Oríginal Article

World Journal of Engineering Research and Technology



**WJERT** 

www.wjert.org

SJIF Impact Factor: 7.029

Article Accepted on 17/04/2025



# DEEP LEARNING FOR DEEPFAKES CREATION AND DETECTION

# Dr. Chetanpal Singh<sup>1\*</sup>, Jatinder Warraich<sup>1</sup>, Numan Ahmed<sup>1</sup> and Rahul Thakkar<sup>2</sup>

<sup>1</sup>Future Technologies Industry Cluster College of Vocational Education RMIT University, 124 La Trobe St, Melbourne VIC 3000.

<sup>2</sup>Victorian Institute of Technololgy, Department of Computer Science, 14/123 Queen St, Melbourne VIC 3000.

Article Revised on 28/03/2025

Article Received on 08/03/2025



\*Corresponding Author Dr. Chetanpal Singh Future Technologies Industry Cluster College of Vocational Education RMIT University, 124 La Trobe St, Melbourne VIC 3000.

### ABSTRACT

The technology of deep learning is often used for the resolution of some of the most intricate problems starting from analytic of big data to use of computer vision and in the control of human level applications. Advancements of deep learning is made and applied for the development of software for causing a threat situation to democracy, privacy as well as is risk to nations overall safety. Deep fake is a prominent application of deep learning which is creating havoc in the world by creating false image and videos. It is very hard to distinguish these images from real ones and hence it can be used to

hurt societies and human community. This is also a threat at the national and global level. Hence in this paper many studies are studied to understand how the algorithm works and how detection method can be used to detect the deepfakes. The literature are reviewed which proposes how to effective detect deep fake and it is found that using artificial technologies and block chain this can be controlled to a large extent.

**INDEXTERMS:** Deep Fakes, Deep Learning, Survey, Artificial Intelligence, GAN, Forensics.

### I. INTRODUCTION

Deepfakes is a term which stems down from a combination of words "deep learning" and "dake". In a deepfake application often images are superimposed to victimize a human being

or over video of real individuals.<sup>[1]</sup> Often these are fake cases where one falls victim to the malicious intention of another. Face swap is a common form of deepfake application. Deep fakes when defined in a wider manner are artificial intelligence of AI technology used content used to make lip sync events and puppet master events. The lip sync category of the deepfake application is where videos are changed for creation of mouth expressions and involves an audio running at the background. Puppet master, on the other hand is a deepfake application covering the videos of a human who is the puppet here.<sup>[2]</sup> Facial gesture, eye and other facial movements are impersonated or copied in an animated manner and this is run in the camera front.

The use of the visual effects technology or the computer graphics applications is found in much deep fake creation. The present time is where advanced technologies such as deep learning algorithms, auto encoding technologies and GAN or generative adversarial networks are used for the creation of deep fakes. All these are applied in the domain of computer vision.<sup>[3]</sup> Such models are utilized for the examination of the human being movement and face gestures and expressions and then use synthesis of the facial imageries of other human to develop the analogous movement and expressions. The methodology of the deep fake in general needs massive image repositories so that the algorithm can be given ample training materials so the output imagery or videos is photo realistic.

The celebrity humans and political are the public figures. They are engaged in many public works and their promotions are done mainly through a massive database of image and videos. These are found on the digital platform itself since public figures are engaged socially in many events. Back in the year 2017, the first image which was made using deepfake emerged. Here a very well-known celebrity face was swapped with porn actor. The real threat and risks of this technology was right on the face of the society.<sup>[4]</sup> The world was threatened to know that their image and life is at risks since images made with deepfakes can be used to demoralize them or create false image and impression in the minds of the common man.

Deepfake technology in the wrong hands can create havoc between country and public relations. They can be used to create false statements about religion and culture. Chaos can be the result of such activities in the global market with such news and promotions. The technology can be used for the creation of satellite imagery which is false and show the world components on the planet which has no real existence. This will confuse the military army and defense system in the country. The military can be confused to see sudden bridge

architecture over a river which in reality has no real existence. Such things can be constructed as a bait to harm the defense personnel (Bansal & Joshi, 2021).

