Artificial Intelligence into Multimedia Deepfakes Creation and Detection

Moaiad Ahmed Khder
Department of Computer Science
Applied Science University
Bahrain
moaiad.khder@asu.edu.bh

Samer Shorman
Department of Computer Science
Applied Science University
Bahrain
samer.shorman@asu.edu.bh

Duaij Thabit Aldoseri
Department of Computer Science
Applied Science University
Bahrain

Mohammed Mansoor Saeed
Department of Computer Science
Applied Science University
Bahrain

2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD) | 978-1-6654-6372-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ITIKD56332.2023.10099744

***Abstract*— Artificial intelligence has enabled deepfakes, or fake videos that closely resemble real ones, commercially viable (AI). As according to our research, people are less likely to believe social media news, but they are also more likely to be dubious than to be fooled by deepfakes. This finding combines ideas about how effective visual communication is and how uncertainty weakens public discourse confidence. The methodology for this article will based on reviewing the related articles to explore the main components and solutions in the deepfakes. In this article will study and review the deepfakes and it impacts in different multimedia with AI tools. Therefore, that conclude that deepfakes may increase already existing hazards to countries and online civic cultures by increasing general mistrust and doubt in the most of online resources.**

Keywords*— *Artificial intelligence, deepfakes, Photo deepfakes, Audio deepfakes, video deepfakes

I. INTRODUCTION

Deepfake videos make it possible to swap out a human personality in a photo or video. With the aid of continually advancing technology, these films were produced. The false videos get more realistic because of the algorithm's learning and improvement of the subject's motions, speech, and changes. If the algorithm has access to enough video and audio on the topic as a starting point, it may create a phony video in which the person says things they haven't stated. Soon, it won't be possible to distinguish between these videos and real footage with the human eye. [1]

This type of "computer-generated imagery," or computer imaging, technology was previously solely available for Hollywood's high-budget film productions (CGI). Like other technologies, the usability of this one is enhanced for consumers by faster CPUs, powerful graphics cards, and cleverer algorithms. Anyone can now make face-swap movies by downloading the Deepfake app and following the program's video lessons. The video clips can feature anyone's face and can be used to create X-rated videos. They originate from sources like interviews or images used in advertisements. This essay concludes by discussing the issues Deepfakes provide to legal experts and criminal justice professionals as well as how they affect the probative value of video evidence. [2, 3].

This study is organized as the following: the current section is introduction which contains the literature review, deepfake which explain the main concepts, artificial intelligence, Artificial intelligence, deepfake, deepfake creation, deepfake detection, types of deepfakes, deepfake impact and uncertainty, , some benefits of deepfake, proposed solution and conclusion.

II. ARTIFICIAL INTELLIGENCE

The ability of a technology operated by a computer to perform activities that intelligent individuals typically complete This word is typically used to refer to efforts to create artificially intelligent systems with cognitive capacities similar to that of people, such as the ability to think, find meaning, generalization, and learn things. Ever since advent of the digital computers with in 1940s, it has been demonstrated that computers can be taught to accomplish extremely difficult jobs, such as finding proof for logical theories or play expert chess. [4] There are currently no systems that can replicate person flexibility across a larger range of tasks or those needing a significant amount of previous knowledge, despite continual improvements in computer processor speed and large memory. In this specific sense, artificial intelligence is employed in a variety of contexts, such as speech or handwriting, internet google search, and medical diagnostics. On the contrary, some computers can do tasks as well as specialists and professionals. [5].

III. DEEPFAKE

To make a video of the target acting or saying things the sources person says, deepfakes are made utilizing methods that may overlay face photographs of the target person onto such a video of the source person. Deepfakes is a subgenre that includes face swapping. A more basic definition of deepfakes is artificial intelligence-produced entertainment that can also be lip-synced or puppeteer. Lip-sync deepfakes are videos which have been edited to match the audio and the mouth motions. Videos of an aiming (a puppets) that are generated to mimic the expressions, eye movements, and head movements of a different subject (a master) sitting in front of the camera are known as puppet-master deepfakes [2]. Although some deepfakes can still be created using traditional visual effects or computer graphics methods, machine learning designs like auto - encoders and formative adversarial systems (GANs), that have been widely used in the field of computer vision [2-8], now serve as the similar underlying mechanism for creating deepfakes. After analyzing a person's facial emotions and movements, these models are utilized to create face pictures of various people with similar emotions and movement patterns [9] [10]. Models are being taught how to create lifelike videos and images using deepfake methods. Because there are so many films and pictures of well-known people, such as politicians and celebrities, deep fakes frequently begin with them.

