LUNDS
UNIVERSITET

Deepfake

An Emerging New Media Object in the Age of Online Content

A Master's Thesis for the Degree of Master of Arts (Two Years) in Visual Culture

Emily Hsiang

Division of Art History and Visual Studies
Department of Arts and Cultural Sciences
Lund University
KOVN12, Master's Thesis, 30 credits
Supervisor: Lila Lee-Morrison
Spring 2020

Acknowledgements

First and foremost, I would like to thank my supervisor Lila Lee-Morrison for her encouragements, insights, and for inspiring me to take charge of my thesis.

Thank you to artist Bill Posters for the screener to *Big Dada*. Thank you, directors Francesca Panetta and Halsey Burgund, for the press kit, photos, and screener for *In Event of Moon Disaster*.

I would like to thank my friends and the Visual Culture class of 2020. Thank you to my Visual Culture hearts: Christina, Karen, and Mimosa. I especially want to thank my thesis buddies, the girls from Media and Communications: Cheryl, Juliet, and Mirjam, for our shared laughter and lunches.

Lastly, I want to say a huge thank you to my loving parents, Hua-mo Shen and Jieh Hsiang, without whom I would not be here today. Lots of love to my loveliest grandma, who cheered for me across the Atlantic Ocean. Thank you for your unconditional love and faith.

Abstract

Thesis: Deepfake: An Emerging New Media Object in the Age of Online Content
Author: Emily Hsiang

This thesis aims to understand the phenomenon of deepfake as new media objects and how deepfake content circulate and are consumed online. Deepfake is a newly emerging technology that face swaps targeted individual faces in existing media using machine learning. To document and recognize deepfake in the field of media studies, this thesis attempts to translate the technicality deepfake technology through the perspective of media studies.

I apply new media theory from Lev Manovich and medium theory from McLuhan and examine deepfake properties as new media object and its symbiosis relationship with social media. Deepfake relies heavily on its symbiotic circulatory relationship with the Internet. With major players – the user, the technology, the Internet – influencing the other by playing a part at being mediums and contents. Through examining artworks made with deepfake technology, it is revealed that deepfake could be used to deliver powerful message by being viewed as both content and medium.

Keywords

deepfake, new media, medium, machine learning, face swap, fake news

Table of Contents

List of Figures

Introduction

At present, two decades into the 21st century, the world has swiftly adapted to the digital. Online networks are now an inseparable virtual part of physical life. The introduction of social media and personal smartphones – highly functional personalized mini-computers with the ability to continuously be connected online – has made it even more possible to have real-time exchanges via text, voice calls, and video images across the world. We as users of the Internet have become not merely consumers of mass media but are ourselves producers of our own contents, which in turn are consumed by others. This is the age of online contents. With the exception of certain areas where internet speed curbs the rate of information delivery, the experience of space and time has been dramatically changed through global connectivity.[1] Among this environment of increasingly complex and expanding digital realm of visual new media, deepfake emerges as a new visual phenomenon that is increasing in sophistication and popularity.

Deepfake is, like many digital objects circulating in and around the Internet, a new media object that is uniquely at the crossroads among emerging visual contents and technologies. Deepfake is widely understood as a deceptive video where a subject's face is swapped in to replace another body's face, making it so that the subject is shown to act upon or say something that did not occur.[2] Deepfake videos is a technological product that is made through a specific type of machine learning technology. Currently, deepfakes are predominantly pornographic videos and memes. Approximately 96% of deepfakes online are of pornographic nature, many of which are faces of female celebrities slapped onto the bodies of adult porn stars without their consent.[3] Meanwhile, many deepfake memes[4] feature faces of celebrities in absurd situations and tend to be humorous, such as juxtaposing celebrity faces onto another celebrity of a different gender or posting adult's faces onto toddlers.[5]

Deepfake is non-traditional in a sense that the technology can create highly realistic face swaps using machine learning. With the utilization of machine learning, deepfake technology

---

[1] N. Mirzoeff, *An Introduction to Visual Culture*, 2nd edn, USA and Canada, Routledge, 2009, p. 221.
[2] M. Maras and A. Alexandrou, 'Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos', *The International Journal of Evidence & Proof*, 2019, Vol. 23(3) p. 255–262, Available from: SAGE journals, (accessed 5 February 2020).
[3] R, Metz and CNN Business, 'The number of deepfake videos online is spiking. Most are porn', *CNN*, 7 October 2019, https://edition.cnn.com/2019/10/07/tech/deepfake-videos-increase/index.html, (accessed 10 May 2020).
[4] A meme is typically an image, video, or text that can easily be copied and spread. Memes typically incite a certain emotional response so it can be shared among Internet users.
[5] *Baby Elon Musk Montage Deepfake*, [online video], 8 May 2019, https://youtu.be/WHwQeetjLwk, (accessed 10 May 2020).

learns to mimic an individual's facial expressions, gestures, speech patterns, and creates realistic looking face swapped virtual avatars with radically relatively hands-free human interaction. Deepfake is not the first technology to create digitally animated people or altered faces and voices, but it is the first to be automated with machine learning. Deepfake automates and industrializes the form of art of digital manipulation, abating the time-consuming digital manipulation that Photoshop and CGI require to create realistic-looking virtual avatars.

Research question

This thesis aims to study deepfake as a technology contributing to a specific type of visual phenomena in the present. Like many new media emerging from the age of online contents, deepfake is a new type of content that is in a mutual symbiotic relationship with online interaction between individual users. But unlike many other new media, deepfake is part of an unprecedented generation of media that can be fully automated with machines. There are two questions this thesis seeks to understand: How can we understand deepfake as new media objects and what can come out of it? How do these types of content circulate and how are they consumed?

Background and Relevance

Deepfake has arrived at a time when social life is tightly connected, entwined, and inseparable from online social media. Social media has been hailed as a leading baton for individuals to find other likeminded communities. With the proliferation of portable devices and applications, news and visual content can rapidly spread instantaneously from devices to devices at the speed of light. Different social media users – family members, friends, acquaintances, strangers – can all share, admire, and experience great quantities of online content anytime of the day. Many traditional media functions are moving online, such as television broadcasting transitioning to streaming or reading the newspaper turning to reading social media feeds. Interpersonal connections have also moved online. More and more people are consuming information and daily news off social media platforms like Facebook or Reddit. Social media users tend to curate their

experience so they mostly encounter news perspectives that they already agree with (which were then highlighted by the platform's algorithms to push content a user may like so that they may stay on the platform longer), or receive reposts from friends and family, which in turn makes people less susceptible to checking if the information they receive is true or false.[6] In this environment, facts are becoming harder to separate from fiction.

Together with social media platforms' unique ability to spread viral content, deepfake has received wide attention for their potential use in spreading disinformation such as fake news – a form of news content that delivers deliberate misinformation. Indeed, many news media publications reflects this anxiety by reporting mainly speculations about the future of deepfake in the aspect of potential abuse deepfake could spawn: such as face swapping an individual's face on porn to use as revenge porn or blackmail,[7] swaying political elections by creating deepfake videos of politicians doing things they did not do,[8] or deepfake posing as a threat to journalistic trust.[9]

News media's panic surrounding deepfake videos' potential ability to abuse political outcomes is given credence by the fact that, as I write this thesis, the first known usage of deepfake video in politics was reported. On 7th February 2019, just before legislative election day in India, a local politician had manufactured and intentionally spread a viral deepfake video of himself speaking in a minority language as an attempt to sway voters.[10] At the time of writing this thesis, it is unknown if this deepfake video changed the outcome of the election, but the implications are that deepfake can be used to influence political outcomes. I have also found a case where a video of the President of Gabon addressing his nation was claimed to be deepfake – causing immediate social unrest and a military coup.[11] Both videos went viral on social media –

---

[6] R. Chesney and D, Citron, 'Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics ', *Foreign Affairs*, January-February 2019, v. 98, iss. 1, pp. 147-55, (accessed 15 February 2020).
[7] M. Westerlund, 'The Emergence of Deepfake Technology: A Review', *Technology Innovation Management Review, Vol 9, Iss 11*, 2019, p. 40-53. Available from: Directory of Open Access Journals, (accessed 23 February 2020).
[8] HB. Dixon Jr., 'Deepfakes: More Frightening Than Photoshop on Steroids', *Judges' Journal, vol. 58*, no. 3, 2019, p. 35–37.
[9] F. Marconi, 'Enablers: The AI Technologies Driving Journalistic Change', p. 55-128, *Newsmakers: Artificial Intelligence and the Future of Journalism*, Columbia University Press, New York, 2020, JSTOR: www.jstor.org/stable/10.7312/marc19136.7, (accessed 12 May 2020).
[10] N. Christopher, 'We've Just Seen the First Use of Deepfakes in an Indian Election Campaign', *Vice*, 18 February 2020, https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp, (accessed 19 February 2020).
[11] A. Breland, 'The Bizarre and Terrifying Case of the "Deepfake" Video that Helped Bring an African Nation to the Brink', *Mother Jones*, 15 March 2019, https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/, (accessed 27 April 2020).

the deepfake from India spread within WhatsApp as user-to-user transmission of media, the Gabon video was posted on Facebook where rumors quickly gained a footing on Twitter.

Yet despite the urgency and anxiety associated with deepfake, there exists a gap between creating deepfake videos and the experience of the visual consumption of deepfake. The process of creating deepfake is transparent and is generally knowledge open for all. Knowledge to learn how to create deepfake are freely available online. Deepfake programs with user friendly interfaces are also available online and can be easily found. There are forums and webpages of Frequently Asked Questions (FAQ) for anonymous users to learn from each other, offer help to other amateurs, and share their creations and creative processes. This development is part of a recent trend due to the democratization of technologies, democratization of technical knowledge, and increasingly available hardware. What used to be technologies (with knowledge and the equipment required to use them) only available to highly trained technicians, are now available to almost any individual with a computer can now access, thanks to the proliferation of the Internet.

Empirical Material and Delimitations

A part of my interest in deepfake videos also comes from curiosity about how the deepfakes are made. As such, understanding the technical specificity of how deepfakes are created is vital to understanding the meanings created at the site of production. I chose to focus on the technical specificity of creating deepfakes with the desktop computer program FaceSwap because it has several open online forums with detailed steps to learn from.

Deepfakes created by amateur hobbyists are what computer vision academics refer to as amateur deepfake, and I will refer to these deepfakes as such in this thesis. However, there is another aspect of deepfake that is shrouded in mystery, and that is what I called high-tech deepfake: a kind of deepfake that is made by corporate or startup companies, where the process of deepfake creation tend to be company secrets. In addition to these handful of high-tech deepfake created by companies, there are also deepfakes created for academic research that I would consider as not amateur deepfake. This thesis will focus on amateur and high-tech deepfake created by companies.

To delimitate my empirical materials, I have chosen four specific deepfake videos to ground my analysis. To demonstrate amateur deepfake, I use a deepfake meme that has become a

common example referenced in academic articles to explain the technology of deepfake; as well as a porn video that is among one of the first uploaded deepfakes. For the high-tech deepfake, two artworks are chosen: *Imagine This…* was a deepfake video released online via Instagram, as a challenge to Facebook's algorithm.[12] *In Event of Moon Disaster* was an installation in the form of a 1960s television.[13]

It is vital to understand what kind of environment enabled the emergence of deepfake. Deepfakes are primarily distributed through social media platforms such as Reddit, YouTube, or other specialty websites. I observe these websites through phenomenological approach as part of my empirical material, as the websites are connected to the circulation and consumption of deepfakes.

Like a great many online visual media contents, such as photos or videos, there exist plenty of deepfake content that are either original uploads or copies of the original that are uploaded somewhere or sometimes later. Original copies of deepfakes tend to be inhumed by many other copied and reposted copies of the same videos. Sometimes they may be deleted. This nebulous existence of media on the Internet made it somewhat tricky to attain the original copy of deepfakes in question. Although deepfakes are posted on numerous online websites, I have narrowed down and selected a handful of websites to base my observations on. The websites I take knowledge from are GitHub, Reddit, YouTube, and MrDeepFakes. All these websites have different priorities, attract different content and different users. GitHub is an online hosting website, where the free accounts are commonly used to host open source projects for developers or tech enthusiasts. The deepfake program FaceSwap that create deepfake videos is hosted on GitHub.[14] It is from this program and its associated forums I am able to learn how deepfake as a technology works. Reddit hails itself as 'the front page of the internet', home to thousands of communities and endless conversations. It is the first platform where I encountered my first deepfake and which continues to be a site where active members upload and share deepfake meme videos. YouTube is an online video-sharing website and also another popular site for users to share deepfake meme videos. MrDeepFakes is a deepfake porn site where users take requests

[12] C. Metz, 'A Fake Zuckerberg Video Challenges Facebook's Rules', *The New York Times*, 11 June 2019, https://www.nytimes.com/2019/06/11/technology/fake-zuckerberg-video-facebook.html, (accessed 27 February 2020).
[13] MIT Center for Advanced Virtuality, *IEOMD-press-release-2019.pdf*, Retrieved: with the filmmaker's permission, Cambridge, Massachusetts, 22 November 2019, (accessed 11 April 2020).
In text: (MIT Center for Advanced Virtuality, 2019).
[14] *GitHub,* [website], https://github.com/deepfakes/faceswap, (accessed 3 March 2020).

to create specific deepfake porn, upload deepfake porn for streaming, and download. It is not the only existing deepfake porn website out there, but this site is where I select my empirical material.

Method

For my interpretation of the deepfake examples, I lean heavily on Gillian Rose's critical visual methodology. Rose describes the critical approach to visual methodology as an analytical meditation on the social conditions, effects, and circulation of a visual image. The observer's own observations towards the examined images must also be considered.[15] I apply this method by treating my deepfake examples as visual images to be dissected using Rose's four sites of critical visual methodology. I examine how a deepfake's site of technology (how deepfake is produced) and site of circulation (how deepfake travels from its production site to online platforms), together with the site of image (what the content of this specific deepfake empirical material is) influence its perception at the site of audiencing (where the deepfake's meaning is made by its viewers).[16]

The quality of deepfake videos can vary differently depending on its source material, the kind of machine learning programs, and how good the hardware is used to create them. Depending on the quality of deepfake, the videos can be visually realistic or uncanny. These different qualities of the image produced provide practical steps to both detect and deconstruct deepfake as sites of image, where the image's meanings are made in the image itself.[17]

Part of my curiosity towards deepfake technology is to understand how deepfake can be recognized and documented as part of a new type of medium that is emerging within the realm of visual media. Therefore, part of my methods is focused on studying the websites where deepfake first emerged. This includes social media websites Reddit, GitHub, YouTube, and a porn website MrDeepFakes.

---

[15] G. Rose, *Visual Methodologies: An Introduction to Researching with Visual Materials*, London, Sage, 2001, Rose, *Visual Methodologies*, p. 24-47.
[16] Rose, *Visual Methodologies*, p. 22.
[17] Rose, p. 32.

The social modality of audiencing images examines 'the social identities of those doing the watching.'[18] While there are ways to quantify what kind of audience that are using certain platforms, it is not a topic this thesis covers. Instead, I provide my own perspective and perception of this technological development. So instead of asking *who is doing the looking,* I am more interested in asking 'for whom this video is created to do the looking?'

My position here is both observer and participant. As an interested layman, my taking part of these websites come from my own experience in participating in these websites as users, as well as objective observations within these sites. I consume content on Reddit and YouTube on a daily basis. There are different ways of coming across these types of deepfake content – either by deliberately searching for it or getting it recommended from a social media's algorithm – on the Internet. Different websites create worlds for different distribution, interaction and consumption of deepfakes. My own observation on Reddit is that new posts with deepfake content, being sufficiently popular, would occasionally pop into my news feed without me actively going looking for it. This is not to say I have not actively sought out deepfakes on Reddit, which may have influenced the Website's algorithmic decisions. But this points out that deepfake is a visual phenomenon with rising popularity. On GitHub, the discussion among deepfake tends to be users asking each other engineering related questions. Therefore, I use GitHub as a site of knowledge to learn about the technicality of deepfake.

Theory

I analyze deepfake from a media studies perspective. I apply Marshall McLuhan's theory on how *communication mediums that deliver content shapes behavior*, with the help of Paul Levinson's assertion that the users and the Internet are new forms of content and medium and take it one step further by applying their claims and look at online social media platforms as empirical material. Internet, social media and users are the distribution, content, and medium. Applying this concept to deepfake, I hope to find deepfake as mediums and define them as sites of meaning.

In this thesis, I borrow heavily from Lev Manovich's method in identifying new media to examine the deepfake as a new media object. Manovich identifies new media as the

[18] Rose, p. 40.

synthesization of media merging with the computer, where new media – graphics, moving images, texts, have become computable and become another set of codes.[19] Manovich draws theoretical tools from both the humanities, including art history, social studies, and social theory - and computer science[20], and grounds his analysis of the language of new media by 'placing it within the history of modern visual and media cultures.'[21] As such, I apply this theory and method of approaching my analysis of deepfake as a new media object. I trace, document, and describe deepfake's short history, emergence, and technicality.