Deep fake technology also has some positive applications. Today humans are focusing enormously to create photo relativistic digital imagery and video. The digital avatars for example are used in varied social media platform as avatars representing the real human.<sup>[5]</sup> Use of deep fake coding in such social platform allows the system use real image of the user and create digital avatar which closely resemble the real human. The world of movie making also makes use of deep fake technology. Many events shown in a movie such as shooting a human or a bomb explosion cannot be done in real as it is not practical, legal or ethical. The use of deepfake resolves the dilemma. It provides images and videos as real as it can be. In the world of video games, deep fakes are creating an enormous appeal to the player by giving them realistic experience of the game world. Some other areas where it is used in positive way are in photography, movie making, virtual reality, education entertainment, shopping and others.

It is an irony to find that the deepfake use in negative and malicious events outperforms its positive applications. Creation of this advanced technology and the data availability in digital platforms leads to creation of a channel where the model learns extensively about human. The computer platform and algorithms are made stronger and thus they are able to create image and videos of human which are as close to a real photo. Social media is where the humans share images and videos in different formats. Hence they fall target to such malicious schemes. Still imagery is used today for the creation of the deep fake applications based on the current advancements.<sup>[6]</sup> The use of this technology is indeed a threat not just too any public personnel but also to the common humans.

Here is another example where a CEO was scammed a massive sum of 243,000 where voice based deep fake application was made.<sup>[7]</sup> DeepNude is software released where threatful imagoes and videos are created where a real human being is mapped by the technology for creating porn content.<sup>[8]</sup> The Chinese application called Zao is in viral focus in today's world where users can swap their own face with the movie celebrity and star bodies. These common users can hence use their own self in real movies and share those clips. This is a massive threat caused which compromises the privacy and confidentiality of the real person and impacts life and property massively.<sup>[9]</sup>

For the digital platform it is now extremely critical to discover the truth. This is a major challenge since deep fakes are used widely today for many malicious intentions. Today applications are easily found and free tools are used to create such content by anyone. Hence detecting of the deep fake is important. These technologies are based over the technology deep learning. There is a constant tug of war between its positive and negative application of the methods of deep learning (Fouzi et al., 2023).



Figure 1: Paper of deep fake in year range 2016-2021.<sup>[1]</sup>

### (Taken from https://app.dimensions.ai)

DARPA or Defense Advanced Research Projects Agency of the United States began research and develop in the field of media forensics and this is called as MediFor or Media Forensics to address this use of swapping faces using deep fake technology.<sup>[10]</sup> The detection process and systems to detect this technology uses is enhanced. The company Facebook Inc. is partnered with the Microsoft Corporation to start the challenge of DeepFake Detection. This has enhanced the eagerness of the public to identity deep fake and use ways to prevent it from happened. Data is collected for year end 2021 where deep fake related papers which is published in current years is shown (Fig 1). These paper numbers are however less than the real number which is present in real since the research is happening continuously.

Many papers exist where survey is made with regards to the creation and the identification of the deepfake.<sup>[11]</sup> A paper is analyzed which is researched about the re-enactment techniques, replacement techniques used in deep fake detection.<sup>[12]</sup> In reenactment technique the moth, expression, body, eye look or gaze and expression is changed while in the technique of replacement the face of victim is swapped. Such techniques of detection of the deep fake are differentiated based on traditional process and using deep learning processes.<sup>[13]</sup> Some other

paper makes classification of formation and identification using the technique used to form the deep fake such as swapping identity, synthesis of face, changes in attribute and in swapping of expressions.