IV. AI AND DEEPFAKE

Artificial intelligence allows computer systems to learn directly from examples, data, and experience etc. As will carry out complex processes by learning from data, rather than following instructions to be used to create deep-fake films (simulate human behaviour). Computational representations of human behaviour and mental processes that are supposed to work logically and Intelligently are what artificial intelligence (AI) and machine learning. AIs continually improve their performance by learning from past mistakes and changing their behaviour in response to both new inputs and past performance[11]. The development of AI technology is still in its early phases. Siri from Apple, Alexa from Amazon, Nest from Google, and Pandora's automatic music recommendation service are examples of first-generation AI technology. These computer systems use machine learning technology, which recognizes questions and requests made in spoken English and provides database responses.

During 2015, Google released their accessible TensorFlow machines learning and image analysis engine. TensorFlow is said to have been used by Google translation and Gmail teams to interpret messages they receive and anticipate the responses they would provide (referred to as "smart replies"). Diabetic retinopathy, the most common cause of blindness in people with diabetes who are working age, has been predicted by physicians using TensorFlow. To train TensorFlow to recognize diabetic retinopathy, the Google team fed TensorFlow a set of images that had been labelled and grouped by ophthalmologists. Medical professionals can observe a brand-new image thanks to the device. TensorFlow can identify images that have diabetic retinopathy by comparing them to other images in its collection. The technology allows doctors to see a brand-new image. TensorFlow can determine whether a picture has diabetic retinopathy by comparing it to others in its database. Although TensorFlow is a useful tool for machine learning and image learning, the source code TensorFlow frontend has been utilized to create Deepfake videos. [12].

V. DEEPFAKE CREATION

Deepfakes has gained popularity because of the high quality of the changed videos and the simplicity of its applications for a wide range of users—from professionals to beginners. These apps are mostly created utilizing deep learning techniques. Deep learning has a proven ability to represent complex, high-dimensional data. For dimensionality reduction and picture compression, deep autoencoders, a deep network variant with such capabilities, are commonly utilized [13]. The first attempt at deepfake creation was a Reddit user's autoencoder-decoder pairing structure-based Fake program [7, 14]. After the autoencoder has retrieved latent features from the face images in that technique, the decoder is used to reconstruct the face photos. To switch faces between source and target photos, two encoder-decoder pairings are needed, each trained on a distinct collection of images. The two network pairs share the encoder's settings. In other words, two pairings' encoder networks are identical. The common encoder can identify and learn the differences between two sets of face images using this method since features similar like the placement of the eye, nose, and lips are frequently present in face photos.

VI. DEEPFAKE DETECTION

Deepfakes are getting worse for democracy, social security, and personal privacy [15]. Deepfake detection methods were released as soon as this threat manifested. Early attempts relied on characteristics that were produced utilizing mistakes and flaws while fabricating video synthesis. However, the use of deep learning in more recent approaches made it possible to automatically extract prominent and distinctive qualities that made it possible to detect deepfakes. Deep fake detection is typically thought of as a binary classification issue where classifiers are used to distinguish between real movies and ones that have been altered. For the classification algorithms to be trained using this approach, a sizable collection of actual and fraudulent videos is needed.