To further ground my analysis of the selected empirical material, I draw concepts from philosopher Walter Benjamin's concept of aura from his work 'The Work of Art in the Age of Its Technical Reproducibility'; philosopher Judith Butler's concept of gender performativity in 'Imitation and Gender Insubordination'; as well as social psychologist Shoshana Zuboff's concept of surveillance capitalism in 'The Age of Surveillance Capitalism.'

Previous Research

Many writings about deepfake focus on the technical present, or the speculations of the future. There are many online academic journals in the fields of computer vision, open information science, crime and law, as well as news journal publications on politics and technology reviews that discuss deepfake in terms of its potential impact on society. The technological evidence research was found in scientific reports, academic journals, as well as open access to technological online forums such as GitHub. Many of these academic journals detail the technological process of creating deepfake videos,[22] the ways to identify deepfake videos with the naked eye,[23] as well as with computer vision,[24] or developing new technology to accurately

[19] Manovich, p. 20.
[20] Manovich, p. 10.
[21] Manovich, p. 8.
[22] A. Hauser, and M. Ruef, 'Deepfake - An Introduction', *Scip Labs*, 2018, Available from: Academia.edu, (accessed 3 March 2020).
[23] X. Yang, Y. Li, S. Lyu, "Exposing Deep Fakes Using Inconsistent Head Poses", ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, Calgary, Alberta, Available from: ArXiv.org, (accessed 2 March 2020).
[24] Y. Li and S. Lyu, 'Exposing deepfake videos by detecting face warping artifacts', 2018, Available from: arXiv:1806.02877, (accessed 2 March 2020).

detect[25] or identify[26] deepfake. There are increasing scholarships from the perspective of the social science and gender critique, as I have found journals that have touched upon how the community of deepfake video creators are seemingly uninterested in examining deepfake's role of perpetuating a history of male gaze[27] or toxic geek community culture.[28]

Without discerning the potentiality and technical development of deepfake technology, there is a lack of attention to deepfake analysis from a visual cultural and media studies perspective, as well as a historical account of how the technology rose to be. This may partly be due to deepfake being a newly emerging technology, having the term coined only in late 2017, and many of the publications are only recently published in 2018-19.

In terms of public discourse, deepfake has garnered widespread attention with its potential to disrupt both the scale of micro and macro. Many crime and law academic journals place their attention on addressing the potential harm of deepfake, and how laws could be drawn up to prevent future malicious use of the technology.[29] News articles contain many speculations about the future of deepfake mainly in the topics of potential abuse deepfake could spawn, such as face swapping a private individual's face on porn to use as revenge porn or blackmail, swaying political elections by creating deepfake videos of politicians doing things they themselves did not do.[30]

News articles, however, are a vital process in helping me piece together the short history of deepfake. I was particularly dependent on Vice Media Group Motherboard's Sex & Tech reporter Samantha Cole's reports and interview with u/deepfakes in finding more information about the emergence of deepfake and deepfake porn. As many of these original posts of deepfake videos are now deleted online or nearly untraceable during my own research.

---

[25] Y. Li, et al., 'Celeb-DF: A New Dataset for DeepFake Forensics', 2019, Available from: arXiv:1909.12962, (accessed 2 March 2020).
[26] Y. Li, et al., 'Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics', 2019, Available from: arXiv:1909.12962, (accessed 2 March 2020).
[27] T. L. Wagner and A. Blewer, '"The Word Real Is No Longer Real": Deepfakes, Gender, and the Challenges of AI-Altered Video', *Open Information Science, Volume 3: Issue 1*, 19 July 2019, p. 32-46. Available from: doi.org, (accessed 6 February 2020).
[28] R. Winter and A. Salter, 'DeepFakes: uncovering hardcore open source on GitHub', *Porn Studies*, p. 1-17, 2019, DOI: 10.1080/23268743.2019.1642794, (accessed 26 February 2020).
[29] R. Chesney and D. Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security', *California Law Review, Inc.*, 2019, Available from DOI: https://doi.org/10.15779/Z38RV0D15J, (accessed 14 May 2020).
[30] E. Thomas, 'In the battle against deepfakes, AI is being pitted against AI', *Wired*, 25 November 2019, https://www.wired.co.uk/article/deepfakes-ai (accessed 20 May 2020).

Dissemination

In Chapter 1, I hold a general discussion about deepfake. I historicize the visual phenomenon of face swap, the technological changes that lead up to the normalization of digital manipulation, and the short online history of deepfake technology. Then, I discuss the technical specificity of deepfake, breaking down its technical process of face swapping by placing my gaze on open-source program[31] FaceSwap.[32] I discuss deepfake as new media and introduce social media platforms through which deepfake circulates. I then go on to examine its dynamics and functions on social media driven platforms Reddit and YouTube.

In Chapter 2, I analyze amateur deepfake empirical material *Man of Steel deepfake gif* and *Gadot-porn video*, two of the earliest deepfakes. I discuss the sites of circulation, production, and audiencing for amateur deepfake; specifically, how social media platforms encourage and drive the creation of deepfake. In Chapter 3, I analyze high-tech deepfake by looking at artworks *Imagine This…*[33] by British artists Bill Posters and Daniel Howe, and *In Event of Moon Disaster* directed by Francesca Panetta, Halsey Burgund, and created by Massachusetts Institute of Technology Center for Advanced Virtuality.[34] I explore the general discourse surrounding these works of art and discuss deepfake as content and as a medium.

[31] To clarify, the phrase 'open source' pertains to products where the source code, design, or product is transparent to the public for use. The phrase 'open-source' refers to a product or content that is created with decentralized software development model, but the product may not be open source itself.
[32] *Faceswap.dev*, [website], 2019, https://forum.faceswap.dev/, (accessed 3 March 2020).
[33] B. Posters, *Bill Posters,* [website] http://billposters.ch/projects/big-dada/, (accessed 27 March 2020).
[34] *In Event of Moon Disaster*, [website], 2019, https://moondisaster.org/, (accessed 28 February 2020).

Chapter 1: General Discussion

What is Deepfake

A deepfake is a type of synthetic media in which a person's facial features in an existing video is replaced or swapped by someone else's facial features. A synthetic media is an umbrella term for a type of media that is algorithmically created or altered with machine learning.[35] The term deepfake is a portmanteau of 'deep learning' and 'fake', and, as the name suggests, is a fake media created by deep learning. Deep learning is a kind of algorithm called "neural networks" that learns to replicate patterns by going through large data sets. With computing prowess created with machine learning, deepfake can create convincing fake videos depicting individuals doing things they did not do with strong potential to deceive the video's audience.

For this thesis, I use deepfake's generally accepted definition in various other computer vision research journals as a deceptive video, created with machine learning that shows a face swapped subject doing or saying things that did not occur. Deepfake isn't the first technology to swap or replace faces with realism, but it is the first technology to automate this process. The automation allows re-animation to be done with minimal human intervention and with a collection of images from a person rather than the person being performers of these actions. The faces and expressions can be so expertly generated, that it becomes difficult to distinguish what is a true footage or false footage.

The History of Swapping Faces

The act of creating manipulated images or videos is not new, nor is the swapping of faces in images. The practice of switching out faces has existed historically, often occurring from the need for a certain type of image. One well known historical face swap (Fig. 1) is the face swapping of U.S. president Abraham Lincoln's portrait head and rival U.S. senator John C. Calhoun's body, originally made by engraver A.H. Ritchie, in an 1852 print by portrait painter

[35] A. Ovadya and J. Whittlestone, 'Reducing malicious use of synthetic media research: Considerations and potential release practices for machine learning', 2019, Available from: arXiv:1907.11274, (accessed 12 May 2020).

Thomas Hicks. This face swapping is created by superimposing the head of Lincoln's portrait onto Calhoun's body on a woodprint. It was commonly believed that the print came from during a period of high demand for a heroic, presidential looking portrait following Lincoln's assassination. Lincoln himself was known at the time for being notoriously ugly with a crooked stature, which was deemed, not heroic looking enough.[36] As such, the printmaker cut off a portrait of Lincoln, flipped his face, and slapped the face onto Calhoun's body. Despite the weird proportion of the face and Lincoln's infamous mole appearing on the wrong side of his cheek, this image is widely cited as one of the most important presidential images of Lincoln. The implication being that, because there is a desire for a heroic looking president, falsehood can be pardoned as long as the desire is satisfied.[37]

Fig. 1. A composite print of a face swapped presidential image of Abraham Lincoln.[38]

From analog photo manipulation arises digital manipulation. Computer-generated imagery (CGI) is commonly used for computer-generated effects in films, television, video

[36] H. Holzer, 'How the Printmakers Saw Lincoln: Not-So-Honest Portraits of "Honest Abe"', *Winterthur Portfolio* Vol. 14, No. 2 (Summer, 1979), p. 143-170, JSTOR: www.jstor.org/stable/1180612, (accessed 25 March 2020).
[37] K. Lum, 'From Analog to Digital: A Consideration of Photographic Truth, 2012.', *Everything is Relevant: Writings on Art and Life*, 1991-2018, Montreal: Concordia University Press, p. 214, (accessed 4 March 2020).
[38] Pate, W., *Abraham Lincoln*, 1865, [photograph], Retrieved from the Library of Congress, https://www.loc.gov/item/2003654314/, (accessed 23 March 2020).

games, or other visual media. While these new technologies may have seemed out of this world when they were innovations, they have now been accepted as part of the norm. Realistic digital look-alikes, explosions, dinosaurs, aliens, actors digitally de-aged, or digitally resurrected actors are all commonly used within entertainment. Although the technology uses complicated computer software, it requires a team of highly skilled technicians and artists to manually digitally manipulate the face of the subject in question. A recent controversial case of CGI face swapping in film is the case of the digital resurrection of Peter Cushing, who died in 1994, in *Rogue One: A Star Wars Story* (2016). Where, instead of recasting an actor for the role of Grand Moff Tarkin, the filmmakers opted to recreate and reanimate Cushing as a CGI character.

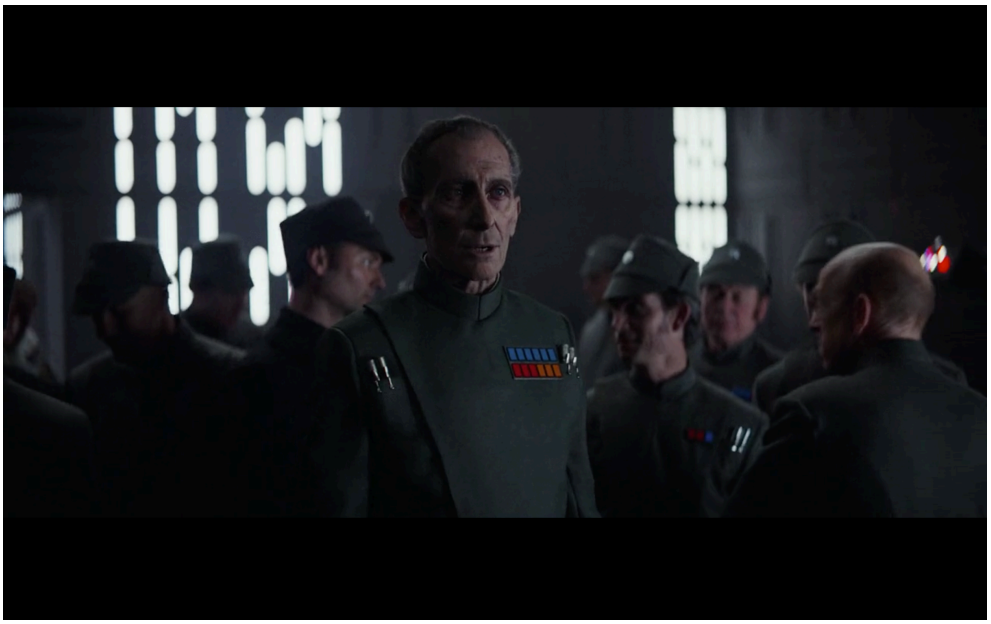

Fig. 2. A still screenshot of digitally resurrected Peter Cushing as Grand Moff Tarkin standing against a backdrop of live actors.[39]

This digitally enabled posthumous performance is played by British actor Guy Henry, hired for his physical and facial resemblance of Peter Cushing. Just like Abraham Lincoln's face swap was created due to the need for a heroic looking presidential image, Cushing's digital resurrection stems from a need to see a Cushing's persona presented as a digitally resurrected avatar. Cushing's face is rendered, and his facial movement is so smooth it seemed unnatural (Fig. 2). As a still image, his facial features stand in contrast to the other real, live actors behind him, whose facial complexion is less smooth, grainier looking. The face features an inhumanly

[39] *Rogue One: A Star Wars Story*, Gareth Edwards, USA, Walt Disney Studios Motion Pictures, 2016, [still frame of digital film], (accessed 25 February 2020).

shiny sheen, what we could call 'handmade flaws', or an uncanniness of a face that is strangely familiar. Henry provided the voice and physical performance, which was then hidden underneath a digitally woven mask of Cushing's face diligently created by the VFX team. Cushing is not a simulation, but an approximation of a simulation – a dead character portrayed by a living actor imitating the dead actor. Commodified in death, the chimera moves with the grace of the actor enveloped in a digitally coated skin of a dead man.

CGI is a different technology from deepfake, but it is part of a historic trend of digital manipulation that has led to more general acceptance for digital manipulation. Although CGI uses computations and computer technology, it takes a team of technicians to labor in the details. The technicians have to meticulously carve out facial and vocal manipulations frame by frame. This type of CGI digital manipulation also requires state of the art hardware, time for rendering, which makes it less accessible for the regular engineers who may not have the same training or equipment. Deepfake is set apart from traditional digital manipulation such as Photoshop and CGI because it requires less human attention in that much of the reproduction is automated through machine learning.

History of Deepfake

The term deepfake was originally coined by an anonymous Reddit user who referred to himself as u/deepfakes.[40] In 2017, u/deepfakes released several pornographic deepfake videos featuring female celebrity faces imposed on porn stars in a subreddit called r/deepfakes.[41] In an interview with Vice Media Group Motherboard's journalist Samantha Cole, u/deepfakes revealed that the codes he used are based on multiple open-source libraries like Keras with TensorFlow.[42] To compile enough facial image material for the deepfake videos, u/deepfakes scraped images from

[40] For clarity, I've added the Reddit user handle "u/" in front of Reddit user "deepfakes" to differentiate u/deepfakes the user from deepfake the technology. We know that the user uses masculine pronouns from when he was interviewed in Samantha Cole's Vice Motherboard article. S. Cole, 'AI-Assisted Fake Porn Is Here and We're All Fucked', *Motherboard*, 11 December 2017, https://motherboard.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn, (accessed 3 March 2020).

[41] In Reddit lingo, "u/" in front of a subject indicates a user. "r/" in front of a subject indicates a subreddit, which is a community board with specific hobbies or interests.

[42] Keras is an open-source neural-network library and TensorFlow is an end-to-end open-source machine learning platform. Together, both programs can achieve fast experimentation with deep neural networks. Both programs are open sourced, meaning any user can use them and improve it.

databases such as Google, YouTube screenshots and stock photo sites. u/deepfakes then shared the codes he used to create deepfake videos on Reddit, and a plethora of copycat porn videos quickly emerged. The contents of these videos soon attracted the attention of various media outlets, creating an overflow of media attention to deepfake porn. Soon the subreddit r/deepfake and other deepfake related subreddits were banned.[43] Many major social media platforms soon follow suit and banned deepfake porn on their sites.[44]

u/deepfakes has since deleted his Reddit account, but the codes remain accessible by any amateur engineer. Quickly after u/deepfakes released his codes, programs to create deepfakes like FaceSwap[45] – a desktop based program – and FakeApp[46] – a mobile based program – emerged and developed into more user friendly versions with user-interface, making it possible for less tech savvy users to create their own deepfakes. Both programs (and other copycat programs) are publicly available on open source and crowd sourced website GitHub, which includes directions on how to use them.