Figure 2: Paper categorization in dividing them in – fake imagery detection and video detection.<sup>[2]</sup>

In Fig 2. It is witnessed that this is another survey where papers of deep fake are reviewed and then divided in image based and video based detection. This is to use a varied perspective on the subject and in its taxonomy. The fake imagery based detection process is where more and more focus is given over the features being used. This is to analyze whether they are crafted in hand. Some sub category which is detected here is based on whether it is using temporal characteristics in frames or using visual based artifacts in video. The challenges are also used along with the trends and the detection direction and the challenges of forensics in multimedia (Chen et al., 2022).

### A. Research Objective

The purpose of this study is to analyze and classify the methods utilized in the production and identification of deepfake technology in detail. This includes analyzing the techniques used in handling images and videos such as lip sync and puppet master applications as well as evaluating the progress in the areas of deep learning algorithms, autoencoding, and GANs. In addition, the research shall endeavour to assess the positive and negative uses of deep fake technology in areas like media forensics, social networks, in movies and making digital avatars, with a focus on increasing the detection of such technology.

### **B.** Research Motivation

The primary rationale for this study is rooted in the increasing complexity and availability of deepfake technology that threatens personal rights, safety, and credibility. Deepfakes for various purposes generating fake and dangerous content, political manipulations, or identity

theft prove the necessity for the development of efficient protection mechanisms. Furthermore, analyzing the capability of the technology being beneficial in entertainment, education, and digital media brings the necessity of the research enforcement studying the negative impact while recognizing and developing the positive applications (Gamage et al., 2022). Thus, the study seeks to enhance the creation of better forensics and detection solutions that will help protect citizens and the society at large from the negative impacts of deepfakes.

### II. LITERATURE REVIEW

### A. Reviews on the Application

Deepfake has emerged as a popular subject in academic and public discussion because of the sprawling possibility impacts a different aspect of society. Research reveals a range of examples that demonstrates this duality, pointing to various positive and negative ways OM is used. On the benign side of deepfakes is when people use deepfakes in entertainment, for instance, to develop lifelike characters in movies and computer games. For instance, deepfake technology made it possible to bring back actors who are now late or made characters continue growing or regress in age making visual stories or movies more interesting.

On the other hand, deepfake as a tool for spreading fake news and scamming has caused much concern. Videos in the form of deepfake with similar policies to the real news have made some concerns about the deepfake which might pose as fake yet easy to convince the general public leading to social disorder. The effects are even more political, such as cases in which deepfake has been employed to interfere with an election campaign or to slander a leader. The academic reviews suggest that the idea of regulation is essential to establish the right level that fosters innovations while at the same time protecting the general population. As argued by scholars, regulatory measures should aim at developing principles good at identifying deceptive disinformation techniques, raising the public's awareness and increasing the effectiveness of legal action against their utilization while not hindering technology progression.

# B. Algorithms of the Application

This kind of technology has undergone significant improvement in the algorithms used in the creation of deepfakes and the primary basis of this form of development has been the GANs. It was proposed in 2014 by Ian Goodfellow and his team and is made of two predominant neural networks, often referred to as the generator and discriminator that play a zero-sum

game. This is a completely revolutionary architecture which has really been very successful in generating realistic synthetic data.

As for the proofs, the authors focused on early works that introduced theoretical approaches to GAN that later became the foundation for additional practical applications. One of the newest solutions that attracted much attention is StyleGAN which is developed by NVIDIA, and which utilizes a new architecture that provides a vast improvement in the quality of generated images. Breaking free from the limitations of ProGAN's quality and resolution, StyleGAN brings about a host of liberation to GANs.

Other improvements are extensions, for example, Conditional GANs (cGANs), where the generation process depends on additional information, used to interpret the received data; cGANs are employed in image-to-image translation and video synthesis. Another modification that can be mentioned is CycleGAN which enables translation of images without reference to the paired data which can be beneficial for such tasks as converting photos into the paintings or the opposite (Fouzi et al., 2023). The literature also provides a detailed examination of these improvements and specifies the key features, and fields where deepfake methods are most effective, which demonstrates the effectiveness and further development of deepfake algorithms.