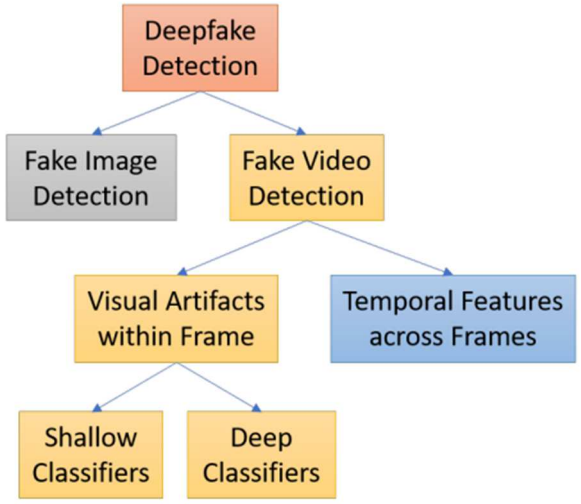

Fig. 1. Categories of peer-reviewed articles that are relevant to deepfake detection techniques, with publications broken down into fake picture identification and face video identification as the two primary subcategories.

VII. TYPES OF DEEPFAKES

The early incubation stage of artificial intelligence-based tools for creating false material is soon giving way to a period of rapid development and enhanced performance, like previous technologies. Table 1 [7] summarizes the different deepfakes and the potential commercial applications that could emerge as their underlying technologies advance.

TABLE I. TYPES AND EXAMPLES OF DEEPFAKES

Type	Description	Current example	Businessapplication
Photo deepfakes	Face and body-swapping Making changes to a face, replacing or blending the face (or body) with someone else's face (or body)	FaceApp's aging filter alters your photo to show how you might look decades from now	Consumers can try on cosmetics, eyeglasses, hairstyles, or clothes virtually
Audio deepfakes	Voice-swapping Changing a voice or imitating someone else's voice	Fraudsters used AI to mimic a CEO's voice and then tricked a manager into transferring $243,000	The voice of an audiobook narration can sound younger, older, male, or female and with different dialects or accents to take on different characters
	Text-to-Speech Changing audio in a recording by typing in new text	Users made controversial Jordan B. Peterson, a famous professor of psychology and author, say anything they wanted until his threat of legal action shut the site NotJordanPeterson down	Misspoken words or a script change in a voiceover can be replaced without making a new recording
Video deepfakes	Face-swapping Replacing the face of someone in a video with the face of someone else	Jim Carrey's face replaces Alison Brie's in a Late Night with Seth Meyers interview.	Face-swapped video can be used to put the leading actor's face onto the body of a stunt double for more realistic-looking action shots in movies.
	Face-morphing A face changes into another face through a seamless transition	Former Saturday Night Live star Bill Hader imperceptibly morphs in and out of Arnold Schwarzenegger on the talk show Conan	Video game players can insert their faces onto their favorite characters
	Full-body puppetry Transposing the movement from one person's body to that of another	"Everybody Dance Now" shows how anyone can look like a professional dancer	Business 1leaders and athletes can hide physical ailments during a video presentation.
Audio & video deepfakes	Lip-syncing Changing the mouth movements and words spoken in a talking head video	In "You Won't Believe What Obama Says In This Video!" Jordan Peele edits Obama to use profanity in a public service announcement	Ads and instructional videos can be 'translated' into other languages using the same voice used in the original recording

Like other technologies, fake content creation tools based on AI are quickly moving from their early incubation stage to a time of rapid growth and improved performance. The deepfakes that were displayed in Table 1 (Photo, Audio, Video, and Audio & video) highlight the many sorts of deepfakes and the targeted applications that will evolve as their fundamental methods mature.