Why Study Technical Specificity

The face swapping produced by deepfake is interesting to me because it produces a specific type of visual phenomenon. By understanding the technical specificity to how a deepfake is created, it will also help to understand how deepfake travels, circulates, and displays. There are two different mode where deepfake is produced: amateur deepfake and high-tech deepfake. Amateur deepfake are what academic computer journals referred as deepfake created by a hobbyist that uses open-source software and are the most common kind of deepfake today. High-tech deepfake are much less common and created by corporate or startup companies. They are currently used primarily as art installations. The process for creating high-tech deepfake are relatively shrouded in mystery, as so far they remain company secrets. Therefore, I will describe the creation process

[43] *Reddit*, [website], 2018, https://www.reddit.com/r/SubredditDrama/comments/7vy9cw/rdeepfakes_the_aigenerated_fake_celebrity_porn/, (accessed 4 March 2020).
[44] S. Cole, 'Twitter Is the Latest Platform to Ban AI-Generated Porn', *Motherboard*, 7 February, 2018, https://www.vice.com/en_us/article/ywqgab/twitter-bans-deepfakes, (accessed 3 March 2020).
[45] *GitHub*.
[46] *Malavida*, [website], 2019, https://www.malavida.com/en/soft/fakeapp/, (accessed 3 March 2020).

of deepfake by placing my attention on amateur deepfake program FaceSwap as a site of production.

Peeking into the realm of engineers as a layman, I have chosen this program for practical reasons: The codes to amateur deepfake programs like FaceSwap are continuously updated by developers and are designed so that anyone interested in creating deepfake can use them. As such, FaceSwap is one of the more accessible deepfake software designed for less tech-savvy people to start making their own deepfake. The codes of FaceSwap themselves, although were not compiled by u/deepfakes, are created with the original codes released by u/deepfakes. The FaceSwap forum on GitHub[47] and its own website[48] are also vital sites of knowledge for my understanding of the technology. These sites provide detailed practical steps to help users create their own deepfakes, as well as relevant information for machine learning terminology. By looking at these websites as sites of knowledge, it is also revealed that these sites provide insight to how open source knowledge circulate, spread and permeate through online platforms.

How are Deepfakes Made

The main machine learning techniques for deepfakes tend to be a combination of using autoencoders and generative adversarial networks (GANs).[49] GANs are a class of machine learning system in which two neural networks are pitched to compete against each other, thereby making it possible for the machine to learn rapidly. FaceSwap is a software that relies on autoencoders, which is a type of neural network that can learn efficient data coding in unsupervised learning. Within autoencoders, there are encoders and decoders: an encoder maps the input into the code, and a decoder that maps the code to a reconstruction of the original input.[50]

To teach the machine to learn facial expressions from a face and generate a new face as closely as the original face as possible, several steps need to be taken. Images that give prominence to a person's facial features are fed into an encoder, and after processing, the same

[47] *GitHub*.
[48] *Faceswap.dev*, [website], 2019, https://forum.faceswap.dev/, (accessed 3 March 2020).
[49] D. Güera and E. J. Delp, 'Deepfake Video Detection Using Recurrent Neural Networks', *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Auckland, New Zealand, 2018, p. 2.
[50] *Faceswap.dev*, [website], 2019, https://forum.faceswap.dev/viewtopic.php?f=6&t=146, (accessed 3 March 2020).

face is learned and generated by a decoder. This step allows the encoder to create an algorithm that can be used later to reconstruct faces from the input images that it learns from into a representation in the form of a 'vector'. To generate a new face, a decoder then takes the vector and turns the representation into a newly generated face that is as similar to the original face as possible. The faces are not mechanically replicated but are learned faces and generated anew to look as close to the original faces as possible.

The creation process of a deepfake involves two training sets, a target face and a source face. The target is the subject in a video whose face is replaced by the face of the source. The faces first need to be extracted from an image, so the machine can accurately learn which elements from an image it can learn from. Within the extraction phase, the face is detected (finding the face within a frame), alignment is found (which identifies the 'landmarks', such as the eyes, nose, and mouth of a face), and finally, a mask is generated (a 'mask' identifies if the features in an image is a face or the background). Extracting generates a set of faces, called the face set, which it will be used to train the neural network.[51]

Fig. 3 is an image that describes the process of creating deepfake using autoencoders. This image is found in a computer science paper 'Deepfake Video Detection Using Recurrent Neural Networks' and it portrays a simplified flow to how encoders learn faces and swap faces in the decoding stage. In this image, the target is the feminine face (Face A, actress Amy Adams), the source is the masculine face (Face B, actor Nicolas Cage). To clarify, the images show one face to make it seem like the training set consists of a single face, but this is not the case. Training set needs many samples of the same face from different angles and lighting conditions. This image's representation of one face in a training set is merely a demonstration of what the facial features of the faces to be trained appear as.

[51] *Faceswap.dev*, [website], 2019, https://forum.faceswap.dev/viewtopic.php?f=5&t=27, (accessed 3 March 2020).

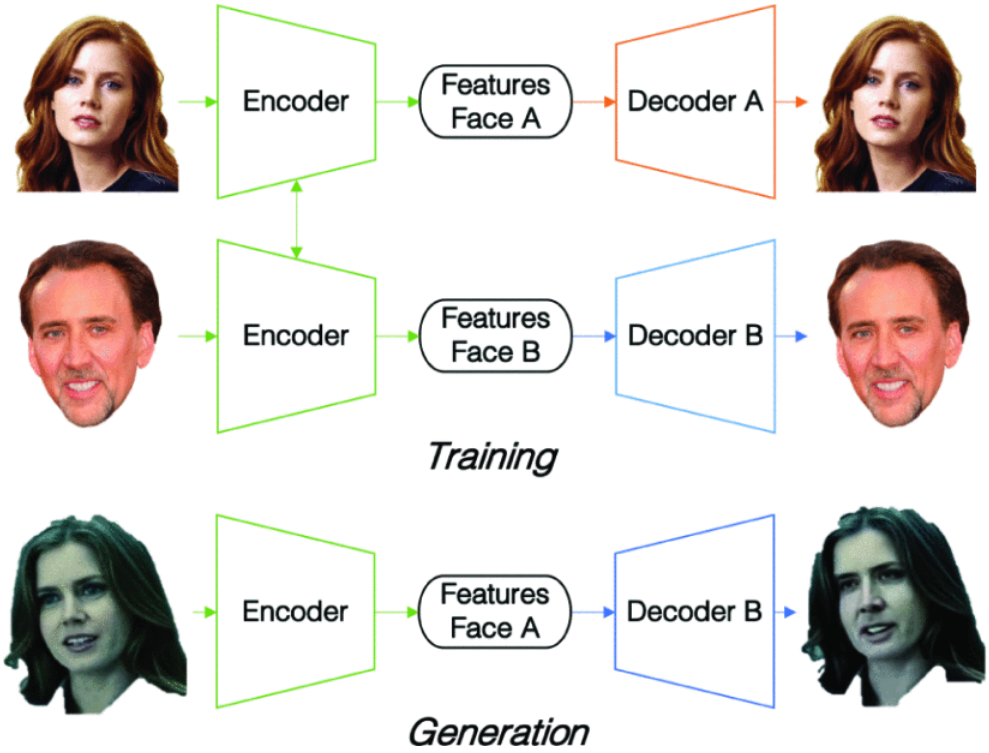

Fig. 3. Image visualizing the process where face swapping with encoders and decoders take place.[52]

The learning of faces takes place during the training stage. The neural network trains two sets of faces – the face of the target and the face of the source – with a shared encoder. This process allows the machine to learn to create a single algorithm that learns from two different people's facial expressions. The same encoder learns to output two different faces into Decoder A and Decoder B. Both decoders can be seen in Fig. 3, having learned to generate the faces they learned from the input. As the training progresses, the program learns to map the different facial expressions of the two sets of faces and learns to replace the faces of the target and the source.

In Fig. 3 the result of the face swapping is shown in the third row of faces, taking place during the generation phase. The program swap faces by using the vector of face A created by the shared encoder, mimicking the expressions and behaviors of face A, but generating the 'look' of the face using output from Decoder B. In Fig. 3, we see target face's features are replaced by source face's features in the output. The facial expression is the same as Face A's (Amy Adams) facial expression, but her features have been replaced by Face B's (Nicolas Cage). The result

[52] Güera, D. and Delp, E. J., *Deepfake Video Detection Using Recurrent Neural Networks,* [online image], https://ieeexplore.ieee.org/mediastore_new/IEEE/content/media/8637085/8639076/8639163/guera2-p6-guera-hires.gif, (accessed 7 February 2020).

faces are taken out of a well-known deepfake that parodies 2013 Hollywood film *Man of Steel*, of which I will describe later in chapter 2.

There are many factors that determine the success of a convincing deepfake. Roughly 500 images of a source are needed to create a deepfake that is persuasive.[53] FaceSwap's website FAQ recommends 500-5000 images per facial angle. In general, if the user is able to provide images with high resolution, high variety of facial expressions and different illumination conditions, this will give the machine a lot of material to learn from. As well as give the machine sufficient time to learn, the processed video could achieve a high level of realism.[54] The time to train a model can take up to 12-48 hours to a week depending on the types of computational hardware a user has. The more sufficient time given to the machine to learn, the better the results of processed video could look.

Despite the promise of hyper realistic deepfake, the deepfakes publicly available on the Internet currently tend to be unconvincing for (mostly) technical and contextual reasons. Many videos are low-quality synthesized faces. Differences in skin tone or lighting of source face and target face may result in visible splicing boundaries and color mismatch in complexion.[55] Many deepfakes can also be detected with the lack of eye blinking[56] and weird face poses[57] due to the training data often being still images of frontal facial features with eyes wide open for aesthetic reasons.

Deepfake as New Media Object

I aim to establish deepfake as a new media object as theorized by Manovich by grounding my analysis of deepfake with his method. In Lev Manovich's seminal book 'The Language of New Media,' he recognizes the disruption to the state of all media with the arrival of computer art. He acknowledges that we live in a present where 'we are witnessing the emergence of a new medium – the meta-medium of the digital computer.'[58] The combination of the shift in technological

[53] A. Hauser and M. Ruef, 'Deepfake Analysis - Amount of Images, Lighting and Angles', *Scip Labs*, 2018, Available from: Academia.edu, (accessed 3 March 2020).
[54] *Faceswap.dev*, [website], 2019, https://forum.faceswap.dev/app.php/faqpage, (accessed 3 March 2020).
[55] Y. Li, et al., 'Celeb-DF: A New Dataset for DeepFake Forensics'.
[56] Y. Li and S. Lyu, 'Exposing deepfake videos by detecting face warping artifacts'.
[57] X. Yang and Y. Li, and S. Lyu, "Exposing Deep Fakes Using Inconsistent Head Poses"
[58] L. Manovich, *The Language of New Media*, Cambridge, Mass: MIT Press, 2001, p. 6.

hardware such as the personal computer, and the arrival of the Internet in 1995 represents 'the most material and visible sign of globalization.'[59] As such, the gradual computerization of culture will eventually transform all that is around us. Manovich defines new media as a representation of what he calls 'the computer layer' converging with 'the cultural layer.'[60] The convergence result of technological rise of computers, transitioning existing media into numerical data. Results in the creation of new media where images, sounds, shapes, and other media becomes computable. He then recognize the language (here the word 'language' is used as an umbrella term 'used by designers of new media objects to organize data and structure the user's experience'[61]) of new media, which aims to describe and understand the logic driving the development of new media by grounding his method 'by placing it [new media] within the history of modern visual and media cultures.'[62]

Borrowing from Manovich's principles of new media, I apply some of these principles to deepfake. The principles of new media include a list of five principles: 'numerical representation, modularity, automation variability, and cultural transcoding.'[63] It is worth noting that not every new media object follows these principles, as they describe and reflect the general tendencies of new media that indicate the meeting of the cultural layer and computer layer.[64] I will be referring back to this principle throughout my analysis.

Created on computers, deepfake originates in numerical form. The numerical representation in Manovich's principle of new media describes new media objects that can be characterized mathematically. New media are made of, and made from, codes. Their materiality is mathematical and programmable. The images of a deepfake are processed through autoencoders, which are themselves composed entirely of codes. Although the source and target training data images could have originated as analogue images, they are nonetheless fed into a machine, and become processed and stored as pixels and mathematical codes. Deepfake is a media with a set of moving images that can be presented with mathematical function.

The modularity of new media is the 'fractal structure of new media.'[65] The principle in which many new media objects can be described as an object that can exist as their own element,

[59] Manovich, *The Language of New Media*, p. 7.
[60] Manovich, p. 20.
[61] Manovich, p. 7.
[62] Manovich, p. 8.
[63] Manovich, p. 20.
[64] Manovich, p. 27.
[65] Manovich, p. 30.

as well as exist as other elements in different mediums. A deepfake media can exist as a video or a gif, or a still image. Because the media exist not as a material object but as a digital object – composed of data and codes – they can pass through wires spontaneously and be split into new duplicated copies (or multiple copies) while maintaining their own independence. The face sets that are fed into a deepfake program are learned, digested and concocted to come out freshly generated as a new face that strongly resembles the face set the machine learned from. And yet the hundreds of images within the face set themselves are also individual new media themselves.

These characteristics of new media then enable 'the automation of many operations involved in media creation, manipulation, and access.'[66] Photoshop, the program of transforming or altering a photograph to achieve desired effects, is so ubiquitous it now has its own verb, 'photoshopped'. Computer-generated imagery (CGI) has also made its place in the entertainment industry. On our phones, there are popular phone apps, such as Instagram, that offer versions of facial filters to alter color or crop photos taken on the phone. They are all forms of digital manipulation and utilize different degrees of what Lev Manovich describes as 'low-level' automation of media creation, where 'the computer user modifies or creates from scratch a media object using templates or simple algorithms.'[67] Which is to say, information generated and assembled through generic, pre-programmed templates. Photoshop, CGI, and phone apps with facial filters use different degrees of software sophistication, but at the end of the day all follow templates with a certain degree of human handling. These programs may automatically digitally manipulate faces but are not deepfake. Instead, deepfake is what Manovich would call a high-level automation of media creation. This requires a machine to understand the semantics embedded in the objects being generated. The autoencoders in amateur deepfake programs still need human guidance to learn to recognize faces, but the autoencoders learn to reconstruct faces through vigorous encoder learning and decoder generating and enable a relatively hands free human interaction by learning against each other.

I have demonstrated that deepfake follows the principles of new media, and here I am taking it one step further and establishing deepfake as new media object. Manovich invokes the word *object* to emphasize his concerns with new media grounds at the culture at large, rather than limiting the definition to new media art alone. New media objects describe 'the general principles

[66] Manovich, p. 32.
[67] Manovich p. 33.

of new media that hold true across all media types, all forms of organization, and at all scales.'[68] Therefore new media objects encompass all types of new media. Deepfake is an emerging visual phenomenon that is a combination of different efforts. Open-source deepfake programs are, as far as we know, developed by and continuously tweaked by an online community of machine learning enthusiasts. Amateur deepfake technology and deepfake media does not belong to any specific organization or person, and deepfake exist across multiple media types (as videos, gifs, still images) online. As such, amateur deepfake thrives in the medium of the Internet.

The Age of Online Content

Canadian media theorist Marshall McLuhan had famously proclaimed that 'the medium is the message,'[69] an assessment that our way of using any type of tool for communication has a bigger impact than the content such medium conveys or delivers. McLuhan assures that technology that transfers the message transforms society and alters the consciousness. 'What we are considering here, however, are the psychic and social consequences of the designs or patterns as they amplify or accelerate existing processes. For the 'message' of any medium or technology is the change of scale or pace or pattern that it introduces into human affairs.'[70] Technology affects the individual and society that uses it. The content of all electronic media is not that shapes or changes our behavior. It is the medium itself that delivers the content that has the greater impact on the environment to which we use every day, and after a while, alter the ways we behave. Hence, alter the way we experience the world. This is not to say we should ignore the content of the media entirely, but rather to study what can be revealed when we study the medium and not just the content.

The medium can be any form of apparatus for communication. In McLuhan's time, the 1950s and 60s, the dominant medium was the telephone, radio, and television. They are physical mediums which are operational by the invention of electricity, an invisible medium. In the 21st century, computers have quickly replaced television as the dominant physical medium, with the Internet as an invisible medium.

---

[68] Manovich, p. 14.
[69] McLuhan, *Understanding Media: the Extensions of Man*, p. 7.
[70] McLuhan, p. 8-9.

The Internet is the invisible medium of our time. In McLuhan's time, the invisible medium is the electric light, which McLuhan describes as 'a medium without the message'.[71] McLuhan points out that the electric light is without content, but the light illuminates content that otherwise would not have been observable without it.[72] The light could be used for brain surgery or night baseball, even if the light itself has no meaning. In other words, the electric light becomes meaningful in that it illuminates something. Indeed, with the introduction of the electric light, human behaviors have changed. The nights are now illuminated with night lights, and natural light no longer dictates human nighttime behavior.