Figure 3: Two pair of encoding and decoding made.<sup>[3]</sup>

### C. Tools Used

Deepfake tools has become widespread and they are available to anyone to use and develop and new tools are created by professionals. This section highlights some of the main tools that are used in creating deep fakes as well as those used in detecting them. FakeApp, one of the earliest apps was used to create deepfake videos with comparative ease. It was free to use as it was an open-source that made it popular among hobbyists but it was soon outcompeted by upgraded versions. Another similar open source project is FaceSwap which provides more customization and management over the faces in still images and videos which can be easily and accurately swapped. Regarding the best programs for faceswap, DeepFaceLab has become one of the most popular tools as it offers a wide range of features and works stably. It also supports the creation of multiple deepfake generating processes like face swapping, face reenactment and attribute change. An active community works to enhance the platform based on its users' input to guarantee it remains optimized in terms of deepfake applications.

StyleGAN by NVIDIA is notable for their high image quality of synthetic images production. However, it has been used and tweaked to create realistic avatars, images, and virtual characters which are hallmarks of deepfake. The article compares these tools in terms of interface, performance, and the complexity of algorithms and gives a repertoire of deep fake technologies at the present stage (Boutadjine et al., 2023).

## D. Scope of Deepfake Technology

Thus, the depth of deepfake technology does not stop at image and video forgeries only. Current works predict a wide coverage of uses, including synthetic media and augmented reality, voice deepfakes, and more. Deepfake is one of the applications of synthetic media that has almost transformed content creation. Sponsored content and advertisement figures that have avatars in the virtual world seem to gain popularity. These synthetic personalities can be extremely well thought out and focused on projecting specific character traits, which may appeal to the intended followers more than real-life everyday influencers.

Virtual reality is another domain where the use of deepfake technology is quite promising. The proposed VR application, in particular, the Deepfake technology used, can make the experience more real and engaging for users. For instance, moderately interactive realistic avatars driven by deep fake algorithms can reproduce users' facial and bodily gestures synchronously that augments reality in a livelier way.

Deepfake voice synthesis is somewhat a new and promising area of the AI development in which an artificial voice is created in such a way that it sounds like a real person speaking. It can be used for producing narratives of the videos, producing dialogues of the smart products, and can even clone the voices of the dead personalities for the educational purpose. The ramifications to accessibility are strong as it can create more essay and natural sound for people with speech disorders (Chen et al., 2022).

The literature also outlines the ethics and the society where these advancements will take place. Currently, there are debates and discussions on the necessity of setting specific ethical norms to regulate the application of deepfake. Scholars call for integrated approach, which involves both the use of IT tools and legal intervention and ordinary people's recognition of the possibilities to misuse.

### E. Future Directions and Challenges

As for what the future holds for the field of deepfake technology there are great opportunities and daunting prospects. AI and machine learning is set to progress significantly and it only means that there will be further and better developments to come in the future, making these applications more lifelike and fully functional. But such development is not without hardships.

Another major issue can therefore be identified as the need to create efficient technologies for early detection of tumors and cancer. As deep fakes continue to improve one has to wonder how one can tell a fake from the original. To prevent this, researchers are coming up with different strategies of solving the problem, including the use of forensic analysis and detection ideals which utilizes artificial intelligence (Bansal & Joshi, 2021). But the tendency in which deepfake creators and detectors play the cat-and-mouse will continue.

Another important field is the ethical and legal one. There is a need to set rules and laws that would allow the incorporation of advanced technologies in industries while at the same time putting in place mechanisms to prevent their use inappropriately. It entails defining procedures of checking the content's authenticity, defending the rights of individuals, as well as addressing the issue of punishment for misuse.

Conclusively, from the existing literature, deepfake technology offers an elaborate insight into the practicable adoptions, algorithms, tools, and prospects into the future. Consequently, the advantages are enormous, but the threats and issues are equally immense. The implementation of this technology will be equally important as innovation should be supported without compromising the welfare of society.

