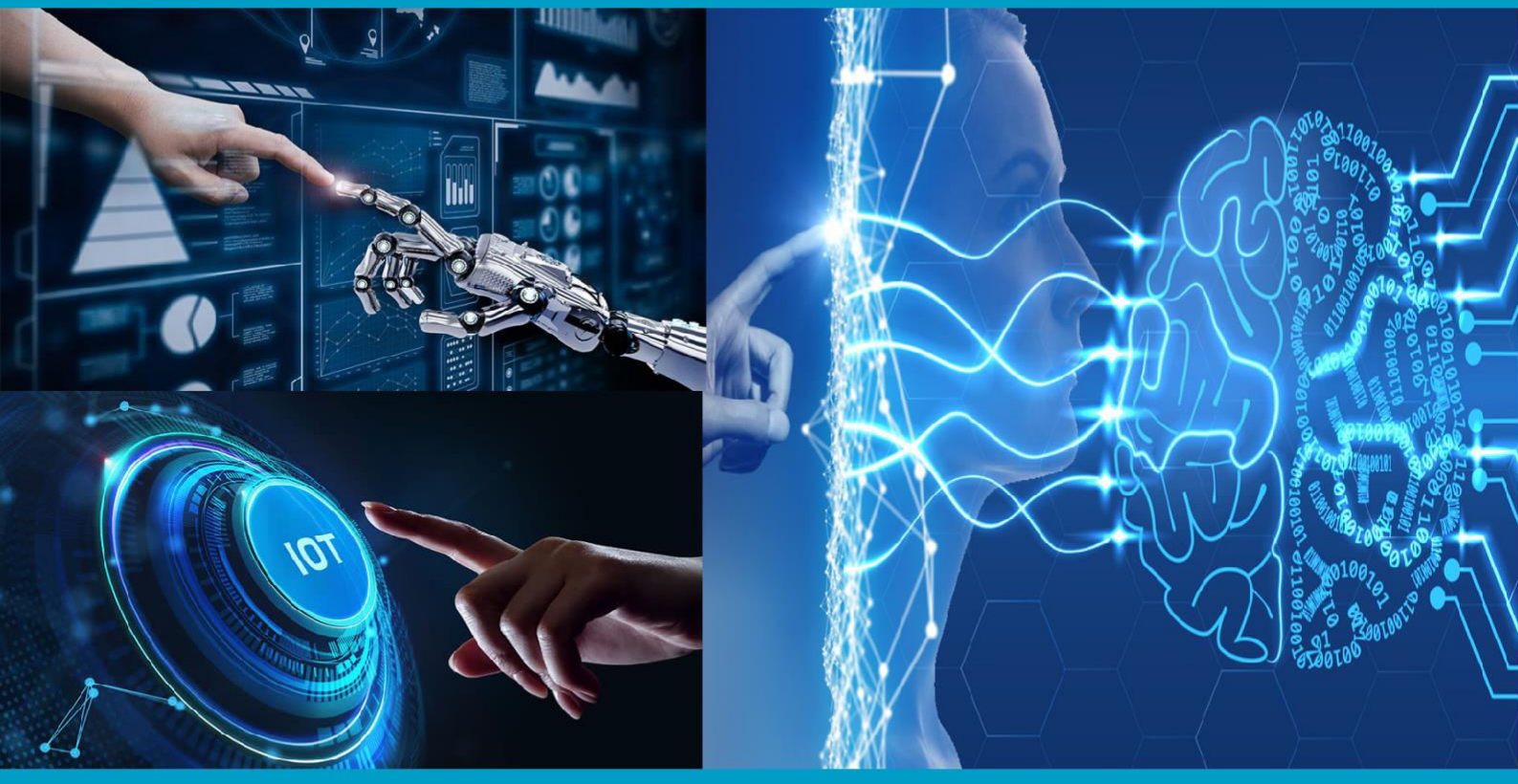
IJMRSETM

e-ISSN: 2395 - 7639



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 10, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.580

International Journal of Multidisciplinary Research in Science, Engineering, Technology &Management (IJMRSETM)

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

Unveiling the Digital Mirage: An Introduction to Deepfakes

Dr. Archana Sahai

Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow, India

ABSTRACT: The world that we live in today is flooded with information from sources that sometimes are true and sometimes are not. Just a simple change of narrative on the same thing might drastically change people's opinions. With fabricated media already rampant, a new player has emerged in the synthetic media industry, which goes by the name of 'deepfake', which is a type of fabricated video made using deep learning, accurate to such an extent that it is sometimes impossible to differentiate the actual video from the fake. The report consists of information about deepfake videos, their origins, their impact on the world, their good and bad uses and how to spot them. Their place in today's world and their legality are also talked about.