Manovich notes that personal computers are so prevalent that 'the computer layer' – the hardware and software, the data and computer language, code and hardware sockets – and 'the culture layer' – story and plot, point-of-view and compositions – are bound in a spiral dance, influencing each other and coming hand in hand.[73] Not only are computers the new viewing apparatus, it is also a tool that could visualize old media into digital media.[74] These days, new media are created on computers, distributed via computers, and stored and archived on computers. With the prevalence of computers and the Internet comes new ways of entertainment, interaction, work, and life. Having access to computers and the Internet are changing the way people view content and behave. Images can be viewed as online images, shows can be streamed online, texts are read through a screen, etc. The result of this composite is a new computer culture: a blend of human and computer meanings, of traditional ways in which human culture modeled the world and the computer's own means of representing it.[75]

A video can now be viewed on YouTube, while existing on a different social media platform such as Facebook, Instagram, or Reddit. These platforms allow contents to amass different viewership and create different conversations within different online communities. These platforms can be interchangeably viewed on different hardware like a desktop computer, laptop computer, or cell phones. The many different ways where these software systems can deliver content also complicates the way images circulate. Online digital media travel differently from analogue media and from digital media that are not online. Once a digital media is put

[71] McLuhan, p. 8.
[72] McLuhan, p. 8-9.
[73] Manovich, p. 46.
[74] Manovich, p.130.
[75] Manovich, p. 46.

online, it travels by broadband, downloading byte by byte, simultaneously and summoned to the surface of screens everywhere.

It is through this complicated and intricate ecosystem of hardware and digital software, the medium of screens and Internet, that deepfake emerge as a form of user generated content. Deepfake program is created by a Reddit user and improved by other online enthusiasts. Deepfake content, as new media objects, are generated by other users, which is then shared for more users to consume. Deepfake, by its nature, is a successful form of online content that is user generated.

The Medium as Online Content

McLuhan was concerned that media studies place too much emphasis on content and not enough research on the mediums that deliver the content. McLuhan observes that 'the 'content' of a medium is like the juicy piece of meat carried by the burglar to distract the watchdog of the mind.'[76] And indeed, we think and talk about the things we see online more often than we think about *where* and *how* we come into contact with them. Did these contents come to us non-deliberately by scrolling through an online forum with algorithmic design, or did we actively seek them out by searching for them? These questions are of interest as it sheds insights to how deepfake as content are distributed and mediated through the Internet.

The Internet itself does not have a content, but its role as a medium that mediates between mass amounts of information passing through screens to our eyes, has been very definitive of the 21st century way of life. Many daily life activities: Reading, watching, writing, talking to others, navigating the world, consuming media, are all today intricately connected with the Internet. Just like the electric light brightens the night with an electric lamp, the connectivity of the Internet enables a society of individuals that could be linked regardless of physical location or time.[77] The virtual world of the Internet also facilitated a progressively expansive virtual trove of user generated visual media. The Internet connects individual users with screens of desktop computers, personal laptops, or phones. All screens different in sizes but all providing 'personalized' experience for content consumption.

[76] McLuhan, p.18.
[77] McLuhan, p. 9.

The Content of Any Medium is Another Medium

But just because the medium should be the focus of a study, it is not to say that the content from which the medium delivers is not important. McLuhan suggests that '[t]he content of any medium is another medium,'[78] and that '[t]he content of the press is a literary statement, as the content of the book is speech, and the content of the movie is the novel.'[79] Meaning that the content that is brought forth from a medium themselves are a prior medium, transformed from their own worlds and brought forth as content into a new medium to be presented before our eyes. The content of the medium is not only not unimportant, it is also a way to examine how a medium delivers its effect.

Communication scholar Paul Levinson expands and applies McLuhan's thesis to the Internet. Levinson asks: 'What medium serves as content for the Web…?'[80] and answers the question by revealing that the answer is 'not one medium, but many media, for the Web has taken as its content the written word in forms ranging from love letters to newspapers,.., and moving images with sound which can be considered a version of television.'[81] In Levinson's analysis, the written word drives the content of the Web. And that most, if not all, of the 'message' of the medium of the Internet comes from writing. Indeed, most of what comprises as 'content' on the Internet can be boiled down to texts: messages in FAQs, comments and replies, searching for and typing in things. Reduce the display of a website or a video or an image, and they are made of codes that are composed with texts with rules. Therefore, the Internet encourages not only reproduction of content, but also interactions of content between individuals. In 2020, the dominant invisible medium has become the Internet, a global system of connection of networks.

Levinson assures that 'the user is the content of the Internet.'[82] That the online experience is a two-way street where users consume content submitted by other users and themselves engage by submitting more content for other users to consume.[83] The initial development of social networking sites was a 'scaling-down of public broadcasting to become individuals posting to

[78] McLuhan, p. 305.
[79] McLuhan, p. 305.
[80] P. Levinson, *Digital McLuhan: A Guide to the Information Millennium*, New York, Routledge, 1999, p. 37.
[81] Levinson, *Digital McLuhan*, p. 37-38.
[82] Levinson, p. 39.
[83] Levinson, p. 39.

groups'.[84] Normally these group postings would consist of hundreds or so people, and they interact with each other only. The human user creates new content with every interaction they have with the Internet.

Different online platforms attract different online communities and are part of a world where different deepfake content relate and interact with its viewers. Levinson explains that '[t]he heightening of human choice over media on the Internet – to write, to search, to view, to speak,' echoes McLuhan's notion that 'the human subscriber of the Internet chooses which prior medium, serving now as content, to use online.'[85] The many different choices and ways for interaction for both mediates deepfake as the content, deepfake as the medium, and users as part of that intricate system of mediation.

Deepfake and Social Media Platforms

Manovich argues that 'computer operations encode existing cultural norms in their design'[86] and that '[t]he practice of putting together a media object from already existing commercially distributed media elements existed with old media, but new media technology further standardized it and made it much easier to perform.'[87] The medium of social media platforms encourage sharing content and reiteration of old content. Instead of remixing old things together to create something that is a modified or remixed version of the original, social media's upvoting and viewing system encourage people to repost and rehash materials for optimization of viewership. Given how easy it is to copy and paste and post and share new content, it is with this logic that deepfake as content thrives to circulate.

Social media platforms, like real life cliques, have different sets of social rules in the matters of engagement between users, and most users follow a set of rules. Many platforms also have rules that encourage or discourage certain behaviors. Accordingly, the uploading and submission or uploading of deepfake content to different platforms tend to follow certain 'customs' of the sites. I am an Internet user who does not use GitHub or frequent engineer

[84] D. Miller, et al., 'What Is Social Media?', *How the World Changed Social Media*, 1st ed., vol. 1, UCL Press, London, 2016, p. 1–8. *JSTOR*, www.jstor.org/stable/j.ctt1g69z35.8, (accessed 21 February 2020).
[85] Levinson, p. 41.
[86] Manovich, p. 129.
[87] Manovich, p. 130.

websites, therefore certain interactive 'customs' are invisible to me. I can, however, describe the 'rules of conduct' on Reddit and YouTube, as I am a more active user of these sites.

Reddit is a social news aggregation – a platform that collects content from other social networks into one unified presentation – website. Reddit's upvoting system encourage users to post new content to rake in new 'karma', an arbitrary point system where the more upvotes the posts get, the algorithm will encourage more people to see the content, which means more people will see it and upvote it, giving the user more karma points. Certain users can even purchase 'Reddit gold' with real money and gift these golds to users who publish content they like. These contents can be original user generated content (any form of content that is posted by users onto online platforms), or more often, reposted items. Reddit's versatile posting system also allows for multiple media to be reposted, or re-uploaded onto the site, so often people would recycle old content by posting them as a new post to get new karma.

YouTube is a website that offers video hosting services and promotes video content upload and views by rewarding its creators with financial rewards. Reposting is more difficult on YouTube, as there is a more rigorous copyright check and balance. However, since YouTube is one of the most popular and largest video platforms, a lot of its videos get reposted on other platforms, including Reddit. YouTube encourages views and comments through its comment system and is known for its algorithm that aggressively pushes for similar videos a user may like, to entice the user to spend more time on its site.[88]

New media objects like deepfakes are primed to thrive in the medium of social media platforms for their modular structures. Take deepfake video on YouTube as example. The webpage consists of numerous different objects, all consisting separate media elements and are independently functional on its own: A video for the audience to play, images of the video and user, texts, hyperlinks in the form of texts. All are stored independently on a network, to be summoned at the whims of a viewer's click.

It is common for deepfakes to be posted on YouTube first, then posted on Reddit. This way it can collect as wide a range of views as possible. Both platforms have different users that can interact with the deepfake content differently. All these activities feed into an advertising system where the social media can then get advertisement money from marketing companies. Anyone can recycle and rehash any content they find interesting; anyone can be a content

[88] K. Roose, 'The Making of a YouTube Radical', *The New York Times*, 8 June 2019, https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html, (accessed 27 February 2020).

distributor with the click of a button or press of a thumb. Anyone can be a user, or own multiple user accounts. The encouragement of content circulation makes it possible for users to be engaged online, but also enables users to be the content that drives the Internet.

Deepfake as a content also reinforces the behavior of modification and remixing existing content. That iconic scene in a film featuring Actress A? Her face is now replaced with Actor B. This funny show where professional Impersonator A plays a parodic version of Politician B? Now impersonator A's face is replaced with Politician B. The faces deepfake makes are a swap of different faces extracted from different images, all existing originally in their own media and could be used independently before becoming a merged product of deepfake. The videos or gifs of deepfake can be viewed by themselves, or exist online within a platform, or within a website. They can exist in different, and infinite versions. As moving images, videos, gifs. Deepfake recycles images and remixes them into new synthesized media. Social media platforms then allow them to exist in variable online forms. To be downloaded and reposted and remixed into infinitum.

Chapter 2: Amateur Deepfakes

Amateur deepfakes tend to be made by enthusiasts using open-source programs like FaceSwap, which are then shared by uploading onto social media platforms or enthusiast websites. Deepfake is produced the most in two different kinds of content. Some of these deepfake could be videos or in gif format. A few of them may even become viral, circulating rapidly from one Internet user to another. Deepfake technology, although open sourced, is still a relatively niche hobby limited to engineering enthusiasts. The two most commonly use contexts in which deepfake is produced are deepfake memes and deepfake porn. Many deepfake memes are parodic or satirical and are created to provoke a certain type of reaction. Many deepfake porn exist on specialty websites. This chapter discusses these two kinds of content and examines their symbiotic relationship with online platforms.

Deepfake as Memes

Fig. 4 is a screenshot of a gif – a type of graphics file format that contains an animated moving image – demonstrating the result of a deepfake. This particular gif is commonly used by scientific journals as an educational demonstration of the face swapping of deepfake technology. The faces here are used as an example to provide visual explanation of deepfake technology in D. Güera and E. J. Delp's article 'Deepfake Video Detection Using Recurrent Neural Networks', as well as is Wikipedia's default example image.[89] The film clip, the frame in the left featuring female actress Amy Adams' face, is originally taken from Hollywood film *Man of Steel* (2013). The deepfake result, featuring male actor Nicolas Cage's face, is created by Reddit user u/derpfakes, a multiple platform account dedicated to creating Safe For Work – deepfakes that are not pornographic, hence, something safe to view even at work – deepfakes. On Reddit[90] and YouTube[91] are different medium versions of this gif without the side by side comparisons.

[89] *Wikipedia*, [website], https://en.wikipedia.org/wiki/Deepfake, (accessed 9 March 2020).
[90] u/GifFakes, 'Nicolas Cage in Man of Steel by /u/derpfakes', *Reddit*, [website], 6 February 2018, https://www.reddit.com/r/GifFakes/comments/7vntk0/nicolas_cage_in_man_of_steel_by_uderpfakes/, (accessed 12 April 2020).
[91] *Nick Cage DeepFakes Movie Compilation*, [online video], 2018, https://youtu.be/BU9YAHigNx8?t=39, (accessed 12 April 2020).

Fig. 4. A still screenshot of a deepfake comparison gif.[92]

Fig. 4 is an example of a deepfake that is unconvincing for contextual reasons despite the good technical result. Here, a face swap is aptly completed. Adams' body, mannerisms, and facial expressions are still the same as its original input source, but the face is replaced in the form of Cage's face. From a technical side, Cage's face is generated and precisely mimics Adams' expressions. The side by side comparison makes apparent the copy of the Cage face, as the two expressions match each other. The lighting condition of Cage's faces closely matches with the blue toned environment around it, their skin tones are similar, and there are no visible face splices. But take a closer look and one may observe that Cage's facial features do not take into account to proportionate onto Adams' smaller frame. The nose, eyes, mouth all appear to be lifted from a bigger head and slapped onto a smaller head. The juxtaposing Cage's masculine looking face onto Adams' otherwise feminine and petite frame makes the obvious distinction that this video is deepfaked.

In the case of Fig. 4, the humor is built upon the assumption of binary gender performance. Philosopher Judith Butler described gender as 'a kind of imitation for which there is no original.'[93] She explained that binary gender categories such as female and male are

[92] u/derpfakes, '*Deepfake example.gif*', [still frame from online gif], 19 August 2019, https://en.wikipedia.org/wiki/File:Deepfake_example.gif, (accessed 9 March 2020).
[93] J. Butler, 'Imitation and Gender Insubordination' in D., Fuss (ed.), *Inside/Out: Lesbian Theories, Gay Theories,* New York and London, Routledge, 1991, p. 21.

reproduced through performance. Gender identity is based on repetition of performative acts, which in turn defines the meaning of what is considered femininity and masculinity. These performances define social interaction, body language, and facial expressions. The humor in gender-swapping of faces in deepfake memes then is built upon the subversion of expectations of heteronormative performativity. The replacement of Adam's feminine face with Cage's masculine face but still retaining mannerisms that would traditionally be interpreted as feminine expressions aims to subvert expectations of heteronormative performativity. Such subversion of expectations stuns the viewers into disbelief out of how absurd it looks, therefore providing a humorous, and contextually unbelievable, result.

The Site(s) of Circulating Online Deepfake

Gillian Rose calls the site of circulation a study of the how and where the movement of visual image takes place. She gives the example of a film. The film could be first seen in the cutting room, where it is viewed by its editor or director. Once it is finished, the film travels to another site, such as the theater, to be displayed to other audiences.[94] Once the image is produced at a site, the image moves to another site. In the case of digital images, they travel through a diverse array of software and hardware in their production and circulation. Digital images travel differently from material images. A painting could be transported by ships or planes to move across exhibitions. And just like many other digital images online today, the circulation is intensified by the widespread use of social media platforms.[95]

The line for the site of circulation and production blurs in amateur deepfake. To create amateur deepfake, many source materials are downloaded from other online databases and then distributed and redistributed through online websites and social media platforms. Take example of the *Man of Steel deepfake gif*. The original clip is from a professionally produced big budget film. The scene shows a dimly lit scene where Amy Adams' character is having a discussion in a blue tinted room. After the footage is filmed, the film is transferred into an editing room, to be edited and post-produced. Then, it is released in theaters for mass consumption, circulation, and economic gains. After the film finished its theater run, it was released online to circulate. As the

[94] Rose, p. 34.
[95] Rose, p. 35.

film circulated online, it caught the eye of a deepfake user, deepfake creator u/derpfakes. The user then downloaded the film onto his personal computer, went online, and amassed hundreds of thousands of Amy Adams and Nicolas Cage's facial image from Google, YouTube, or other search engines. Their faces are then extracted and fed into a deepfake program. Once the face swapped deepfake video is created, it is then uploaded onto online platforms, such as YouTube. Later on, it was released on Reddit, where it became viral.

Depending on the context from where this *Man of Steel deepfake gif* is circulated and visually consumed can provoke different reactions. The first time I saw the deepfake gif was accidental: I was browsing Reddit and this gif scrolled into my newsfeed. Like many other visual contents, that pops in and out of my visual scope on a daily basis while using the Internet, the experience of seeing it was so banal that I looked at it, reacted to it, and immediately moved beyond it and forgot about it. My only recollection of this encounter was that this gif was one of the first deepfake I have ever seen, and that I thought it was, as many other videos and images I see every day, just another frivolous video media. To me, it was a simple, harmless, fun meme.

The second time I looked at the *Man of Steel deepfake gif* was because I went out of my way to find it. Somehow, this gif has become a textbook image to demonstrate deepfake technology by its frame-by-frame comparison of Adams and Cage's faces. This comparison gif is taken from online encyclopedia Wikipedia and is a frame by frame comparison of a target video and a deepfaked video using two well-known actors. The inconceivability in this deepfake is essential, as the uncanniness of Cage's face on Adams' petite body not only makes it obvious that this is fake, but also makes it possible to notice how uncannily the machine can make the face mimicry to look realistic. By looking at it with deliberation, my experience of viewing deepfake as a content also changed. Instead of seeing it as something that is simply silly and amusing, or that it is just another new thing that's coming out of the Internet on a daily basis, I instead found myself wondering how these types of content circulate, and also, stay relevant, online.

