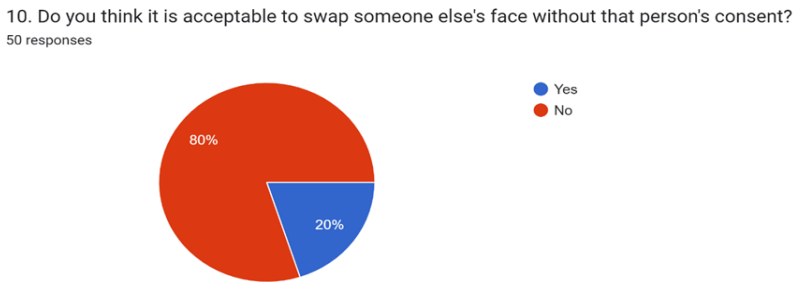

Figure 5.10 Swapping of face without consent

According to the pie chart above, 80% agreed that it is acceptable to swap someone else's face without that person's consent using Deepfake technology. Binder (2018) claims that the technology is now more accessible than ever, in part because of the launch of the "Fake App" software, which allows users to make Deepfakes. Anyone can create films using this fake app if they have one or two good quality videos of the faces they want to spoof. Nevertheless, there are 20% disagreed with this statement. When a person's face is used without his express consent, the right to publicity is directly violated. The individual whose resemblance appears may be able to file a right of publicity lawsuit if a creator makes money. He also exploits that person's image in a Deepfake without that person's permission (Langvart, 2013).

5.3 To Understand the Strategies to Mitigate Irresponsible Usage of Deepfake Technology Among Users.

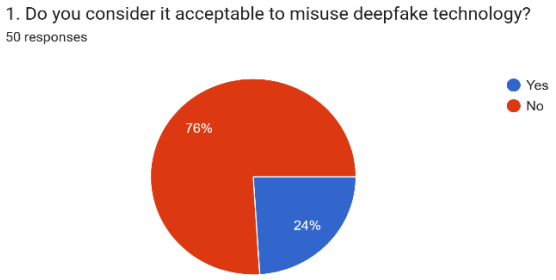

Figure 5.11 Misuse of Deepfake technology

According to the above pie chart survey, the highest number of respondents (76%), agreed that it is unacceptable to misuse Deepfake technology. In addition, only 24% agreed that it is acceptable to misuse deepfake technology. Hence, we can say that most of the respondents agreed that it is unacceptable to misuse Deepfake technology. This is mainly because Deepfake technology offers a serious concern if it ends up in the wrong hands because of its capacity to manipulate and generate audio visual contents so convincingly. False information can incite rage, particularly when it concerns sensitive social issues amid emergencies like war, fire, earthquakes, or protests (Temir, 2020). False information can constitute a serious threat to society, and even an entire nation. Therefore, this makes it possible for individuals with little technical knowledge and no artistic training to edit films, swap faces, change expressions, and synthesize speech with a high degree of accuracy. The worries of the intelligence community are one specific illustration of a potential threat. According to Westerlund (2019) showed his concerns that the nefarious use of Deepfakes could endanger national security. It becomes possible for political propaganda to be disseminated and for critical election campaigns to ruin the image of people. Through the use of this technology, opponents might stoke unrest, manipulate public opinion, and threaten the very underpinnings of democratic processes. Obviously Deepfake targets social media platforms because they facilitate dissemination of rumors, conspiracies, and false information.

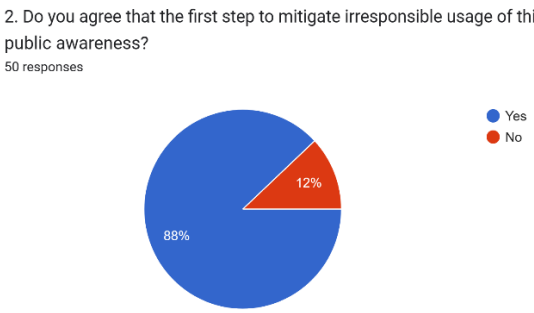

Figure 5.12 First step to Mitigate irresponsible usage

Other than that, it is found that 88% of respondents agreed that the first step to mitigate irresponsible usage of this technology is to first raise public awareness. Many respondents agree with this statement because most of the people are unaware of current technologies like Deepfake and its effects. Bangladesh, for instance, is undergoing fast digitization because of its expanding economy and rising youth population (Ahmed et al, 2021). As technological advancements bring numerous benefits, Ahmed et al. (2021) stated that it also exposes the people to heightened risks of cybercrime. The unfortunate reality is that a sizable section of the populace is still ignorant of the possible repercussions linked to these risks. As a result, spreading knowledge about the risks of cybercrime, particularly the pernicious effects of Deepfakes, is even more important in Bangladesh's fight against these malevolent practices. People can be better prepared to recognize and thwart these digital manipulations by being more knowledgeable about the dangers and ramifications of Deepfakes.