#### **II. METHODS AND MATERIALS**

#### A. How Does This Work

Deepfake creating is mostly dependent on deep learning, or specifically, GANs – Generative Adversarial Networks. A GAN consists of two neural networks: The two main parts of the network the generator and the discriminator which are trained in an adversarial manner. The generator synthesizes the fake images with the help of distribution of the training data, whereas the discriminator decides whether these images are real or fake.

1. Training Process: The process training starts with the generator creating random images. This is how the discriminator evaluates the given images and informs the user if they seem real or fake. This feedback is used to make alteration to the parameter of the generator where the generator has the potential of producing realistic images over time's feedback (Boutadjine et al., 2023). The discriminator also strengthens, aiming to perfect the recognition of fake images from the actual ones. Thus, the game between the generator and the discriminator proceeds in cycles until the generator synthesizes images that the discriminator cannot distinguish from the authentic images.

2. *Data Requirements*: Sophisticated deepfakes depend on large database of actual pictures for learning the patterns and models. Described datasets for facial deepfakes are CelebA and VGGFace. These datasets incorporate a large variety of facial expressions, angles, and lighting which is crucial when making realistic deepfakes.

*3. Computational Resources*: The training of the GANs is generally very computationally expensive and sometimes entails the use of GPUs and large memory. It takes several days to several weeks from the conception of the model and the data set to crack down on the best solution. The generation of today's deepfake often utilizes cloud computing to offer required computational resources (Gamage et al., 2022).

### **B.** Advantages

Deepfake technology offers numerous advantages across various domains:

1. Entertainment: In the entertainment industry deepfakes will change the face of CGI and special effects completely. They allow for production of very convincing graphics and that is why actor's faces can be easily grafted to body doubles. It has been applied in minimizing actors' ages, get actors who have died and bring them back for performances, and indeed invent new people. For example, in the movie "Rogue One: Such cases as in film, "Solo: A

Star Wars Story," deepfake technology was used to create a young Princess Leia thus giving a retro to movie and amazing to watch.

2. *Education:* This locations also called deepfakes can enhance educational interactions and presentation. Effective computer animated teachers and personalities of notable historical figures can be used to give students individual lessons which can also be fun. For instance, a deepfake of Albert Einstein could be created to deliver knowledge about physics, physics classes would be more engaging and easy to grasp.

3. *Healthcare:* In healthcare, its application is quite beneficial since it can help in developing convincing scenarios that mimic the real-life situations that healthcare professionals are likely to encounter. Increased Laparoscopic surgical manipulations may be performed on a simulation model of the human body hence increasing the surgeon's expertise without jeopardizing the lives of real patients. Deep fake voices can be useful in speech therapy needed in the generation of natural sounds for those with speech disorders (Bansal & Joshi, 2021).

### C. Importance

Control over deepfake technology is a necessity to gain advantages where it is possible and to avoid the dangers that come together with its usage. This progress is forging to change the lives of different companies and sectors, including entertainment, education, and healthcare. While great strides have been made in improving the safety, reliability, and functionality of the technology, it is also capable of being used maliciously, sparking sorely notable social and ethical repercussions.

*1. Advancement of Technology:* Further investigation of possibilities in fake face generation is required to advance the field further. The problem focuses on the development of generating high-quality and realistic content for different target purposes and enhancing the approaches that make the technology more offered and intuitive.

2. *Robust Detection Mechanisms*: Just as important as its creation though is the establishment of reliable preventive measures to detect deepfake. Over the past few years, deepfakes have continued to become more complex and accurate looking, making it even harder to differentiate between what is real and fake. Further research in explaining and mitigating the occurrence of both fake news and privacy violations is crucial.