KEYWORDS: Deepfake, Artificial Intelligence, Machine Learning, Cyber Security, Cyber Crimes.

I. INTRODUCTION

Deepfakes, is a combination of "deep learning" and "fake," refer to a rapidly evolving technology that allows the creation of highly realistic and convincing manipulated videos, images, or audio. Leveraging advancements in artificial intelligence and machine learning, deepfake algorithms have the capability to generate incredibly lifelike simulations of people saying or doing things they never actually did [1].

This emerging technology has gained significant attention and raised concerns due to its potential to deceive, manipulate, and mislead viewers. Deepfakes can be created with relative ease using readily available software and a dataset of targeted individual's media content. By synthesizing facial features, expressions, gestures, and vocal patterns, deepfakes can convincingly superimpose one person's likeness onto another's, effectively enabling the manipulation of visual and auditory information in a variety of contexts.

While deepfakes have found some positive applications in entertainment and creative industries, such as bringing historical figures back to life or enhancing special effects in movies, they also pose significant risks. They have been employed to spread misinformation, defame individuals, and manipulate public opinion. The potential consequences range from political and social destabilization to personal reputational harm [2].

As a response to the growing concern around deepfakes, researchers and experts are actively developing techniques and tools to detect and mitigate their negative impacts. These detection methods utilize advanced machine learning algorithms, analyzing subtle inconsistencies, artifacts, or anomalies that may reveal the presence of a deepfake. However, as deepfake technology continues to evolve, so too do the challenges in accurately detecting them.

Understanding the capabilities, implications, and detection of deepfakes is essential in navigating the digital landscape and promoting media literacy. By raising awareness and staying informed about this technology, individuals can better distinguish between genuine and manipulated content, fostering a more discerning and critical approach to the information they encounter.

However, false information is usually easily spotted. However, with deepfakes now in the market, the game has changed forever, deepfakes are getting harder and harder to detect day by day, and experts have agreed that it is highly likely that deepfakes will be impossible to predict soon.

So, the least we can do is be aware of them so that we are not caught off-guard with the misinformation and our ability to make educated decisions is not hampered. Knowing about deepfake videos is the first step towards spotting them and then stopping them.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

II. DEEPFAKE'S ORIGINS

The concept of deepfakes first gained prominence in 2017 when a Reddit user with the pseudonym "deepfakes" popularized the technique by creating and sharing manipulated adult videos featuring celebrities. These videos were generated using deep learning algorithms, specifically a type of neural network called a generative adversarial network (GAN). GANs consist of two components: a generator network that creates synthetic content, and a discriminator network that tries to distinguish between real and fake samples. Through an iterative process, the generator network improves its ability to generate increasingly realistic outputs that can fool the discriminator network.
Although the term "deepfake" became synonymous with manipulated videos, the underlying techniques and technologies used in deepfakes extend beyond video manipulation alone. Deep learning models have been applied to manipulate images, audio, and even text, allowing for the creation of convincing fake images, voices, and conversations.

III. RISE OF DEEPFAKES

The term first became widely used in **2017**, after a Reddit user by the name 'Deepfakes' posted pornographic videos featuring actresses whose faces were digitally altered to resemble female celebrities, such as Scarlett Johansson and Gal Gadot.
Since then, other examples of disturbing deepfakes have appeared, including a **2018** deepfake created by Hollywood filmmaker Jordan Peele featured former **US President Obama** discussing the dangers of fake news and mocking the current President Trump. The same year one more video of Trump was created; in that video,**Donald Trump** offered advice to the people of Belgium about climate change. The video was created by a Belgian political party,"sp.a" to attract people to sign an online petition calling on the Belgian government to take more urgent climate action. The video provoked outrage about the American President meddling in a foreign country with Belgium's climate policy.
In **2019**, an altered video of American politician **Nancy Pelosi** went viral and had massive outreach; the video was slowed down to make her sound drunk; Donald Trump posted it, and Facebook refused to take it down.[9]

In **2019**, the US Democratic Party deepfaked its own chairman Tom Perez to highlight the potential threat of deepfakes to the 2020 election.[16]