In backtracking my encounter and interaction with this specific deepfake, I am interacting with the Internet and social media platforms, as well as online forums with open access to knowledge, as virtual spaces. Whereby the tools for my entry and interaction to this world being my physical laptop. I type in the key words 'Deepfake Nicolas Cage Superman Reddit' into Google search bar, hit enter, and use the touchpad to navigate through a list of search results presented on the screen. Revealing to me a virtual world of codes and electronic lights that display its form into texts and colors and images in a manner my eyes can understand and

recognize. I click into Google search results one by one and eventually find the Reddit page I have in mind.

Through this interlaying computer layer, I am maneuvering through a series of virtual worlds to find what I need. It is also at this stage that I realize that Reddit, as 'the front page of the internet', is the kind of medium that encourages users to continuously regurgitate content that grabs our attention, which makes excavation of empirical material challenging. Reddit is a social media aggregator platform, meaning that its operational logic is to convert, reprocess, recirculate content as quickly and as often as possible. The same original media content, let's say, a gif, is posted on different subreddits all at once, making maximum viewing impact. The more subreddits an original media content is posted also means it gets more views, so people would oftentimes post the same content in many different subreddits to see which sub it eventually goes viral. After sometimes, a few weeks or months, the same content is reposted to gain new views. This operating logic coupled with its ever expanding nebulous nature of the interconnected webs and online content, makes it so that many of the contents that could be the 'original copies' I have managed to have an encounter in the past eventually get smothered by new content or reposts. This is part of the reasons why I opted to take this image from Wikipedia. I remembered seeing this comparison *Man of Steel gif* on Reddit in 2018, but I could not find it because it has been lost in a sea of newer gifs.

Many materials that circulate on Reddit are hyperlinks, images, or videos. Reddit's home page feature daily top posts from different Subreddits (communities based on certain topics) and is the most visited page on Reddit. Often, they contain content that elicits certain emotional responses. These provoking contents tend to be either political (r/worldnews), humorous(r/funny), rage inducing (r/justiceserved), or bite-sized information (r/interestingasfuck). For a content to stay on the homepage it must inspire shares and upvotes and inspire users to comment and interact with each other. Often an image or video post that has garnered wide attention on Reddit would be reposted months later to gain new karma. Online memes emerge out of this need. Meme templates are not unique, but they are images or videos designed for mass production. A user can then use captions to personalize their stories or jokes onto these templates to circulate in a comprehensive and emotionally provoking way to incite mutual understanding with other individual users who may identify with these memes. Contextually, deepfake meme are often satirical and parodic in nature. Famous public figures such as celebrities, film stars, or politicians are face swapped into situations they themselves have

never been in or participated in. These situations can typically be well known movie scenes that are iconic or ludicrous, making it possible for viewers to easily distinguish the difference between the original video and the deepfaked one.

Deepfake media are not created entirely from scratch. The technology assembles a new media from parts that are ready made. What drives the creation of deepfake content is instead the desire for modified reproduction. Through the act of swapping out faces, the preference or interest of the creator is revealed. A creator that creates a deepfake of actor Nicolas Cage's face on actress Amy Adams' performance from a Superman movie *Man of Steel* must have some kind of knowledge of the existence of this scene. The creator actively participates in deepfake communities, and probably frequents sites where these deepfake media are posted. Just as the user is the content of the Internet, the user is also part of the content of deepfake.

Deepfake creators create deepfake because it is a (currently still) unconventional visual way to stimulate interaction. Deepfake memes are relatable because they can be humorous in nature, but also eccentric enough for them to often go viral. Although the original copy of the *Man of Steel deepfake gif* eventually got lost in a sea of other deepfake memes, it did become more than just another video online to be consumed. The gif is now integrated into the online open source knowledge ecosystem as a featured image in the most well-known open source online encyclopedia Wikipedia and used as a comparison image in academic journals to explain the technological workings of deepfake.

Deepfake as Porn

Deepfake porn makes up the bulk of deepfake content on the Internet. These kinds of deepfakes very often feature female celebrities face swapped with actresses in pornographic films or videos, to appear as if the celebrity is performing sexual acts. Deepfake porn is the first type of amateur deepfake that proliferated and popularized deepfake, and as of the writing of this thesis, it is still the most common type of deepfake content.

Fig. 5. A screenshot of a porn actress with a deepfaked face undressing herself.[96]

This is a 5 minute incest-porn video that is conflating sexual fantasy with Israeli actress Gal Gadot starring as the main protagonist. In Fig. 5, she is lounging lazily on the bed and flirting at the direction of the camera. Gadot is wearing tight jeans, a lace bra peeking through her pink shirt that accentuates her breasts. Her legs are parted enough to suggest she will be taking them off at one point. A hand, belonging to the unseen protagonist, fondles a sex toy close to the frame of the camera. The title of the video is *'Gal Gadot step-sister helps you masturbate with your toy.'*

There are many points that make it easy to deduce that the video is fake. Contextual awareness means that no one who watches it would actually think this is a real video of Gadot having sex with her step sibling. Technical awareness points to the obvious that stepsister-Gadot's moaning and speech pattern sounds nothing like the real life Gadot. But even with the audio turned off, many visual flaws to this video give strong indications to the fact that this is a fake video. At the site of the image, stepsister-Gadot's face appears much less defined compared with the body she is inhibiting. Her face is disproportionate and much too big for her body, as if a mask that encases the original porn actress's head. Certain scenes, such as the screenshot, where the stepsister veers her face to the side, stepsister-Gadot's eyes awfully lose her line of sight. The discoloration and splicing of face can be seen here. Gadot's skin tone is a shade darker than the

[96] *Gal Gadot step-sister helps you masturbate with your toy*, [still frame of online video], 2019, https://mrdeepfakes.com/video/4696/gal-gadot-step-sister-helps-you-masturbate-with-your-toy, (accessed 14 April 2020).

original porn performer PepperXO, and there is visible color splicing right above stepsister-Gadot's eyebrows and PepperXO's forehead. All things considered; this isn't a very well-produced deepfake at all.

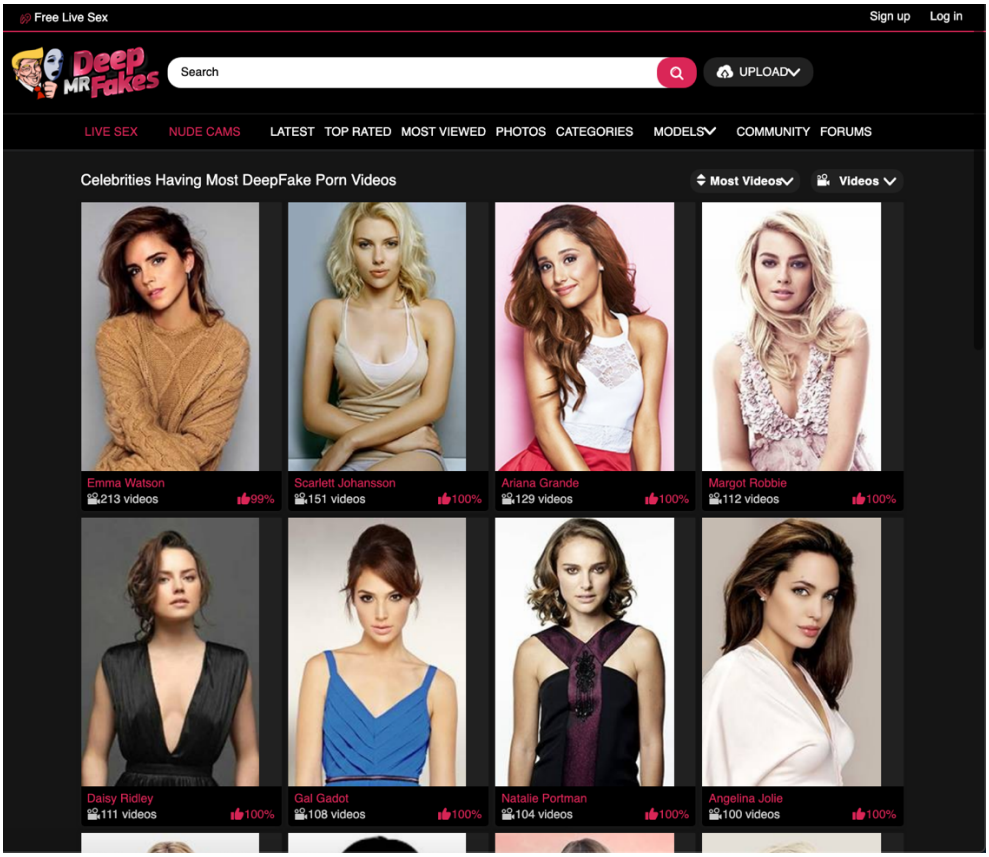

Fig. 6. A screenshot of a page on website MrDeepFakes, featuring a list of celebrities with the most deepfake porn videos.[97]

The final indicator that this video is fake stems from its site of circulation. This video is found on a deepfake porn specialty website called MrDeepFakes, a website that hosts a deluge of deepfake pornography. The website screenshot (Fig. 6) features a page with different profiles of celebrities with their age, height, weight, and various deepfake porn videos headlining their faces catalogued.[98] The semiotic name of MrDeepFakes and its website icon featuring a head of what seems to be a Donald Trump, famous playboy millionaire and controversial U.S. president, putting on a mask of a feminine face moaning for help or of pleasure.

It is also worth noting that although this screenshot only features White actresses, during my own research on MrDeepFakes' home page (a tab listed as 'Celebrity DeepFake Porn Videos

[97] *MrDeepFakes,* [website], https://mrdeepfakes.com/celebrities, (accessed 15 March 2020).
[98] *MrDeepFakes*.

Being Watched'), the majority of videos feature different East Asian and South East Asian celebrities. The implication being that, although uncatalogued, there exists a *market*, or *demand*, for deepfake porn of female celebrities globally and beyond those catalogued actresses.

The website also provides various information such as 'How to make deepfakes', 'Request custom deepfakes', and a 'DeepFake Creators Program' for creators who want to earn money through submitting deepfake videos for users to pay and download. Just as Reddit encourages user interaction with an upvoting system, MrDeepFakes encourages user to create more deepfake by encouraging views and with financial reward. The market is there, and it does not only exist on porn sites.

The materialization and commodification of desire is achieved through deepfake. Deepfake porn is a medium that showcases what Manovich describes as part of new media's principle of variability, that it fulfills the 'postindustrial logic of 'production on demand' and 'just in time' delivery logics that were themselves made possible by the use of computers and computer networks at all stages of manufacturing and distribution.'[99] Instead of relying on imaging of a celebrity performing a sexual act she would never perform on screen, deepfake technology allows for this desire, that which is impossible, to be possible. A viewer or user may desire to see a certain celebrity as a step-sister. In real life this desire remains a fantasy, but with deepfake porn, even though there is an understanding that this is fake, the visual satisfaction is real and can be immediately fulfilled in hours.

The Automated Digital Reproduction of Faces

Deepfake porn and memes overwhelmingly feature faces of public celebrities. The deepfaked faces are forcibly removed from their bodies and put into a false situation. Not only deepfake relies on the 'out of body-ness' to provoke online interaction, the faces themselves – like paintings that are reproduced – lose their 'in situ' uniqueness. In Walter Benjamin's epochal text 'The Work of Art in the Age of Mechanical Reproduction', he assessed that 'even the most perfect reproduction of a work of art is lacking in one element: Its presence in time and space, its

[99] Manovich, p. 36.

unique existence at the place where it happens to be.'[100] A technical reproduction of a painting as a photograph or a digital image of a painting – may be able to reproduce the look of the painting down to its very detail, but the reproduction will lack the unique wear and tear the painting goes through throughout the ages, or the sensation one gets by viewing the painting at the site it is supposed to be admired by.

Furthermore, technical reproduction as Benjamin states, can 'put the copy of the original into situations which would be out of reach for the original itself'[101] and that 'it enables the original to meet the beholder halfway,'[102] via the versatility of technical reproduction. Which further diminishes the value or the original artwork. The *aura*, which is the unique authenticity of the original artwork, is thus lessened in the mechanically reproduction of an artwork. A deepfaked face is placed in a situation that is otherwise out of reach, and it inevitably loses the aura that comes uniquely with the corporeal body, the personality, and the personal temperament that makes the individual unique. The deepfaked person is a simulation, a body equipped with a borrowed face. With deepfake's digitally automated face swapping, the face is forcibly detached from its natural position, and slapped onto an unnatural body, becoming an uncanny unreal creation of the imagination. Images of faces are processed into codes and generated into readily commodifiable bodies and faces.

Comparative to other digital manipulation programs and techniques, the uniqueness of deepfake is not that it can mass reproduce moving images of individuals. Film already reproduces moving images of individuals by its ability to capture the looks and performances of actors. Digital videos further reduce the amount of time for video reproduction, by allowing the images to be instantaneously duplicated into copies.

Deepfake is uncommon for the remixing of faces. It is the type of digital reproduction that not only can generate close to perfect facial expressions of the target and source faces, but it also remixes and recycles its training data into creating entirely new content. The technology doesn't just reproduce, it generates new doppelgangers and remixes existing video mediums.

But why are movie stars and celebrities overwhelmingly targeted in deepfake making? For technical reasons, these celebrities have more readily available or searchable public image data to turn into training sets. But also, there has always been a fascination with celebrity

[100] W. Benjamin, 'The Work of Art in the Age of Mechanical Reproduction', in H. Arendt (ed.), *Illuminations*, New York: Schocken Books, 1969, p. 3.
[101] Benjamin, 'The Work of Art in the Age of Mechanical Reproduction', p. 4.
[102] Benjamin, p. 4.

personalities and this sort of fascination is what makes deepfake relatable online. Benjamin's essay discusses the relationship change of actors and the camera. 'The cult of the movie star, fostered by the money of the film industry, preserves not the unique aura of the person but the 'spell of the personality, the phony spell of a commodity.'[103] How film compensated to the 'shriveling of the aura' by unnaturally creating larger-than-life personalities for the actors.

The audience – the users – enjoy and find images of celebrities engaging, and so celebrities become the first subjects for deepfake training. Any deepfaked face can be amusing, uncanny, but the face of a private person does not resonate and could not compete with the wide range of reaction a face of a larger than life personality celebrities carefully curate in public life could stir. Images of famous people entices emotion, provokes reaction. It is why images of celebrities circulate at bigger quantities and qualities online. It is why celebrities are successfully commodified and created for mass consumption. Part of the spectacle of deepfake comes not by seeing a face that does not belong, but in *recognizing* a face that does not belong. Cage's face does not belong on a female body, Gadot's face is two shades tanner on PepperXO's body. The recognition sparks interest in the viewers, who perceive deepfake as spectacle, and encourage them to react or interact with these media online.

Controversies

Deepfake porn has stirred major controversy in news media outlets. Most notedly, in the question of consent. By deepfaking a celebrity's face onto a porn star's body, consent is circumvented from both the celebrity and the porn actresses.[104] As digital avatars, they raise no objection in suggestive scenarios or lines. No more celebrities saying no to graphic sex scenes or nude performance. These fantasies of seeing certain female faces on certain types of female bodies can be materialized with a program with or without their knowledge. Through deepfake, the porn stars and celebrities transform into a molded, face swapped chimera. As face swapped chimeras, these digital avatars are made to perform as their creators dictate, to be gazed upon and not gaze back. With amateur deepfake, deepfake materializes and embodies the laymen's desire. Deepfake

[103] Benjamin, p. 11.
[104] T. L. Wagner and A. Blewer, '"The Word Real Is No Longer Real": Deepfakes, Gender, and the Challenges of AI-Altered Video'.

porn has not only prompted well known celebrities to speak out about their own faces pasted onto deepfake porn without consent.[105] Deepfake porn has also posed real life implications to private individuals. Such as the case of investigative journalist Rana Ayyub, who was the subject of a series of deepfake porn plots to discredit and silence her.[106]

Deepfake porn is part of many of the tens of thousands of hours of videos that are uploaded onto the web on a daily basis and make up a staggering percentage of deepfake content online. In a report titled 'The State of Deepfakes', found a total of 14,678 deepfake videos online (as of September 2019), through which 96% of deepfakes are pornographic videos. The research found out that videos on 4 pornographic websites dedicated to deepfake porn had over 134 million views, and 100% of the subjects in these videos are female. As such, deepfake porn reveals troubling implications that the female body is to be enjoyed and visually consumed by heterosexual males. Additionally, a third of these subjects feature K-Pop singers, while the rest are mainly British or U.S. actresses or singers.[107] Which suggests that deepfake as a visual phenomenon is potentially global. The gender divide within deepfake video subjects is further illustrated as the report concluded that '[i]n contrast, the non-pornographic deepfake videos we analyzed on YouTube contained a majority of male subjects.'[108]

Out of the hundreds and thousands of deepfake porn, I have selected the '*Gal Gadot step-sister helps you masturbate with your toy'* video because it was one of the first videos submitted by u/deepfakes.[109] It is significant not only for being one of the first deepfake porn, but also that it seemed to have set the precedent for deepfake creators to massively churn out more porn. The Gadot-porn original video copy cannot be found anymore, as not only that it was originally posted on a now-deleted Reddit page, it is now replaced by copies and copies of the same video and uploaded again and again by different users to different porn sites. It actually took me a few tries to find the right keywords for me to find this particular video. At first, I Google searched with the keywords 'deepfake gal Gadot', which lead me to MrDeepFakes. But I had to refine my search and add 'step-sister' and 'sex toy' to find this exact video, as there are hundreds of

[105] S. Hollister, 'Scarlett Johansson slams deepfakes, says she can't stop the internet from pasting her face on porn', *The Verge*, 31 December 2018, https://www.theverge.com/2018/12/31/18163351/scarlett-johansson-slams-deepfakes-internet-lost-cause, (accessed 17 May 2020).
[106] R. Ayyub, 'I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me', *Huffpost*, 21 November 2018, https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316, (accessed 20 May 2020).
[107] H. Ajder, et al., 'The State of Deepfakes: Landscape, Threats, and Impact', *Deeptrace*, 2019, https://deeptracelabs.com/mapping-the-deepfake-landscape/, (accessed 5 March 2020).
[108] Adjer et al, 'The State of Deepfakes', p.2.
[109] Cole, 'AI-Assisted Fake Porn Is Here and We're All Fucked'.

different deepfake Gal Gadot porn on MrDeepFakes. Even in the context of contents, this video I found from DeepFake is most probably another spliced and reuploaded version of the original one.