3. Societal and Ethical Imperatives: It is imperative to underscore the importance of the societal and ethical value of engaging in research related to deepfake technology. In the same breath it shows that the proper utilization of this technology has the potential of leading to positive effects while on the other hand if used inappropriately it has the potential of causing harm (Chen et al., 2022). It is vital to consider the importance of ethic regulation and safeguard measures to protect society while enabling appropriate advancements in the use of technology such as deepfakes.

### **D.** Applications

Deepfake technology has a wide range of applications, both positive and negative:

### 1. Positive Applications

*Film and Entertainment*: Duplicating real life models, realistic impacts, and keeping the memory for the past actors.

*Virtual Reality:* Improving the VR experiences with avatars which reproduces the user's face and body movements in kinesthetic real-time.

*Educational Tools:* Exploring a highly interesting and exciting approach to learning through the incorporation of virtual teachers and characters of historical figures.

### 2. Negative Applications

*Misinformation:* Deepfakes can make fake news videos through the simulation of people's faces and voices, thus deceiving citizens and stability of society.

*Fraud:* This is the process of forging new and realistic identifies for the purpose of perpetrating fraud, impersonation and the likes.

*Privacy Invasions:* Disseminating deepfake videos for sexual purposes without the consent of the Target such as the aspect commonly known as 'revenge pornography'.

### E. Challenges

The creation and detection of deepfakes present significant challenges:

### 1. Creation Challenges

*High-Quality Outputs:* To guarantee that the synthesized deepfakes are fully convincing as the actual content, significant algorithms and huge datasets are needed. This truly requires a lot of computation and requires resources that may not be easily accessible by just anyone.

*Ethical Use:* The effective use of deepfakes while at the same time ensuring that the technology is not exploited has always remains a hard task. Thus, the principles of ethical behavior and norms must regulate deepfakes' production.

### 2. Detection Challenges

Advancements in Deepfake Quality: It is sad to say that with higher and higher advances in deepfake, they become harder and harder to detect. Experts are working on high accuracy detecting algorithms, but the cat and mouse scenario of creators and detectors persists.

*Regulatory Measures*: The use of deepfakes needs to be monitored effectively and that is why regulatory measures are needed. This includes legislation that enhances on the production and sharing of novel deepfakes, as well as raising the consciousness of people and media literacy in order to be able to identify fake deepfake information.

### 3. Technological Innovations

*Forensic Analysis:* They are digital rights management tools that can ensure that content that is distributed online is original and can track the images and videos' origins.

*AI-Driven Detection:* Deepfakes are currently often detected based on dissimilarities with the original footage, which are hard to detect, but machine learning algorithms in development are identifying further inconsistencies in the video, from the movement of the eyes to changes in lighting. All these tools are very important when it comes to the continuing fight against deepfake makers.

### **IV. RESULTS AND DISCUSSION**

### A. Results

### Performance Metrics of GAN Architectures



Figure 4: GAN Architectures.<sup>[4]</sup>

It emerged that GAN architectures' efficiency exhibits marked differences. Standard GANs are the first to have been proposed by the researchers in 2014; however, newer models such as StyleGAN and StyleGAN2 have been developed. The StyleGAN model proposed by NVIDIA performs well, particularly in high-resolution samples, and set new FID record. Visual artifacts, as well as the quality of generated images, are reduced in StyleGAN2 (Fouzi

et al., 2023). Conditional GANs (cGANs) and CycleGANs showed an improved performance in particular applications such as image-to-image conversion and artistic style transfer.

#### Effectiveness of Detection Methods

The detection methods have also been developed in parallel to deepfake technologies as they advance. The previous approaches were based on detecting contrasting high and low image areas, the later approaches use deep learning for temporal and spatial features in-coherencies, physiological signs and pixel-level contradictions (Boutadjine et al., 2023). CNNs and RNNs demonstrate high accuracy in recognizing deepfakes due to the focus on artifacts and distinctions in temporal sequences. Multi-modal detection techniques that involve the use of visual, auditory and contextual information provide credible countermeasures to deepfakes.