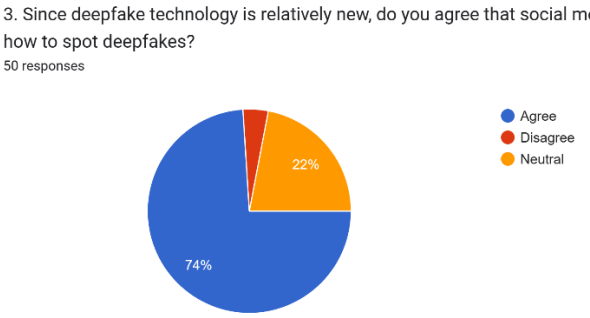

Figure 5.13. Do you agree that social media users need to learn how to spot deepfakes?

The above pie chart shows 74% of the respondents agreed that social media users need to learn how to spot Deepfakes since Deepfake technology is relatively new. There are respondents who chose to be indifferent about this statement (22%).

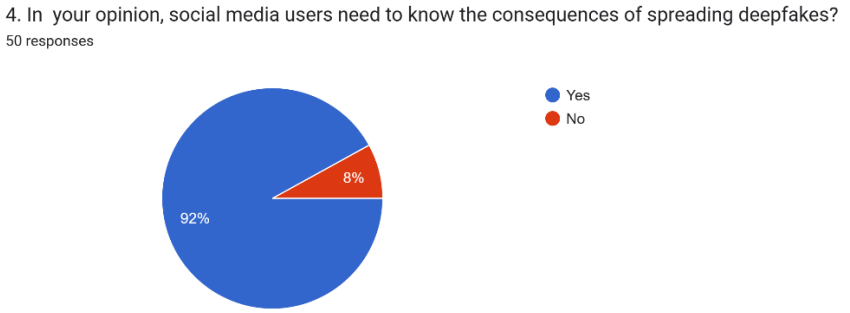

Figure 5.14 Need to know the consequences of spreading deepfakes

The highest number of respondents (92%) chose “yes” to the social media users who need to know the consequences of spreading Deepfakes. This is because Deepfake technologies are recognized for producing fake media content such as face and voice swaps, fake celebrity pornography, and other types of false information. Hence, it is very important for the public to know the bad sides of Deepfake technology,

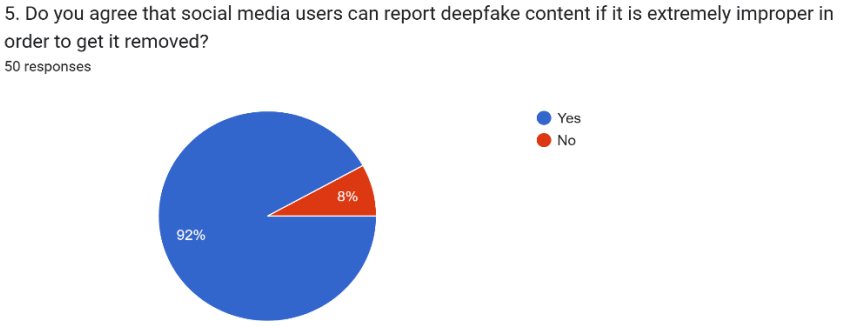

Figure 5.15. Can social media users report deepfake contents?

The pie chart shows that 92% of the respondents chose yes about the statement that social media users can report Deepfake contents to get it removed. Therefore, most of the respondents agreed that social media users can report Deepfake contents if it is extremely improper to get it removed.

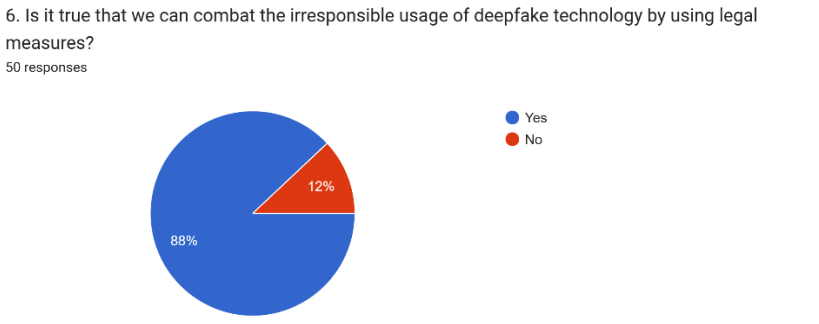

Figure 5.16. Deepfake technology can be combatted by legal measures

The above pie-chart shows that 88% agreed with the statement that they can combat the irresponsible usage of Deepfake technology by using legal measures. However, 12% of the respondents strongly disagreed with the statement. Finally, a great majority of the respondents chose “yes” to this statement. People who become victims of Deepfake creations may have the choice to file civil lawsuits against the authors in common law nations including the US, UK, and Malaysia (The Sun Daily, 6 June 2019). To prevail in their lawsuit, the victim, or claimant, would have to show that they have been misrepresented in a way that the typical person would find embarrassing or offensive. Through this legal avenue, people can seek redress and hold those accountable for the production and distribution of harmful Deepfakes. However, law scholars think that Deepfake victims would have little to no legal options (Caldera, 2019). There would be restrictions on who could be sued by victims due to the threshold issues. Due to the ubiquity of anonymity on the Internet, a person who has been injured by a Deepfake may not have a party against who to file a lawsuit if they are unable to track down the video's developer.