Hitherto, the online deepfake developer's community have stepped up to the controversy deepfake technology has caused. Both FaceSwap and FakeApp – on GitHub and their own sites – have disclaimers on their pages that the programs are not for creating malicious content. To quote FaceSwap's manifesto: 'FaceSwap is not for changing faces without consent or with the intent of hiding its use.'[110] Explicitly meaning that FaceSwap is not for creating deepfake porn, and a creator cannot claim the deepfaked videos are *real* videos depicting a subject doing things that did not happen. However, as much as FaceSwap's manifesto points to their rejection of illicit, unethical, or questionable purposes, this hasn't stopped people from using them to create porn. Any google search results with keywords 'deepfake + video' will quickly point to more deepfake celebrity porn.

Conclusion to Chapter

The *Man of Steel deepfake gif* and *Gadot-porn* video were chosen as empirical materials because they are both part of deepfake's historic development. The *Man of Steel deepfake gif* is now a commonly used frame by frame educational demonstration. The *Gadot-porn* video was one of the earliest examples of deepfake. They are new media objects that are only found on the Internet and have purely existed on the Internet. Deepfake, be it memes or porn, are distributed on online platforms and rely on vigorous user to user interaction to stay visible amongst a deluge of visual content constantly being uploaded online. Many of the original copies of deepfake content can no longer be found on the Internet, circulating and lost within a sea of ever expanding other independent media copies of themselves.

Deepfake technology demonstrates a unique type of automated digital production in that it takes the material that is fed into the machine, remixes the materials, and creates completely new digital avatars. A face that is deepfaked loses its aura. Like a mechanically reproduction

[110] *GitHub,* [website], https://github.com/deepfakes/faceswap/wiki, (accessed 3 March 2020).

diminishes the aura of an artwork in situ, the process of automated digital face swapping detaches the face and takes away the aura of the original face.

Deepfake's relationship with online platforms not only enable their circulation and consumption, it points to deepfake as compelling user generated content. Deepfake memes are funny, eye catching, and are still unconventional visual content that stimulate interaction online. Deepfake porn caters to a demand for sexual fantasies to be visually fulfilled. There is also a market incentive to creating deepfake, as some sites actively reward creators with either financial gains or attention from other users. Deepfake porn has also roused quite a controversy, mainly because deepfake porn circumvent consent from targeted individuals. While both news media and the online development community have spoken out about the misapplication of deepfake porn, deepfake porn remain the majority of deepfake online.

Chapter 3: Deepfake as Art

Another form of deepfake comes in the shape of corporate high-tech deepfake for the purpose of art. The technical specificities tend to be corporate owned and not open sourced. To create these deepfake, an actor (the target) is hired to stand in as an imitator of the source material character, to which their faces will be swapped out later in post-production. This chapter examines *Imagine This…* by British artists Bill Posters and Daniel Howe and *In Event of Moon Disaster* by directors Francesca Panetta and Halsey Burgund. These deepfakes demonstrate how visually perfect deepfake can look, but also places its focus on the future potentials of deepfake. Unlike amateur deepfake that relies on open-source deepfake programs and tend to be used with generating more content to share online in mind, these deepfakes tend to be commissioned for specific projects, and, as art, have specific themes and issues that it wants to raise. These political themes include digital privacy surveillance capitalism, fake news, and deepfake's potential impact to politics.

Imagine This…

In June 2019, a 16 second video of Facebook founder and CEO Mark Zuckerberg was uploaded onto Instagram, an online photo and video sharing social media platform owned by Facebook. In the video (Fig. 7), Zuckerberg appears to be delivering a message through a news segment from CBSN news streaming channel about the power of a mysterious program *Spectre*. Zuckerberg sits comfortably in what appears to be a generic lounge room at a generic start-up open space office. He leans forward and painstakingly explains: 'Imagine this for a second: One man, with total control of billions of people's stolen data, all their secrets, their lives, their futures, I owe it all to Spectre. Spectre showed me that whoever controls the data, controls the future.' A breaking news style video where Zuckerberg 'confessed' to the storm of controversies Facebook have been a part of in the past year: that the King of Social Media, Mark Zuckerberg, is indeed stealing data from us. As deepfake-Zuckerberg speaks, the breaking news media eye – designed to look like commercial broadcast television network Columbia Broadcasting System, also known as the 'eye network' – quietly observes the viewer, much like the invisible surveillance system around those of us who use social media.

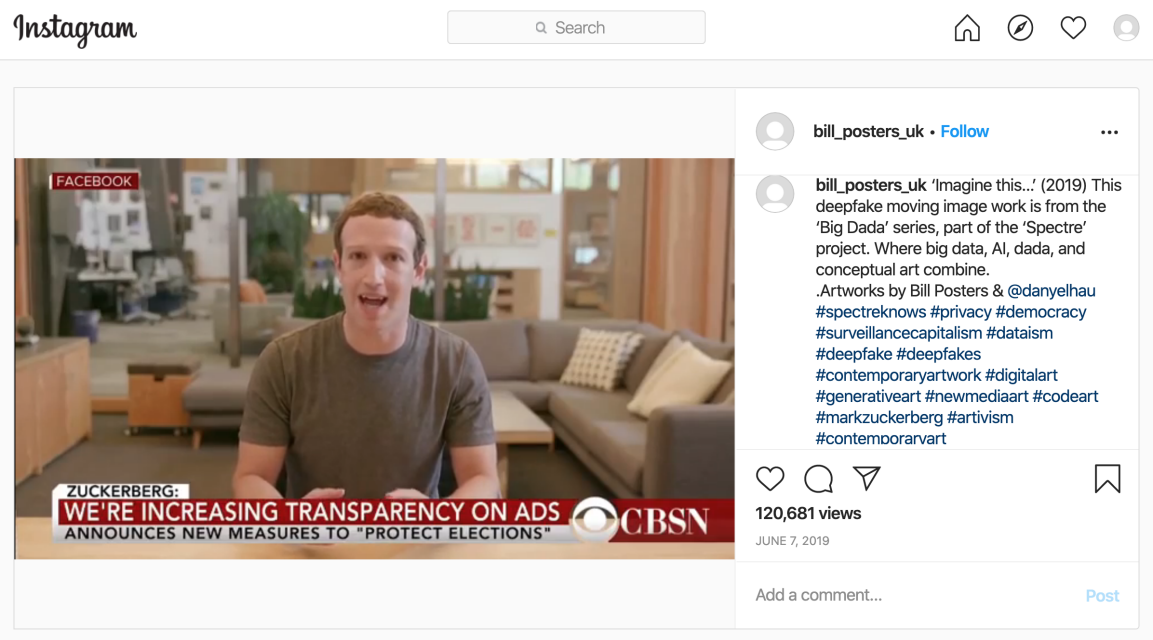

Fig. 7. A screenshot of *Imagine This...* on the desktop version of Instagram.[111]

The video, titled *Imagine this…* (2019), is part of a video series *Big Dada*, where it featured several deepfake synthesized personas of celebrities such as Mark Zuckerberg, Kim Kardashian, Marina Abramović, and others. British artists Bill Posters and Daniel Howe collaborated with actors to read the written texts, and A.I. startups that specialize in creating deepfake videos to create the deepfakes. *Big Dada* is also part of an art series within a physical installation called *Spectre,* an installment of media artworks which aims to talk about how technology intertwined social life by using tactics and methods used by corporation and political campaigns to shape human behavior. The physical exhibition premiered at the Site Gallery, Sheffield, United Kingdom, as part of the Sheffield Doc/Fest 'Alternate Realities' exhibition as a site for the audience to interact and connect with the artwork in June 2019. The film, however, was released on Instagram on 7th June 2019.

The artists made sure that it is easy to discern this film as fake. The contextual suspension from the caption of the film explicitly states that: 'This deepfake moving image work…', and one hashtag literally is #deepfake. Deepfake-Zuckerberg sounds exactly like an actor poorly imitating the real Zuckerberg. However, turn the audio off, it is easy to see how flawlessly the deepfake-Zuckerberg recreates the expressions of the real Zuckerberg. It had none of the amateur deepfake

[111] *Imagine This…*, [still frame from online video], 7 June 2019, https://www.instagram.com/p/ByaVigGFP2U/, (accessed 2 April 2020).

choppy colors or facial imperfections. The eye sights align and there are no facial splices. The film had perfectly created a virtual avatar that recreates the superficial 'spell of personality' of the real Zuckerberg. By creating a vivid Zuckerberg, the film tricks Instagram users to stop their scrolling to watch and observe. He is, after all, the CEO and founder of the world's biggest social media technology company. And what is more curious than the CEO of multiple online digital privacy scandals proclaiming that he is 'Increasing transparency on ads.' Revealing that he is, indeed, spying on the users of Instagram? Deepfake-Zuckerberg was designed to be eye catching, to elicit an initial surprise or confusion to be had when one first watches this film.

Digital In Situ

Imagine This… brings awareness to the attention social media users give social media platforms by demanding the attention itself. The film places its focus on the circulation of fake news and the relationship of the individual with social media. The video is an emulation to look like a sensational news media. It brings awareness of fake news on Facebook, by being itself a type of 'fake news' on social media. The video is packaged to look as if it is a piece of news segment as a bite-sized media content. It is a short viral clip that can highlight a specific statement without taking too much attention away from the audience. Which are the types of content social media sites encourage.

Viewing the film on Instagram provides a rare online in situ experience. Granted, the film has since been reposted on different media sites and that the Bill Poster's Instagram account (@bill_posters_uk) has reposted the same video cropped in different sizes at least twice. But the video is still unique in a sense that it is a new media object where the original copy online is traceable. The viewing experience may be different.

By being published on Instagram, the film can be viewed in various different software and hardware: it can be viewed on a webpage version of Instagram presented by a computer, or on the Instagram app installed onto a phone. The various apparatus for viewings also makes it possible to view the artwork anywhere, anytime. It may be viewed in a public space or at home. Each individual viewing from different screens and in different locations provides a different viewing experience. And yet, all the looking eyes congregate in the same online space – the

Instagram page where the film is published and hosted on – keeping its uniqueness by being in situ.

Imagine This… was released as part of a response to test Facebook's policy on disinformation and misinformation on its platform. To test the 'official responses from Facebook and Instagram concerning their policies regarding computational forms of propaganda on their platforms.'[112] The video was released in adjacent to a viral video controversy on Facebook, where a distorted video of U.S. White House Speaker Nancy Pelosi appears to be drunkenly slurring her words as she gives a speech on stage. The viral video was later revealed to be slowed down by 75% to make her appear drunk.[113] This video is what is known to be called 'shallowfake' – malicious videos and images made crudely with relabeling, photoshop, or video editing that are circulating online – and it is part of tens of thousands of videos circulated online with malicious intent.[114] Facebook has acknowledged the video to be doctored, but did not take it down from its platform.[115] While Facebook later responded that they will not take *Imagine This…* down from Instagram, the company chose to de-prioritize it, so that the short film appears less frequently in user feeds. After its release on Instagram, *Imagine This…* quickly went viral and generated mass media coverage regarding online data privacy and fake news policies on Facebook.

The Facebook–Cambridge Analytica Data Scandal

The release of *Imagine This…* on Instagram, and the film's intentional use of Zuckerberg as its main subject is deliberate. It is to insert itself among the spectacle of the storm of controversy surrounding Facebook's use of private data. Most notedly, pointing to the Facebook–Cambridge

[112] B. Posters, *Bill Posters*.
[113] *Film Club: 'This Video May Not Be Real' Can we trust what we see and hear?*, [online video], 17 April 2020, https://www.nytimes.com/2020/04/17/learning/film-club-this-video-may-not-be-real.html, (accessed 9 May 2020).
[114] B. Johnson, 'Deepfakes are solvable—but don't forget that "shallowfakes" are already pervasive', *MIT Technology Review,* 25 March 2019, https://www.technologyreview.com/2019/03/25/136460/deepfakes-shallowfakes-human-rights/, (accessed 27 April 2020).
[115] E. Stewart, 'A fake viral video makes Nancy Pelosi look drunk. Facebook won't take it down.', *Vox*, 29 May 2019, https://www.vox.com/recode/2019/5/24/18638822/nancy-pelosi-doctored-video-drunk-facebook-trump, (accessed 27 April 2020).

Analytica data breach scandal that broke in March 2018.[116] The scandal revealed that the 2016 U.S. presidential election was heavily influenced by aggressive and persuasive online marketing that targeted individual voters with highly personalized ads. The marketing scheme worked and resulted in the controversial election of U.S. president Donald J. Trump. The hints to this spectacle lie in the texts of the broadcast chyrons, as they sensationally display, in verbatim, 'We're increasing transparency on ads' and 'Announces new measures to 'protect' elections'.

The Facebook–Cambridge Analytica data scandal was of interest to me in that it perfectly demonstrates the push and pulls of social media as medium to the users that use it. The medium of social media saturates around its users. American academic John M. Culkin, in an essay about McLuhan, summarized McLuhan's thesis as: 'We shape our tools and thereafter they shape us. The extensions of our senses begin to interact with our senses.'[117] Users creating content shapes the dynamics of the social media in return, social media reshape our behavior. Corporate players are also users of the Internet. Just as a user can consult another user on how to create a better deepfake on the FaceSwap forum, influential players can influence users by persistent and targeted advertisement to change user behaviors.

Cambridge Analytica is a political consulting firm that illegally harvested 50 million Facebook profiles from Facebook to build a system that could profile individual U.S. voters and target them with personalized political advertisements.[118] Based on the user's Facebook activity, a psychographic profile is created that determines if the individual user is a potential Trump voter or potential swing voter. What is known as 'Persuasion Digital Marketing' is then implemented to market their target: potential Trump voters. For individuals who are likely to vote for Trump, they were shown a 'triumphant-looking image of the nominee, and help finding their nearest polling station.'[119] To the potential swing voters, ads instead present images of people who

[116] M. Rosenberg, N. Confessore and C. Cadwalladr, 'How Trump Consultants Exploited the Facebook Data of Millions', *The New York Times*, https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html, 17 March 2019, (accessed 7 May 2020).
[117] J. M. Culkin, 'A schoolman's guide to Marshall McLuhan', *The Saturday Review*, p. 51-53, 70-72, March 1967, Retrieved from http://www.unz.org/Pub/SaturdayRev-1967mar18-00051, (accessed 4 May 2020). p. 53.
[118] C. Cadwalladr and E. Graham-Harrison, 'Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach', *The Guardian*, 17 March 2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election, (accessed 27 April 2020).
[119] P. Lewis and P. Hilder, 'Leaked: Cambridge Analytica's blueprint for Trump victory', *The Guardian*, 25 March, 2018, https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory, (accessed 27 April 2020).

support Trump and graphic data designed to look like news journalism emphasizing on the corruption of Trump's political opponent, thereby bolster Trump as a better candidate.