### Comparative Analyses of Deepfake Tools



Figure 5: StyleGAN based mix styles.<sup>[5]</sup>

Deep learning studies show that there is a variation in the functionalities offered by different deepfake tools and the quality level of the generated outputs. Simple tools such as FakeApp and FaceSwap, while quite useful for hobbies are less sensitive in the output quality than the DeepFaceLab which supports facial swap, reenactment, and attribute swap. While not a deepfake application, StyleGAN reaches comparable levels of realism when generating synthetic images and avatars.

#### **B. DISCUSSION**

### Ethical and Legal Challenges

Evaluating the impacts of deepfakes, we can identify numerous ethical and legal concerns that appear due to this technology, for instance, manipulating the information, fraud, and violation of privacy. Many of them are insufficient in terms of addressing new developments, requiring new provisions concerning consent, intellectual property, or liability. There is a need to create an elaborate legal regime that will ensure the protection of IP assets and promotion of innovation at the same time.

### Effectiveness of Current Detection Methods

Nevertheless, methods to detect deepfake scenarios should remain improved to cover emerging types of deepfake video. The development of highly sensitive and specific biosensors for the detection of pathogens requires interdisciplinary efforts from academic institutions, industries, and governments. Recognizing deepfakes and judging their credibility is important where public awareness and education are adopted (Chen et al., 2022).

### Role of Policy in Managing Deepfake Technology

Effective policies should only encourage innovation that is sensible and useful; at the same time, they have to prevent misuse. Blockchain and watermarking as the methods can ensure media credibility and can fight against fake deepfake videos circulation.



Figure 6: Detection of manipulation of face using 2 phase process.<sup>[6]</sup>

### Future Trends

Such future trends will include artificially intelligent countermeasures, adversarial training, and assimilation of deepfakes into the mainstream media, there is a dire need to develop ethics into the frame work. The development of deep fake technology will require adaptive strategies that encourage innovation but also offer safeguards of public interest.

### V. CONCLUSION

### A. Future Work

Further research on deepfake technology needs to improve the existing approaches to creation and detection of fake videos. Thus, creation of complex models of artificial intelligence that can recognize visual, audio as well as context-based errors is vital. The use of multiple detection techniques along with large datasets to alleviate the issues associated with accuracy. To mitigate the effects of deepfakes, the detection tools can be installed directly into the social media applications (Gamage et al., 2022). AI, ethics, and law are important areas of the study as they create a synergy to address the challenges posed by innovation while being socially useful.

### B. Open Research Questions

Key research questions include

*1. Robustness Against Adversarial Attacks*: How to improve the models to tackle new methods of deepfakes?

2. *Effective Regulation*: The delicate equilibrium between strictly enforcing legislation against deepfake technology and ensuring that it is developed for the right purpose.

*3. Blockchain and Content Authenticity:* A process of proving the authenticity of digital content with the help of blockchain.

4. *Ethical Considerations*: Ethics fail to advance at the same rate with the technology, and this makes it difficult to set the right standard to be followed in the social media.

5. *Public Awareness and Education:* Reducing the level of people's trust in images and raising awareness to question the credibility of information present online.