In **2020**, before the **Delhi legislative assembly elections**, BJP politician **Manoj Tiwari** used deepfake to create a video criticizing the Delhi government in multi-language. The video marked the debut for deepfake in the Indian elections.The manipulated video was vastly leveraged to persuade many regional migrant workers and influence the masses by distributing them across 5800 WhatsApp groups.[4]

In **April2020**, the Belgian branch of Extinction Rebellion published a deepfake video of Belgian Prime Minister **Sophie Wilmès** on Facebook.The video promoted a possible link between deforestation and COVID-19. It had more than 100,000 views within 24 hours and received many comments. On the Facebook page where the video appeared, many users interpreted the deepfake video as genuine.[9]

The list seems to be unending Several deepfake videos have gone viral recently, giving millions around the world their first taste of this new technology. The amount of deepfake content online is growing at a rapid rate. At the beginning of 2019, there were 7,964 deepfake videos online, according to a report from startup Deeptrace; just nine months later, that figure had jumped to 14,678. It has no doubt continued to balloon since then.
According to Sensity, which tracks deepfake videos online, the number has been doubling every six months since 2018, with 85,047 videos detected as of December 2020.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

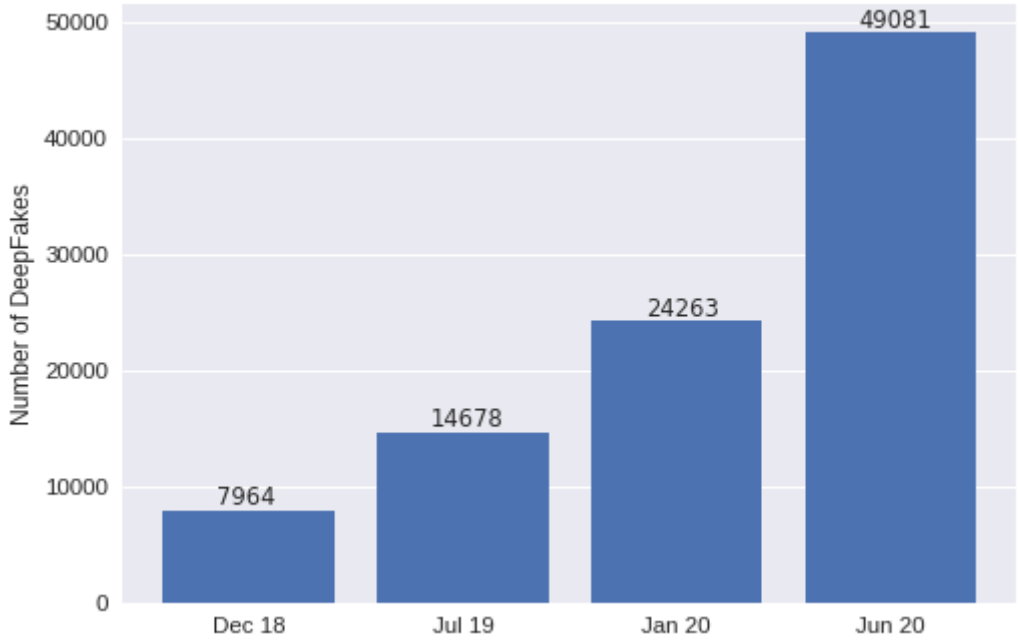


Fig1. Google Trends shows steady global growth in searches for "deepfake" with a significant rise since the beginning of 2021[21].

III. GLOBAL IMPACT OF DEEPFAKE

The global impact of deepfakes is significant and multifaceted, affecting various aspects of society, technology, and public discourse. Here are some key areas where deepfakes have had an impact:

- Misinformation and Fake News: Deepfakes have the potential to spread disinformation and amplify fake news. They can be used to create convincing videos or audio clips of public figures, leading to public confusion, distrust, and the manipulation of public opinion. Deepfakes pose a threat to the credibility of information sources and challenge the notion of visual and audio evidence.
- Political Manipulation: Deepfakes can be employed to manipulate political narratives and elections. By creating fake videos or speeches of political candidates, deepfakes can be used to discredit or defame individuals, influence voter perception, and undermine democratic processes. The potential consequences include the erosion of trust in political systems and heightened polarization.
- Privacy Concerns: Deepfakes raise serious privacy concerns as individuals' likenesses can be used without their consent to create explicit or defamatory content. The ability to superimpose faces onto explicit or illegal activities can lead to reputational harm, cyberbullying, and psychological distress for victims.
- Security and Identity Threats: Deepfakes pose security risks by enabling identity theft and impersonation. Criminals can use deepfake technology to deceive individuals, gain unauthorized access to systems or facilities, or even commit financial fraud. This has implications for cybersecurity, personal safety, and the reliability of biometric authentication systems.
- Media Integrity and Authenticity: Deepfakes challenge the notion of trust in media and raise concerns about the authenticity of visual and audio content. With the increasing sophistication of deepfake technology, it becomes more difficult to distinguish between genuine and manipulated media. This can undermine the credibility of legitimate sources and create an atmosphere of skepticism.
- Legal and Ethical Implications: The rise of deepfakes has sparked debates around legal and ethical considerations. Questions arise regarding the responsibility of individuals and platforms in preventing the creation and dissemination of deepfakes, as well as the potential limits on freedom of expression and artistic creativity.

In response to these global impacts, researchers, policymakers, and technology companies are actively working on developing detection methods, raising awareness, and promoting media literacy. Efforts are also being made to establish legal frameworks and guidelines to regulate the creation, distribution, and use of deepfake technology.

Addressing the global impact of deepfakes requires a multi-faceted approach that combines technical advancements, education, media literacy, and collaboration between stakeholders to mitigate the negative consequences while preserving the positive aspects of emerging technologies.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

IV. POLITICS

Deepfakes have been used to defame and misrepresent renowned politicalpersonalities. Some examples:

- In 2019, a deepfake video of a Belgian politician went viral, showing him making a speech he never delivered. The incident raised concerns about the potential use of deepfakes in political campaigns and the dissemination of misinformation.
- A survey conducted by Pew Research Center found that 77% of Americans were concerned about the potential use of deepfake videos to deceive voters in the 2020 U.S. presidential election.
- **Mauricio Macri**, the President of Argentina from 2015-2019 has had his face swapped with theface of Adolf Hitler in a video that went viral.
- In another video that went viral, **Angela Merkel**, the chancellor of Germany since 2005, has hadher face swapped with Donald Trump.
- In India, in the campaign for the **Delhi Legislative Assembly** in 2020, the **Bhartiya Janata Party**,a prominent Indian political party, used deepfakes in a campaign advertisement in which **ManojTiwari**, a member of Parliament, had his video, which was initially in Hindi, deepfaked toHaryanvi to target Haryanvi voters.
- A voiceover was provided by an actor, and AI trained using video of Tiwari speeches was used to lip-sync the video to the new voiceover. A party staff member described it as a "positive" use of deepfake technology, which allowed them to "convincingly approach the target audience even if the candidate did not speak the language of the voter."
- -**In April 2020**, the Belgian branch of Extinction Rebellion published a deepfake video of Belgian Prime Minister **Sophie Wilmès** on Facebook. The video promoted a possible link between deforestation and COVID-19. It had more than 100,000 views within 24 hours and received many comments. On the Facebook page where the video appeared, many users interpreted the deepfake video as genuine.[6][14]

4.1 Acting

- Virtually made humans were already included in many movie titles and deepfakeshave great potential to contribute stellar developments in the near future. In the world-renownedfilm "Solo: A Star Wars Story," Harrison Ford's young face was deepfaked onto Han Solo's, acharacter in the film.
- **Flawless**,a AI startup helps TV shows ans films effortlessly reach new markets around the world using deepfake dubs as its too distracting to see subtitles dubbed in different languages and mouth movement.As they are entirely out of sync with what they are saying but this AI-powered solution will replace facial performance to match the words in the film dubbed for foreign audiences.[15]

4.2 Fun and Recreation

- Applications like **Zao**, a Chinese app, allow users to substitute theirfaces on famous celebrities or actors in shows like Game of Thrones. In a matter of seconds.
- **In 2020,** an internet meme emerged utilizing deepfakes to generate videos of people singing the chorus of "Baka Mitai," a song from the game Yakuza 0 in the video game series Yakuza. In the series, the melancholic song is sung by the player in a karaoke minigame. Most iterations of this meme use a 2017 video uploaded by user Dobbsyrules, who lip syncs the song, as a template.[10]

4.3Fraud

Audio deepfakes have been used as part of social engineering scams, fooling people into thinking they are receiving instructions from a trusted individual. In 2019, a U.K.-based energy firm's CEO was scammed over the phone when he was ordered to transfer €220,000 into a Hungarian bank account by an individual who used audio deepfake technology to impersonate the voice of the firm's parent company's chief executive.[9]

Many people have started using deepfakes for a lot of other uses as well.Deepfakes have the capability to foil elections by spreading fake information. Even if the fakevideos have been spotted, it would not be long before the video has been circulated thousands oftimes and have been viewed by millions of people who have now formed strong opinions againstthe defamed one.