Furthermore, the Communications Decency Act shields websites from criticism of third-party contents. Consequently, there is little chance of success in using a social media platform for hosting a Deepfake (Farakhmanesh, 2018). Therefore, given the variations in laws across different countries, it is essential for users to exercise wisdom and prudence while navigating the online realm.

CONCLUSIONS

In conclusion, according to our data, we would conclude that the emergence and use of Deepfake technology have positive and negative effects on the social media community. Deepfake technology also plays an important role in the education and enhancement of video or film production and can also be a kind of spiritual support and comfort. However, misusing Deepfake technology will bring a negative impact such as creating pornographic contents, ruining the lives of others, and more. It will also bring negative effects and serious consequences to individuals, organizations, and societies. Finally, through this research we hope that the public or the society will have more awareness to use it wisely.

REFERENCES

[1] Ahmed, M. F. B., Miah, M. S. U., Bhowmik, A., & Sulaiman, J. B. (2021, July). Awareness to Deepfake: A resistance mechanism to Deepfake. In 2021 International Congress of Advanced Technology and Engineering (ICOTEN) (pp. 1-5). IEEE.

[2] Caldera, E. (2019). Reject the evidence of your eyes and ears: deepfakes and the law of virtual replicants. Seton Hall L. Rev., 50, 177.

[3] Chawla, R. (2019). Deepfakes: How a pervert shook the world. International Journal of Advance Research and Development, 4(6), 4-8.

[4] Chesney, R., & Citron, D. (2019). Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics. Foreign Affairs, 98(1), 147–155. https://www.jstor.org/stable/26798018

[5] Fake videos and the law. (n.d.-b). www.thesundaily.my. https://www.thesundaily.my/opinion/fake-videos-and-the-law-FG976531

[6] Farokhmanesh, M. Is It Legal to Swap Someone’s Face into Porn Without Consent?, Verge, Jan. 30, 2018.

[7] Harris, D. (2018). Deepfakes: False pornography is here and the law cannot protect you. Duke L. & Tech. Rev., 17, 99.

[8] Harris, K. R. (2021). Video on demand: what deepfakes do and how they harm. Synthese, 199(5-6), 13373-13391.

[9] Hasan, H. R., & Salah, K. 2019. Combating Deepfake Videos Using Blockchain and Smart Contracts. IEEE Access, 7: 41596–41606. https://doi.org/10.1109/ACCESS.2019.2905689

[10] Helmus, T. C. (2022). Artificial Intelligence, Deepfakes, and Disinformation: A Primer. RAND Corporation. http://www.jstor.org/stable/resrep42027

[11] Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. Journal of information technology & politics, 15(2), 81-93. https://doi.org/10.1080/19331681.2018.1448735

[12] Howcroft, E. (2018). How faking videos became easy and why that’s so scary. Bloomberg: New York, NY, USA.

[13] Kugler, M. B., & Pace, C. (2021). Deepfake privacy: Attitudes and regulation. Nw. UL Rev., 116, 611.

[14] Maddocks, S. (2020). ‘A Deepfake Porn Plot Intended to Silence Me’: exploring continuities between pornographic and ‘political ‘Deepfakes. Porn Studies, 7(4), 415-423.

[15] Maras, M. H., & Alexandrou, A. 2019. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. International Journal of Evidence & Proof, 23(3): 255–262. https://doi.org/10.1177/1365712718807226

[16] Matt Binder (2018), The U.S. Defense Department is Readying for the Battle Against Deepfakes, MASHABLE https://mashable.com/article/defense-departmentfighting-deepfakes/#W6PJhu3Q0aqE .

[17] Qayyum, A., Qadir, J., Janjua, M. U., & Sher, F. (2019). Using blockchain to rein in the new post-truth world and check the spread of fake news. IT Professional, 21(4), 16-24.

[18] Roopa, S., & Rani, M. S. (2012). Questionnaire designing for a survey. Journal of Indian Orthodontic Society, 46(4_suppl1), 273-277.

[19] Sardar, Z., Serra, J., & Jordan, S. (2019). Social Media and Artificial Intelligence. In Muslim Societies in Postnormal Times: Foresights for Trends, Emerging Issues and Scenarios (pp. 93–101). International Institute of Islamic Thought. https://doi.org/10.2307/j.ctv10kmcpb.21

[20] Smith, H., & Mansted, K. (2020). What's a Deepfake? In Weaponised Deepfakes: National security and democracy (pp. 05–10). Australian Strategic Policy Institute. http://www.jstor.org/stable/resrep25129.6

[21] Temir, E. (2020). Deepfake: new era in the age of disinformation & end of reliable journalism. Selçuk İletişim, 13(2), 1009-1024.

Westerlund, M. 2019. The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review, 9(11): 40-53. http://doi.org/10.22215/timreview/1282

[22] Xia, J., & Hua, M. (2021). Numerical Analysis and Optimization Application Trend of Deepfake Video Inspection Technology. In Journal of Physics: Conference Series (Vol. 1746, No. 1, p. 012068). IOP Publishing.

DOI: https://doi.org/10.15379/ijmst.v10i2.2596