#### REFERENCES

- G. Bansal and M. L. Joshi, "DEEPFAKE: A SYSTEMATIC REVIEW," Journal, 2021; 1(1): 1–2.
- A. K. Singh, "Deep Learning for Deepfakes Creation and Detection: Report," Journal, 2024; 1(1): 1–2.
- 3. T. Shen, R. Li, and J. Bai, "Deep Fakes' using Generative Adversarial Networks (GAN)," Journal, 2023; 1(1): 1–2.
- T. Walczyna and Z. Piotrowski, "Quick Overview of Face Swap Deep Fakes," Applied Sciences, Jan. 2023; 12(11): 6711.
- N. Caporusso, "Deepfakes for the Good: A Beneficial Application of Contentious Artificial Intelligence Technology," Advances in Intelligent Systems and Computing, Jul., 2020; 1213: 235–241.
- Ángel Fernández Gambín, Anis Yazidi, A. Vasilakos, H. Haugerud, and Youcef Djenouri, "Deepfakes: current and future trends," Artificial Intelligence Review, Feb. 2024; 57(3).
- T. T. Nguyen et al., "Deep learning for deepfakes creation and detection: A survey," Computer Vision and Image Understanding, Jul. 2022; 223(103525): 103525.
- 8. A. Boutadjine, Fouzi Harrag, Khaled Shaalan, and S. Karboua, "A comprehensive study on multimedia DeepFakes," Mar. 2023.
- D. Gamage, J. Chen, P. Ghasiya, and K. Sasahara, "Deepfakes and Society: What Lies Ahead?," Frontiers in Fake Media Generation and Detection, 2022; 3–43.

- 10. L. Verdoliva, "Media Forensics and DeepFakes: an overview," Journal, 2020; 1(1): 1–2.
- 11. A. A. Abu-Ein, O. M. Al-Hazaimeh, A. M. Dawood, and A. I. Swidan, "Analysis of the current state of deepfake techniques-creation and detection methods," Indonesian Journal of Electrical Engineering and Computer Science, Dec. 2022; 28(3): 1659.
- F. Abbas and Araz Taeihagh, "Unmasking deepfakes: A systematic review of deepfake detection and generation techniques using artificial intelligence," Expert systems with applications, May 2024; 124260–124260.
- 13. A. Heidari, Nima Jafari Navimipour, H. Dag, and M. Unal, "Deepfake detection using deep learning methods: A systematic and comprehensive review," Wiley interdisciplinary reviews. Data mining and knowledge discovery/Wiley interdisciplinary reviews. Data mining and knowledge discovery, Nov. 2023.
- 14. L. Whittaker, K. Letheren, and R. Mulcahy, "The Rise of Deepfakes: A Conceptual Framework and Research Agenda for Marketing," Australasian Marketing Journal, Mar. 2021; 29(3): 183933492199947.
- 15. H. Vyas, "Deep Fake Creation by Deep Learning," nternational Research Journal of Engineering and Technology, 2021; 7(7): 1–2.
- 16. I. Rakhmatulin, "Cycle-GAN for eye-tracking," arXiv (Cornell University), Jan. 2022.
- 17. A. H. Bermano, "State-of-the-Art in the Architecture, Methods and Applications of StyleGAN," Journal, 2022; 1(1): 1–2.
- Aakash Varma Nadimpalli and Ajita Rattani, "ProActive DeepFake Detection using GAN-based Visible Watermarking," ACM Transactions on Multimedia Computing, Communications, and Applications, Sep. 2023.
- S. C. Kumain, M. Singh, N. Singh, and K. Kumar, "An efficient Gaussian Noise Reduction Technique For Noisy Images using optimized filter approach," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Dec. 2018.
- 20. None Miao Zhang and None De-xian Zhang, "Trained SVMs based rules extraction method for text classification," Dec. 2008.
- 21. T. Duc, "Deep Learning for Deepfakes Creation and Detection: A Survey," Journal, 2020; 1(1): 1–2.
- 22. A. A. Agarwal, "FaceOff: A Video-to-Video Face Swapping System," Journal, 2022; 1(1): 1–2.
- 23. Y. Bian, J. Wang, Jaden Jungho Jun, and X. Xie, "Deep Convolutional Generative Adversarial Network (dcGAN) Models for Screening and Design of Small Molecules

Targeting Cannabinoid Receptors," Molecular Pharmaceutics, Oct. 2019; 16(11): 4451–4460.

24. T. Zhao, X. Xu, M. Xu, H. Ding, Y. Xiong, and W. Xia, "Learning Self-Consistency for Deepfake Detection," Oct. 2021.