V. LEGAL ASPECT OF DEEPFAKES WORLDWIDE

This is the question that comes to mind, if deepfakes are so heinous and malicious, then whyaren't they banned yet? The answer is simple, deepfakes are somewhat new, and the law has notcaught up just now. When a famous deepfake video of Mark Zuckerburg emerged on Instagramin 2017, Instagram resisted taking the video down, their head, Adam

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

Mosseri, said, "We do not have a policy against deepfakes currently" he also added, "We are trying to evaluate if wewanted to that and if so, how would you define deepfakes."[7]

It is being worked upon just now that what is the limit at which we draw the line. Social mediaplatforms like Reddit, Twitter, Facebook, and Snapchat have banned any obscene videos fromsurfacing. Still, because they are non-consensual other than that, deepfakes are completely fine andinstead garner much attention from the users.

The cybersecurity of Canada, The Communications Securities Establishment, warned recentlythat there is a significant threat to modern democracy in deepfakes [19]. They wrote in a cybersecurityreport, "Improvements in artificial intelligence (AI) are likely to enable interference activity tobecome increasingly powerful, precise and cost-effective." Deepfakes have also challenged legal systems across the world that are trying to keep themselves abreast of this rapidly evolving technology. In the USA, the Deepfakes Accountability Act (passed in 2019), mandated deepfakes to be watermarked for the purpose of identification. Virginia has also amended its law banning non-consensual pornography from including deepfakes.

In India however there is no explicit law banning deepfakes. Amidst the current laws in force, sections 67 and 67A of The Information Technology Act 2000 ("IT Act") provide punishment for publishing sexually explicit material in electronic form. Section 500 of the Indian Penal Code 1860 provides punishment for defamation, but these provisions are insufficient to tackle various forms in which deepfakes exist.

Another reason why deepfakes are not illegal is because they are often used for good purposes thatprovide excellent value to people. One example is, "the development of deep generative modelsraises new possibilities in healthcare, where we are rightly concerned about protecting theprivacy of patients in treatment and ongoing research. With large amounts of real, digital patientdata, a single hospital with adequate computational power could create an entirely imaginarypopulation of virtual patients, removing the need to share the data of real patients."WritesGeraint Rees, Professor of cognitive neurology, University College London.

Newer and newerways of using deepfakes are emerging day by day, whether it be the creation of the "world's firstsynthesized, presenter-led news reports" by Reuters and Synthesia or the creation of the 'lost'audio of the speech that John F. Kennady, the 35th President of the United States, gave onNovember 22, 1963, the day he was assassinated, or a health charity to have David Beckhamdeliver an anti-malaria message in nine different languages using deepfake technology. The thinghere is that deepfakes are creating value for society and are achieving some great thingswhich could not have been possible had it not been for deepfakes.

VI. SPOTTING DEEPFAKES

Here are some obvious signs that the video that you are seeing is a deepfake [3]–

a. Deepfakes have trouble generating realistic hair, so look for frizz and flyway's.
b. Deepfakes have trouble generating realistic eyes, so weird eye movements or a weird gaze might be a giveaway. An irregular blinking rate might also be a clue.
c. Merging photos sometimes results in double chins, and ghost edgesthat is something to look for.
d. Since deepfakes process the video frame by frame independently, it creates a temporal smoothness problem called flicking. To address that, the makers blur the edges. So, the fake videos will be extra blurry or flickering.
e. Teeth also lie outside the scope of deepfakes so weird teeth are also a red flag.
f. Other techniques use blockchain to verify the source of the media. Videos will have to be verified through the ledger before they are shown on social media platforms. With this technology, only videos from trusted sources would be approved, decreasing the spread of possibly harmful deepfake media.