This parasitic aggressive advertising based on personal data and individual targeting is noted in American social psychologist Shoshana Zuboff's book 'The Age of Surveillance Capitalism', where she coined surveillance capitalism as 'a new economic order that claims human experience as free raw material for hidden commercial practices of extraction, prediction, and sales.'[120] The medium of the Internet is neutral, but the power players – marketeers, political actors, corporations that extract private data – have used the medium for powerful encouragement for certain behaviors. The Internet as medium is a battlement where different power push and pull for dominance. We, the users of social media, live lives that are so normalized with the entanglement of social media that we didn't even know these targeted, personalized ads are influencing our behavior, ideals, morality, which in terms change the way we vote. But what is even more unnerving is that even non-users of social media are being tracked outside of Facebook.[121] As Levinson has assured, the human user, by interacting with the Internet creates new content. In the laws of surveillance capitalism, these contents are then collected meticulously by tech giants and schemes and influence users into behaving certain ways.

However, despite the whirlwind of controversy, there is a banal acceptance to this type of surveillance. Even though Facebook usage dropped by almost 20% after the scandal broke in April[122], user growth of the site increased by 1.8% by the end of 2018.[123] As of Dec 2019, Facebook has 2.50 billion monthly active users.[124]

In Event of Moon Disaster

[120] S, Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, New York: Public Affairs, Hachette Book Group, 2019.
[121] N, Singer, 'What You Don't Know About How Facebook Uses Your Data', *The New York Times*, 11 April 2018, https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html, (accessed 5 May 2020).
[122] A, Hern, 'Facebook usage falling after privacy scandals, data suggests', *The Guardian*, 20 June 2019, https://www.theguardian.com/technology/2019/jun/20/facebook-usage-collapsed-since-scandal-data-shows, (accessed 10 May 2020).
[123] C. Gartenberg, 'Facebook keeps growing despite scandals and privacy outrage', *The Verge*, Jan 30 2019, https://www.theverge.com/2019/1/30/18204186/facebook-q4-2018-earnings-user-growth-revenue-increase-privacy-scandals, (accessed 10 May 2020).
[124] Facebook, Inc., 'Facebook Reports Fourth Quarter and Full Year 2019 Results', *Menlo Park*, 2020, https://s21.q4cdn.com/399680738/files/doc_financials/2019/q4/FB-12.31.2019-Exhibit-99.1-r61_final.pdf, (accessed 27 April 2020).

In Event of Moon Disaster is an alternative history deepfake short documentary film of then U.S. president Richard Nixon's announcement with regards to the Apollo 11 Moon landing in 1969, directed by Francesca Panetta and Halsey Burgund, and created by the Massachusetts Institute of Technology Center for Advanced Virtuality.[125] The film premiered at the IDFA DocLab on November 22, 2019 as a physical installation. In 1969, in anticipation of the world's first attempt at landing the moon, Nixon had a speech prepared in a memo called 'In Event of Moon Disaster', to be read in case the Moon landing was unsuccessful. In this screenshot (Fig. 8), Nixon is sitting by a desk with a solemn expression. He holds a heap of papers which he reads from and addresses the viewers directly. In this alternative history, the moon landing attempt has failed, and the astronauts are left stranded on the moon to die. Nixon solemnly reads the content of the memo, instead of the real, historic out-of-the-world-phone-call to the moon from the Oval Office.[126]

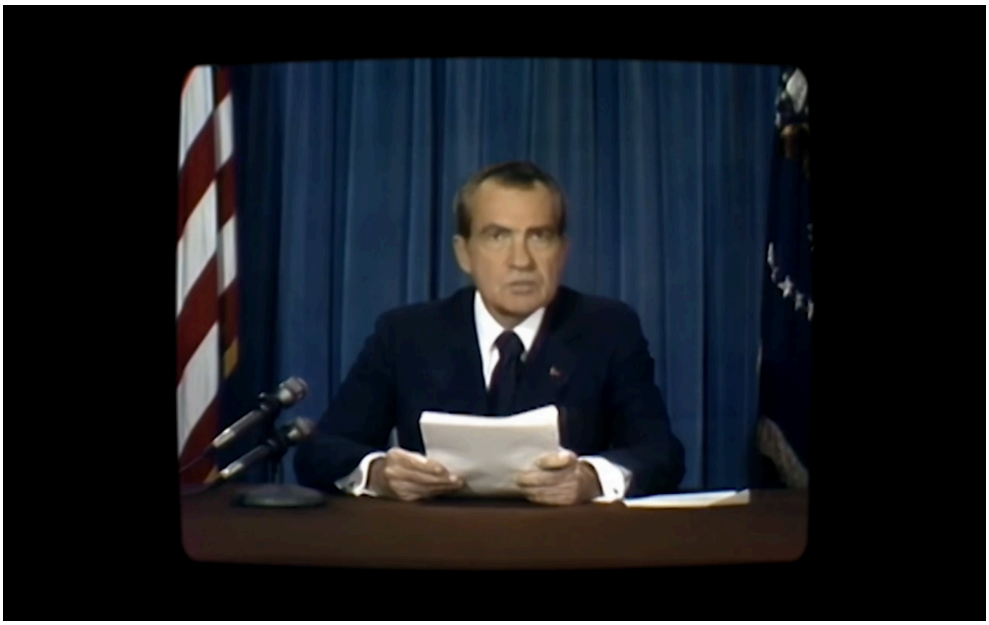

Fig. 8. A screenshot of *In Event of Moon Disaster*, president Nixon addressing the nation of the moon landing disaster. Courtesy of directors Francesca Panetta and Halsey Burgund.[127]

Technological speaking, the film had achieved stunning results. An actor was hired to read the text from the memo while doing an impression of Nixon, filmed sitting in front of a

[125] *In Event of Moon Disaster*.
[126] It is interesting to note that the first Moon landing took place in the pre-internet stage and that Nixon used a telephone call to reach the astronauts on the Moon. *President Nixon speaking with astronauts Armstrong and Aldrin on the Moon*, [online video], 30 March 2011, https://www.youtube.com/watch?v=1Ai_HCBDQIQ, (accessed 12 May 2020).
[127] *In Event of Moon Disaster - PREVIEW*, [still frame from online screener], 2019, (accessed 8 April 2020).

camera that records his voice and facial expressions. Physically, the actor looks nothing like Nixon. And yet, a splitting image of Nixon – his facial expressions, voice, hands, clothes, even the backdrop – was completely generated based off of archival photos, videos, and audio recordings of Nixon during his term of office as U.S. president. The actor was completely replaced by the simulacra of Nixon. The filmmakers collaborated with Ukrainian-based company Respeecher to produce synthetic speech, and Israeli company Canny AI to replicate Nixon's mouth and lips.

In Event of Moon Disaster is about the anxiety of rewriting history using deepfake. The filmmakers chose to use the Moon Landing for several reasons: it is comparatively apolitical which makes it unlikely to alienate potential viewers, and that space is a fun topic, and the Moon Landing is a widely recognized historical event. With this film work, the team wanted to raise awareness to deepfake's ability to blur the lines between what is truth and fake.

Deepfake as the Medium

The joining of mediums – the medium of television and this new medium of deepfake – is done by a physical art exhibition. *In Event of Moon Disaster* is currently the only deepfake in this thesis that is not released on social media. The physical installation (Fig. 9) was made to resemble a 1960s 'average' American family's household living room at the IDFA DocLab. The audience are immersed in this environment.

Fig. 9. An image of the installation of *In Event of Moon Disaster* as a living room.[128]

The 'in situ' installation attempts to capture the spell of the character, the aura of the decade. Against vintage wallpapers, a retro television was set upon a stand. In between the audience and television is a table, where a big document and a stack of newspapers are placed. The newspapers on the table are texts to raise public awareness of deepfake and educate the audience to identify misinformation in the form of deepfakes. The newspapers also give info on how this deepfake was created, and the current technologies researchers are developing to detect deepfake. This gives modern audiences the context that, first, this video is fabricated, and two, a sensorial experience to imagine what the news could be like if deepfake is more common. But by being at the site – a physical installment space within a documentary film festival – the audience already expects and, if not embraces, acknowledges that this piece is a reenactment. A state of suspension of disbelief is a prerequisite to watch this film. This physical installation attempt to simulate a situation where 'one could be watching deepfake, live.' Invoking fear for deepfake to become a tool where historical videos and images become difficult to distinguish.

In *In Event of Moon Disaste*r, the merging of deepfake and television signals deepfake as powerful content showcased to the audience via the medium of the television. Medium is the

[128] F. Panetta and H. Burgund, *IDFAinstallation7.jpg*, [digital image], (accessed 11 April 2020).

message. McLuhan observed that the effect of the television changed society more than the content it delivers.[129] A 1960s American may go home before 6pm to catch a show. A family may arrange the living room so every member in the family can watch television together. In 2020, individuals now have the option to watch television shows on streaming websites and are no longer restricted by the time the show airs. The medium we chose to spend our time with will change the way we use our senses. As such, the medium that carries the message should be the focus of study itself, as it reveals more about how the medium affects society. Discourse around deepfake often falls into the trap McLuhan is wary about. Namely, the content hijacking the study of the medium. In the discourses related to deepfake, the content of deepfake has often outshined the medium of deepfake. The content of *In event of Moon Disaster* points out the potentiality of deepfake weaponized as fake news. But deepfake is also a content of the Internet, born out of social media. And social media is a bigger threat to spreading fake news than deepfake itself.

The merging of deepfake and television also brings forth the focus of deepfake as content towards deepfake as medium. A deepfake is a new media object, it is mathematical and programmable and can exist in its own element as well as other mediums. In the case of *In Event of Moon Disaster,* the content deepfake delivers as medium is a remixing of the synthesized footage of the impersonator actor and the archival audio and visual footage of Nixon. Moreover, the content within this remix of mediums reveals a plot for potential political fabrication.

The filmmakers use this film to explore the potentiality of deepfake's ability to influence one's perception of history. With the installation of the television in a family's living room, the film signals that deepfake could become an integral part of daily lives. It brings awareness that deepfake could become 'fake news' itself and join the prevalent and increasingly intricate network of fake news circulation. Deepfake can also create more room for skepticism. In a sense, the reputation of deepfake already has an effect. If a public person, perhaps a male public figure, proclaims that the footage of himself doing inappropriate conduct is deepfake, who is to say if this is true or not? More skepticism allows more room for a liar to simply proclaim that their actions were deepfaked, making it harder for them to be held accountable.[130] And already, the potential for political disruption the word 'deepfake' carries can be seen in the coup attempt in Gabon.

[129] McLuhan, p. 19.
[130] Film Club: 'This Video May Not Be Real'.

The Coup Attempt in Gabon

Already, the consequence of skepticism regarding a video, whether it is deepfake or not, already has consequences. In January 2018, a military coup was inspired and took place 7 days after Gabon's President was rumored to have released a deepfake video. The President of Gabon, President Ali Bongo, had disappeared from public eyes late 2018 to receive medical treatment in Saudi Arabia and London. After months of absence and little information from the government, it was suddenly announced in December 2018 that Bongo had suffered a stroke in autumn but remained in good shape. And that he will be making his customary New Year's address. But the skeptics weren't convinced. After all, if the president was doing well, why hasn't he made a public appearance? Despite it all, New Year's rolled by and the president released a video. In the broadcast, Bongo is seated in front of a pink backdrop. The camera cuts between two-front camera angles, a common editing style employed to make an otherwise boring interview more interesting, but also used to seamlessly edit different clips into a seemingly continuous take. In other words, Bongo's address could be filmed in multiple takes. Bongo's critics noted the odd body movement of the president: His eyes blinked just 13 times in the 2 minute video, and his eye movements didn't match his mouth, and his upper face barely moved at all. There is even suspicion that his speech pattern was different.[131] Bongo's critics weren't sold. A week later, Gabon's military attempted an unsuccessful coup citing that the video was a deepfake based on the odd behavior of the president. The coup was quickly condemned and resolved.[132] Through investigation, it is revealed that the rumors of deepfake were actually spread by the Gabonese diaspora trying to incite a revolution abroad.[133]

But no matter if the video was fake or real, the skepticism surrounding this video has displayed a prime example of the political power of deepfake. Individuals that do not have robust knowledge of digital literacy or if they live in areas where there is a lack of critical sources that can educate or fact check sources, will have to rely heavily on companies like Facebook and

---

[131] A. Breland, 'The Bizarre and Terrifying Case of the "Deepfake" Video that Helped Bring an African Nation to the Brink', *Mother Jones*, 15 March, 2019, https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/, (accessed 27 April 2020).
[132] 'Gabon coup attempt: Government says situation under control', *BBC News*, 7 January 2019, https://www.bbc.com/news/world-africa-46779854, (accessed 20 May 2020).
[133] 'Breaking Bongo', *Radiolab*, [Podcast], WNYC Studios, 17 November 2019, https://www.wnycstudios.org/podcasts/radiolab/articles/breaking-bongo, (Accessed 27 April 2020).

YouTube to counter the effects of fake news. While no one is immune to misinformation, individuals or groups that are surrounded with robust, independent press and other democratic checks are less likely to be susceptible to government's attempt to manipulate its population with deepfake videos.[134]

Conclusion to Chapter

To quote back to Marshall McLuhan's 'the medium is the message'[135] in simplistic terms, it is the idea that the audience can only receive a story or idea through a medium. It is not so the idea that is expressing itself, but the media that is expressing the story. So, what exactly kind of medium is deepfake as art? What message can be communicated through deepfake?

As a medium, the content deepfake brings forth are interesting as they are remixed old media that are generated into new assemblage. To these artworks, the message is political. *Imagine This…* is a tale of technical surveillance power-fueled by tech giants that is part of the omnipresent medium of the Internet. Released on Instagram, the film shed light on the influence of social media and joins the spectacle of Facebook controversies. *In Event of Moon Disaster* is a tale of the possible consequence of deepfake becoming the norm, bringing front the anxiety of the technology's potential to cause great political harm. Both the artworks explicitly created a 'space' to which the audience have clues to the artwork being fabricated images of real individuals. *Imagine This…* exists in a virtual space, *In Event of Moon Disaster* exists in a physical space.

Both arts are sensational with clear messages. In *Imagine This...*, the human users become raw material for the capitalist machine. Our attention and private lives are the raw material mined and processed into data to either purchase more stuff the companies are trying to sell, and/or become part of an increasing surveillance database in big governments. The art is a call to attention to this phenomenon. In a sense, the deepfaking of Zuckerberg itself is a metamedium in that it demands the same attention to the issue that it is demanding our attention from. The film showcases what deepfakes can do to demonstrate the power of technology by making the king of Facebook as a victim of the technology.

---

[134] M. Westerlund, 'The Emergence of Deepfake Technology: A Review', p. 40-53.
[135] McLuhan, p. 7.

In *In Event of Moon Disaster*, the thrill lies in the fabrication of history and the potential political upheaval deepfake can create. By expertly recreating then U.S. president Nixon's voice and image from archival footage, the film demonstrates how sophisticated this technology can be and endeavor to educate its viewers. A deepfake is a deceptive video created with machine learning that shows a subject doing or saying things they never did. By electing Nixon and appointing him as the victim of historical fabrication, the film successfully attracts attention to the message that: if historical figures can be digitally manipulated to perform an act that has never been done before, the agency images have as evidence will have been taken away. That this technology will turn truth into lies, sway political outcomes and destroy democracy. The technology could even become integrated to the progressively chaotic online network of fake news.

With *In Event of Moon Disaster*, deepfake as a medium manifest a desire for certain narratives. Like regular Internet users and corporate users all have a playground on social media, deepfake in the hands of certain people can be frivolous, sexy, or potentially causing great havoc to democracy.

But art legitimizes technology and what it does. So, what is the point of deepfake art? It is a portrayal of the 'fear of technology'. The sensational, acute freight of the uncertainty technology poses in the wake of deepfake. That deepfake technology can be skillfully done by a few individuals and turn over the perception of reality for everyone else. These works demonstrate the potential of this technology by bringing to attention the big political concern of deepfake. Both films also point to the medium that surrounds us by alluding to the topic of fake news. *Imagine This…* places its focus on Facebook's Instagram. *In Event of Moon Disaster* materializes deepfake via a television installation. Although social media and television are different mediums, they are nonetheless installed and implemented in our everyday lives. This connectivity makes social media a fertile ground for malicious fake news. But without social media's instant connection to different communities, deepfake technology may never have been developed.

Conclusion

This thesis aims to understand deepfake as new media object and how deepfake circulate and are consumed. To ground my thesis, I leaned heavily on Lev Manovich's new media theory to deconstruct deepfake as new media; and Marshall McLuhan's theory on medium which I apply to my analysis of the Internet, where deepfake circulate. As a result of my analysis, this thesis proposes that the emergence of deepfake is a new media object based on its numerical origin, modularity, and variability. This primes deepfake as ideal user generated media that relies heavily on its symbiotic circulatory relationship with the Internet. With major players – the user, the technology, the Internet – influencing the other by playing a part at being mediums and contents. I examine deepfake by observing deepfake media consumption within social media platforms, open knowledge transmission on online forums, and individual enthusiast producers and distributors of deepfake. All these factors contribute to the rise of deepfake as a technology and deepfake as user generated online content.