VIII. DEEPFAKE IMPACT AND UNCERTAINTY

Deepfakes' potential cognitive results will mislead people; that is the problem's clear central issue. Even if viewers are not tricked by a deep fake, they could start to question whether the content is real or fake. Ambivalence and uncertainty are conceptually different. When people are given a decision on which they disagree, they become ambivalent, and "new information merely heightens the inner struggle" [5] [16]. The introduction of fresh information can eliminate ambiguity, which is experienced when there is not enough information available to decide [5]. According to Downs (1957), citizens experience uncertainty because it is too expensive to obtain reliable knowledge. Fakes may make it more expensive to obtain reliable information, which would raise uncertainty. We concentrate on this by examining whether false documents cause ambiguity regarding the information they contain.to question whether the content is real or fake. Ambivalence and uncertainty are conceptually different. When people are given a decision on which they disagree, they become ambivalent, and "new information merely heightens the inner struggle" [5] [16]. The introduction of fresh information can eliminate ambiguity, which is experienced when there is not enough information available to decide [5]. According to Downs (1957), citizens experience uncertainty because it is too expensive to obtain reliable knowledge. Deepfakes may make it more expensive to obtain reliable information, which would raise uncertainty. We concentrate on this by examining whether false deepfakes cause ambiguity regarding the information they contain.

A loss in trust in news on social media, where deepfakes are most likely to spread, might be one of the biggest repercussions if deepfakes and other misinformation efforts are effective in spreading doubt. The psychological implications of deepfakes, namely the perception of political news on social media, are the subject of our second field of study. The public is losing trust in the media [8], and social media news is currently seen with less trust than that derived from traditional sources [17].

Therefore, the worry that deepfakes may, over time, like other sources of misleading information [3] encourage individuals to believe that a fundamental basis of reality cannot be established is what motivates our work. According to research, one factor contributing to the spread of false political stories online is a "thirst for chaos," or the urge to "watch the world burn" without thinking about the

repercussions [15]. State-sponsored propaganda has made creating doubt about what is real and what isn't a top priority.[18] “The objective is... to wreck the information space so the spectator loses up trying for any truth amid the chaos." The aggregation of countless contradictory, ridiculous, and muddled remarks that bad actors introduce into online discourse may lead to a systemic situation of uncertainty [19]. It is vital to pay close attention to whether deepfakes lead to doubt and erode trust in this case.

IX. Some Benefits of deepfake

The deepfake technology have been used in different industry sectors to producing the products such as cinema, educational material, internet technology, games, entertainment, media platforms, health, science, and a wide range of commercial sectors including e-commerce and fashion. Deepfake technology has several advantages for the movie business. It may be used, for instance, to update old film material rather than reshoot it or to offer performers who lost their voices due to illness alternative voices. Moviemakers will be able to re-enact old movie sequences, develop new movies with deceased actors and actresses, employ state-of-the-art facial edits and special effects in post-production, and fix amateur recordings. [20].

X. Deepfake detection

Deepfakes occasionally don't need to reach a large audience to have an adverse impact. Deepfakes may be distributed to target audiences as part of a sabotage plan without utilizing social media by those who make them with malevolent intent. The research community has concentrated on creating deepfake detection algorithms and several successes have been announced because of catching the frightening deepfake problem. This essay has evaluated the most recent techniques and provided an overview of common strategies. A conflict is developing between people who utilize sophisticated machine learning to produce deepfakes and others who try to identify them. Since the quality of deep fakes has been rising, the effectiveness of detection techniques needs to be enhanced. Its premise is that AI is capable of fixing whatever that it has damaged. Several methods have been reported and evaluated in the early phases of detection, albeit using fragmented datasets. One method for improving the effectiveness of detection systems is the development of a deepfake benchmark dataset that is continuously updated and increased. Because deep learning-based models require a large training set, this will make it simpler to train detection algorithms. Performance of the deep fake detection approach must be enhanced, particularly in cross-forgery and cross dataset. scenarios. Since most detection algorithms are developed and evaluated in experiments using the same forgeries and dataset, their generalizability is not assured. Deepfakes, as they are now termed, can be produced in large which requires more research and efforts to benefit from data science mechanisms and web scraping to detect harming deep fake [21][22]

XI. Conclusion

. Since perceiving deep fakes does not imply believing in them, they have begun to erode people's trust in media content. They could irritate the targeted people, promote hate speech and misinformation, heighten political turmoil, stir up public disturbance, lead to carnage, or even start a war. This is crucial right now since deepfake production tools are getting easier to use and social media platforms can quickly spread false information.

Acknowledgment

We thank Applied Science University for useful facilities and its supporting to complete this research.