Detection of deepfake videos is an active area of research, but nothing extremely accurate has come out till now. Two researchers at **UC Berkeley** are working on deepfake detection.**Hany Farid**, a professor in the Department of Electrical Engineering and Computer Science, says, "The basic idea is that we can build these soft biometric models of various world leaders, such as 2020 presidential candidates and as the video starts to break, for example, we can analyse it and determine if the video is fake or not." Their algorithm works by picking up subtle abnormal behaviour and unmatching character traits. Still, their technique is not bulletproof since it only works for famous political figures and for their specific documented behaviour. The Défense Advanced Research Projects Agency (DARPA) of the USA has spent 68 million dollars in the past four years on digital forensic technology to flag deepfake videos[17][20].

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580| A Monthly Double-Blind Peer Reviewed Journal |

| Volume 10, Issue 5, May 2023 |

VII. CONCLUSION

In conclusion, deepfake technology represents a double-edged sword in the digital age. While it offers fascinating possibilities for creative expression, entertainment, and research, its potential for malicious use and deception is undeniable. The global impact of deepfakes spans across areas such as misinformation, political manipulation, privacy concerns, security threats, media integrity, and ethical dilemmas.

However, the challenges of banning deepfakes outright are multifaceted. Striking a balance between addressing the malicious use of deepfakes and upholding principles of freedom of expression is complex. The rapid evolution of deepfake technology, difficulties in detection and attribution, international jurisdictional complexities, and ethical considerations further complicate the regulatory landscape.

Efforts are underway to address these challenges through the development of detection methods, public awareness campaigns, content moderation policies by platforms, and legal frameworks. Collaborative approaches involving governments, technology companies, researchers, and civil society are necessary to tackle the negative consequences of deepfakes while preserving the positive aspects of emerging technologies.Navigating the world of deepfakes requires media literacy, critical thinking, and a cautious approach to the information we encounter. By fostering a discerning mindset, promoting responsible technology development, and encouraging ethical practices, we can mitigate the harmful effects of deepfakes and foster a more trustworthy and informed digital landscape.

REFERENCES

1. Dolhansky, B.; Bitton, J.; Pflaum, B.; Lu, J.; Howes, R.; Wang, M.; Ferrer, C. The deepfake detection challenge dataset. *arXiv* **2020**, arXiv:2006.07397.
2. Akhtar, Z.; Dasgupta, D.; Banerjee, B. Face Authenticity: An Overview of Face Manipulation Generation, Detection and Recognition. In Proceedings of the International Conference on Communication and Information Processing (ICCIP), Talegaon-Pune, India, 17–18 May 2019; pp. 1–8.
3. Westerlund, Mika. (2019). The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review. 9. 39-52. 10.22215/timreview/1282.
4. Deepfake detection by human crowds, machines, and machine-informed crowds, https://doi.org/10.1073/pnas.2110013119
5. Matthew Groh https://orcid.org/0000-0002-9029-0157 groh@mit.edu, Ziv Epstein https://orcid.org/0000-0002-5831-5756, Chaz Firestone https://orcid.org/0000-0002-1247-2422, and Rosalind Picard
6. https://medium.com/@jonathan_hui/how-deep-learning-fakes-videos-deepfakes-and-how-to-detect-it-c0b50fbf7cb9
7. https://www.vice.com/en/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp
8. https://www.techlicious.com/tip/how-to-spot-a-deepfake-video/
9. https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/?sh=5537d97b2f84
10. https://www.thesun.co.uk/news/9379122/what-are-deepfake-videos-and-what-does-the-law-say/
11. https://www.weforum.org/agenda/2019/11/advantages-of-artificial-intelligence/
12. https://en.wikipedia.org/wiki/Deepfake
13. https://library.answerthepublic.net/tag/baka+mitai+meme
14. https://www.newsbreak.com/news/2227739602035/deepfake-lips-are-coming-to-dubbed-films
15. https://informationmatters.net/deepfakes-problem-why-law-needs-to-change/
16. https://timreview.ca/article/1282
17. https://journalism.design/extinction-rebellion-sempare-des-deepfakes/
18. https://dot.la/flawless-ai-dubbing-2652865135.html
19. https://edition.cnn.com/2019/08/09/tech/deepfake-tom-perez-dnc-defcon/index.html
20. https://www.ischool.berkeley.edu/news/2019/hany-farid-race-detect-deepfake-videos-we-are-outgunned.
21. Patel, Mohil& Gupta, Aaryan &Tanwar, Sudeep &Obaidat, Mohammad. (2020). Trans-DF: A Transfer Learning-based end-to-end Deepfake Detector. 10.1109/ICCCA49541.2020.9250803.

Impact Factor: 7.580

INTERNATIONAL JOURNAL
OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com