References

[1] Westerlund, M. “The Emergence of Deepfake Technology: A Review | TIM Review”. (2019).

[2] Pantserev, K. A. “The Malicious Use of AI-Based Deepfake Technology as the New Threat to Psychological Security and Political Stability”. Advanced Sciences and Technologies for Security Applications, 37–55. (2020).

[3] Lyu, Siwei. "Deepfake detection: Current challenges and next steps." In 2020 IEEE international conference on multimedia & expo workshops (ICMEW), pp. 1-6. IEEE, 2020.

[4] Jones, Valencia A. "Artificial intelligence enabled deepfake technology: The emergence of a new threat." PhD diss., Utica College, 2020.

[5] Karnouskos, Stamatis. "Artificial intelligence in digital media: The era of deepfakes." IEEE Transactions on Technology and Society 1, no. 3 (2020): 138-147.

[6] Yadlin-Segal, A., & Oppenheim, Y. “Whose dystopia is it anyway? Deepfakes and social media regulation”. Convergence: The International Journal of Research Into New Media Technologies, 27(1), 36–51, (2020).

[7] Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? Business Horizons, 63(2), 135–146.

[8] Vaccari, C., & Chadwick, A. “Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. Social Media + Society, 6(1), 205630512090340. (2020b).

[9] Karnouskos, S. (2020). Artificial Intelligence in Digital Media: The Era of Deepfakes. *IEEE Transactions on Technology and Society*, *1*(3), 138–147.

[10] Shorman, S, M. Jarrah, and A. Alsayed. "The Websites Technology for Arabic Language Learning Through COVID-19 Pandemic." In Future of Organizations and Work After the 4th Industrial Revolution, pp. 327-340. Springer, Cham, 2022.

[11] Khder, M. A., Bahar, A. Y., & Fujo, S. W. (2022, June). Effect of Artificial intelligence in the field of games on humanity. In 2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS) (pp. 199-204). IEEE.

[12] Yazdinejad, A., Parizi, R. M., Srivastava, G., & Dehghantanha, A. (2020). Making Sense of Blockchain for AI Deepfakes Technology. *2020 IEEE Globecom Workshops (GC Wkshps*.

[13] Wagner, T. L., & Blewer, A, “The Word Real Is No Longer Real”: Deepfakes, Gender, and the Challenges of AI-Altered Video. *Open Information Science*, *3*(1), 32–46. (2019).

[14] Kaur, S., Kumar, P., & Kumaraguru, P. (2020). Deepfakes: temporal sequential analysis to detect face-swapped video clips using convolutional long short-term memory. Journal of Electronic Imaging, 29(03), 1.

[15] Lyu, S. “Deepfake Detection: Current Challenges and Next Steps. *2020 IEEE International Conference on Multimedia &Amp; Expo Workshops (ICMEW)*. (2020).

[16] Shorman, S, and S. Pitchay. "Removing Rain Streaks From Single Images Using Total Variation." The International Journal of Multimedia & Its Applications (IJMA) Vol 10 (2018).

[17] Zhao, B., Zhang, S., Xu, C., Sun, Y., & Deng, C. “Deep fake geography? When geospatial data encounter Artificial Intelligence. *Cartography and Geographic Information Science*, *48*(4), 338–352. (2021).

[18] Karnouskos, S. (2020). Artificial Intelligence in Digital Media: The Era of Deepfakes. *IEEE Transactions on Technology and Society*, *1*(3), 138–147.

[19] Taylor, B. C, “Defending the state from digital Deceit: the reflexive securitization of deepfake. Critical Studies in Media Communication, 38(1), 1–17. (2020).

[20] Gosse, C., & Burkell, J. “Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communication*, *37*(5), 497–511. (2020).

[21] Khder, M.A., *Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application.* International Journal of Advances in Soft Computing & Its Applications, 2021. **13**(3).

[22] Khder, M.A., S.W. Fujo, and M.A. Sayfi, A roadmap to data science: background, future, and trends. International Journal of Intelligent Information and Database Systems, 2021. 14(3):p. 277-293.