Soon into my research, I realized that the production, circulation, distribution and consumption of deepfake technology are considerably interconnected with the online platforms. Deepfake is a new media that has so far primarily existed online. The Internet itself is nebulous in nature. As user generated content, deepfakes are reposted in different places, sometimes at different times. Occasionally, original posted website pages gets deleted. This makes finding information, such as a more comprehensive record of how deepfake emerge or tracking down the discussion threads where the specific deepfake is originally uploaded, challenging. However, the expansive nature of the Internet also allows different communities to thrive. As such, online deepfake content have seem to organically grown into peculiar contexts.

In Chapter 1, I deconstructed deepfake technology and piece together a short history of the technology's development first by establishing the historical precedents of face swapping. I look to amateur deepfake program FaceSwap for my understanding of the technology. Originally developed by, whom we think might be, a lone engineer, the technology has gained wide recognition and has become a part of a growing online digital new media ecosystem. technology was developed online and many of the deepfake videos and gifs today can also be found online. The Internet constructs new forms of communication and behavior, but it is from this interactive space where users post, tinker, and collaborate, that deepfake technology is developed.

I then interpret deepfake as new media object following Lev Manovich's principles of new media. Deepfake media are created by a software, therefore deepfake originate in numerical form. As codes and programmable mathematics. Deepfake media are the codes represented on screen in a way our eyes recognize: as images. Deepfake can exist in its own element, as a video or gif. Like many new media objects, deepfake can function independently as its own. And the modularity allows deepfake media to be instantaneously copied, allowing it easy to be posted online for circulation. McLuhan's seminal phrase 'the medium is the message'[136] runs through the veins of my method and theory. He observes the change of the 'invisible environment' technology created which shapes & changes our behaviors. To which I base my analysis of relationships of deepfake with the Internet, social media platforms, and as an online object. Paul Levinson asserts that the content of online content – the product of digital tools – and relates them to the distributor of these contents, which are, the users.

In Chapter 2 I examined several deepfake media. There are two types of amateur deepfake content: deepfake porn and memes. As with the nebulous nature of the Internet, many videos are reposts or re-edits of the original video, making it difficult to track down the original copy. For this reason, I chose two well-known deepfakes as they are both part of the historical development of deepfake technology. *The Man of Steel deepfake gif* is a commonly referred to comparison image used in in academic journals. I consider it as a successful image to demonstrate deepfake technology as it imposed the face of a well-known male actor onto the body of a well-known actress. The inconsistency of a masculine face imposed onto a female body creates humor by subverting heteronormative gender performativity, but also makes the deepfake contextually obviously fake. I then examine the site of circulation of the *Man of Steel deepfake gif*, which is Reddit. Noting that the medium of Reddit not only encourage deepfake content to flourish through user interaction and constant reposting, but also makes finding the original version of the deepfake difficult. The other amateur deepfake examined is the *Gadot-porn video*. It is one of the first deepfake videos uploaded by the original developer of deepfake technology. Problematic nature of unsolicited face swapping of faces onto porn aside, deepfake porn reveals an online fascination with public faces. A face that has been deepfaked loses its 'aura'. It is forcibly removed from its original location, its in situ, and as such loses the unique personality that a face expresses that comes with the person.

[136] McLuhan, p. 7.

Deepfake creators are incentivized by a feedback loop of online interactions: learning how to use the technology online, creating the deepfake, posting it online, gaining recognition and communication from other users, creating more deepfake... etc. There is also the fascination the technology holds that makes deepfake media noteworthy. Images of celebrities offer more training data and broader recognition, which stimulate more interactions between users. The reproduction of faces from deepfake technology is part of new media object's characteristics. The technology recycles the faces through its encoder and decodes a remixed production. In short, deepfake technology create new visual content by amassing an assemblage of facial expressions. There is also market incentive for the creation of amateur deepfake. As such, the symbiosis relationship between users and the Internet created the environment where deepfake could emerge as a form of content that is user generated.

In Chapter 3, I discuss high-tech deepfakes by examining two artworks produced with deepfake technology. These deepfakes are strategically distributed. Through crafted dialogue and constructed sites of distribution, the deepfake films *Imagine This…* and *In Event of Moon Disaster* tell different political stories. *Imagine That…* aims to draw attention to surveillance capitalism. *In Event of Moon Disaster* draws attention to how technology obfuscates the truth. The technology of deepfake here thus serves as a tool to tell the cautionary tale of technology, specifically, the medium that delivers the deepfake content: social media.

Imagine This… demonstrates the digital in situ of a film published on Instagram and brings awareness to surveillance capitalism. Social media changes and shapes and molds the behaviors of its users, and it reorganizes and reprograms human behavior. The more users are surrounded in the medium, the more reality and behavior are distorted with calculative algorithms fueled with surveillance capitalism. *In Event of Moon Disaster*, the deepfake content is watched through a physical television in an installation. The melding of deepfake as the content of the physical television medium reveals the powerful political potential of deepfake: deepfake as part of the broad system of online fake news, as a daily life norm.

Deepfake as a visual and media phenomenon is intriguing and should continue to be researched. As briefly brushed upon in the thesis, there are big ethical questions regarding the production of deepfake porn on targeted individuals. These questions could also be extended into deepfake meme, as any public figure (male or female) that is deepfaked could and perhaps should be considered a victim of deepfake. Deepfake delivers powerful message by being both content

and medium. Although deepfake is a relatively new invention, its meaning and usage are many and should continuously be explored.

Bibliography

Books

Benjamin, W., 'The Work of Art in the Age of Mechanical Reproduction', H. Arendt (ed.), *Illuminations*, New York: Schocken Books, 1969, p. 1-26.

Butler, J., 'Imitation and Gender Insubordination' in D., Fuss (ed.), *Inside/Out: Lesbian Theories, Gay Theories*, New York and London, Routledge, 1991.

Levinson, P., *Digital McLuhan: A Guide to the Information Millennium*, New York, Routledge, 1999.

Manovich, L., *The Language of New Media*, Cambridge, Mass: MIT Press, 2001.

McLuhan, M., *Understanding Media: the Extensions of Man*, Cambridge, Massachusetts, MIT Press, 1994.

Mirzoeff, N., *An Introduction to Visual Culture*, 2nd edn, USA and Canada, Routledge, 2009.

Rose, G., *Visual Methodologies: An Introduction to Researching with Visual Materials*, London, Sage, 2001.

Zuboff, S., *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, New York: Public Affairs, Hachette Book Group, 2019.

Electronic Journal Articles

Chesney, R. and Citron, D., 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security', *California Law Review, Inc.*, 2019, Available from DOI: https://doi.org/10.15779/Z38RV0D15J, (accessed 14 May 2020).

Chesney, R. and Citron, D., 'Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics ', *Foreign Affairs*, January-February 2019, v. 98, iss. 1, p. 147-55, (accessed 15 February 2020).

Dixon Jr., H.B., 'Deepfakes: More Frightening Than Photoshop on Steroids', *Judges' Journal, vol. 58*, no. 3, 2019, p. 35–37.

Güera, D. and Delp, E. J., 'Deepfake Video Detection Using Recurrent Neural Networks', *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Auckland, New Zealand, 2018, p. 1-6.

Hauser, A. and Ruef, M., 'Deepfake - An Introduction', *Scip Labs*, 2018, Available from: Academia.edu, (accessed 3 March 2020).

Hauser, A. and Ruef, M., 'Deepfake Analysis - Amount of Images, Lighting and Angles', *Scip Labs*, 2018, Available from: Academia.edu, (accessed 3 March 2020).

Holzer, H., 'How the Printmakers Saw Lincoln: Not-So-Honest Portraits of "Honest Abe"', *Winterthur Portfolio* Vol. 14, No. 2 (Summer, 1979), p. 143-170, JSTOR:www.jstor.org/stable/1180612, (accessed 25 March 2020).

Li, Y., et al., 'Celeb-DF: A New Dataset for DeepFake Forensics', 2019, Available from: arXiv:1909.12962, (accessed 2 March 2020).

Li, Y., et al., 'Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics', 2019, Available from: arXiv:1909.12962, (accessed 2 March 2020).

Li, Y. and Lyu, S., 'Exposing deepfake videos by detecting face warping artifacts', 2018, Available from: arXiv:1806.02877, (accessed 2 March 2020).

Lum, K., 'From Analog to Digital: A Consideration of Photographic Truth, 2012.', *Everything is Relevant: Writings on Art and Life*, 1991-2018, p. 213-223, Montreal: Concordia University Press. doi:10.2307/j.ctvwvr2jd.37, (accessed 4 March 2020).

Marconi, F., 'Enablers: The AI Technologies Driving Journalistic Change', p. 55-128, *Newsmakers: Artificial Intelligence and the Future of Journalism*, Columbia University Press, New York, 2020, JSTOR: www.jstor.org/stable/10.7312/marc19136.7, (accessed 12 May 2020).

Maras, M. and Alexandrou, A., 'Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos', *The International Journal of Evidence & Proof*, 2019, Vol. 23(3) p. 255–262, Available from: SAGE journals, (accessed 5 February 2020).

Miller, D., et al., 'What Is Social Media?', *How the World Changed Social Media*, 1st ed., vol. 1, UCL Press, London, 2016, p. 1–8. JSTOR, www.jstor.org/stable/j.ctt1g69z35.8, (accessed 21 February 2020).

Ovadya, A. and Whittlestone, J., 'Reducing malicious use of synthetic media research: Considerations and potential release practices for machine learning', 2019, Available from: arXiv:1907.11274, (accessed 12 May 2020).

Tolosana, R. et al., 'DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection', 2020, p. 1-15. Available from: arXiv:2001.00179, (accessed 03 March 2020).

Wagner, T. L. and Blewer, A., '"The Word Real Is No Longer Real": Deepfakes, Gender, and the Challenges of AI-Altered Video', *Open Information Science, Volume 3: Issue 1*, 19 July 2019, p. 32-46. DOI: https://doi.org/10.1515/opis-2019-0003, (accessed 6 February 2020).

Westerlund, M., 'The Emergence of Deepfake Technology: A Review', *Technology Innovation Management Review, Vol 9, Iss 11*, 2019, p. 40-53. Available from: Directory of Open Access Journals, (accessed 23 February 2020).

Winter, R. and Salter, A., 'DeepFakes: uncovering hardcore open source on GitHub', *Porn Studies*, p. 1-17, 2019, DOI: 10.1080/23268743.2019.1642794, (accessed 26 February 2020).

Yang, X., Li, Y. and Lyu, S., 'Exposing Deep Fakes Using Inconsistent Head Poses', *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, Calgary, Alberta, Available from: arXiv:1811.00661, (accessed 2 March 2020).

Newspaper Article

Ayyub, R., 'I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me', *Huffpost*, 21 November 2018, https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316, (accessed 20 May 2020).

Breland, A., 'The Bizarre and Terrifying Case of the "Deepfake" Video that Helped Bring an African Nation to the Brink', *Mother Jones*, 15 March 2019, https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/, (accessed 27 April 2020).

Cadwalladr, C. and Graham-Harrison, E., 'Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach', *The Guardian*, 17 March 2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election, (accessed 27 April 2020).

Christopher, N., 'We've Just Seen the First Use of Deepfakes in an Indian Election Campaign', *Vice*, 18 February 2020, https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp, (accessed 19 February 2020).

Cole, S., 'AI-Assisted Fake Porn Is Here and We're All Fucked', *Motherboard*, 11 December 2017, https://motherboard.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn, (accessed 3 March 2020).

Cole, S., 'Twitter Is the Latest Platform to Ban AI-Generated Porn', *Motherboard*, 7 February, 2018, https://www.vice.com/en_us/article/ywqgab/twitter-bans-deepfakes, (accessed 3 March 2020).

Culkin, J. M., 'A schoolman's guide to Marshall McLuhan', *The Saturday Review*, p. 51-53, 70-72, March 1967, Retrieved from http://www.unz.org/Pub/SaturdayRev-1967mar18-00051, (accessed 4 May 2020).

'Gabon coup attempt: Government says situation under control', *BBC News*, 7 January 2019, https://www.bbc.com/news/world-africa-46779854, (accessed 20 May 2020).

Gartenberg, C., 'Facebook keeps growing despite scandals and privacy outrage', *The Verge*, Jan 30 2019, https://www.theverge.com/2019/1/30/18204186/facebook-q4-2018-earnings-user-growth-revenue-increase-privacy-scandals, (accessed 10 May 2020).

Hern, A., 'Facebook usage falling after privacy scandals, data suggests', *The Guardian*, 20 June 2019, https://www.theguardian.com/technology/2019/jun/20/facebook-usage-collapsed-since-scandal-data-shows, (accessed 10 May 2020).

Hollister, S., 'Scarlett Johansson slams deepfakes, says she can't stop the internet from pasting her face on porn', *The Verge*, 31 December 2018, https://www.theverge.com/2018/12/31/18163351/scarlett-johansson-slams-deepfakes-internet-lost-cause, (accessed 17 May 2020).

Johnson, B., 'Deepfakes are solvable—but don't forget that "shallowfakes" are already pervasive', *MIT Technology Review*, 25 March 2019, https://www.technologyreview.com/2019/03/25/136460/deepfakes-shallowfakes-human-rights/, (accessed 27 April 2020).

Lewis, P. and Hilder, P., 'Leaked: Cambridge Analytica's blueprint for Trump victory', *The Guardian*, 25 March, 2018, https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory, (accessed 27 April 2020).

Metz, C., 'A Fake Zuckerberg Video Challenges Facebook's Rules', *The New York Times*, 11 June 2019, https://www.nytimes.com/2019/06/11/technology/fake-zuckerberg-video-facebook.html, (accessed 27 February 2020).

Metz, R. and CNN Business, 'The number of deepfake videos online is spiking. Most are porn', *CNN*, 7 October 2019, https://edition.cnn.com/2019/10/07/tech/deepfake-videos-increase/index.html, (accessed 10 May 2020).

Roose, K., 'Here Come the Fake Videos, Too', *The New York Times*, 4 March 2018, https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html, (accessed 27 February 2020).

Roose, K., 'The Making of a YouTube Radical', *The New York Times*, 8 June 2019, https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html, (accessed 27 February 2020).

Rosenberg, M., Confessore, N. and Cadwalladr, C., 'How Trump Consultants Exploited the Facebook Data of Millions', *The New York Times*, https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html, 17 March 2019, (accessed 7 May 2020).

Singer, N., 'What You Don't Know About How Facebook Uses Your Data', *The New York Times*, 11 April 2018, https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html, (accessed 5 May 2020).

Stewart, E., 'A fake viral video makes Nancy Pelosi look drunk. Facebook won't take it down.', *Vox*, 29 May 2019, https://www.vox.com/recode/2019/5/24/18638822/nancy-pelosi-doctored-video-drunk-facebook-trump, (accessed 27 April 2020).

Thomas, E., 'In the battle against deepfakes, AI is being pitted against AI', *Wired*, 25 November 2019, https://www.wired.co.uk/article/deepfakes-ai (accessed 20 May 2020).

Websites

Faceswap.dev, [website], 2019, https://forum.faceswap.dev/, (accessed 3 March 2020).

Faceswap.dev, [website], 2019, https://forum.faceswap.dev/viewtopic.php?f=6&t=146, (accessed 3 March 2020).

Faceswap.dev, [website], 2019, https://forum.faceswap.dev/app.php/faqpage, (accessed 3 March 2020).

Faceswap.dev, [website], 2019, https://forum.faceswap.dev/viewtopic.php?f=5&t=27, (accessed 3 March 2020).

GitHub, [website], https://github.com/deepfakes/faceswap, (accessed 3 March 2020).

GitHub, [website], https://github.com/deepfakes/faceswap/wiki, (accessed 3 March 2020).

In event of Moon Disaster, [website], 2019, https://moondisaster.org/, (accessed 28 February 2020).

Malavida, [website], 2019, https://www.malavida.com/en/soft/fakeapp/, (accessed 3 March 2020).

MrDeepFakes, [website], https://mrdeepfakes.com/celebrities, (accessed 15 March 2020).

Posters, B., *Bill Posters,* [website] http://billposters.ch/projects/big-dada/, (accessed 27 March 2020).

Reddit, [website], 2018, https://www.reddit.com/r/SubredditDrama/comments/7vy9cw/rdeepfakes_the_aigenerated_fake_celebrity_porn/, (accessed 4 March 2020).

u/GifFakes, 'Nicolas Cage in Man of Steel by /u/derpfakes', *Reddit*, [website], 6 February 2018, https://www.reddit.com/r/GifFakes/comments/7vntk0/nicolas_cage_in_man_of_steel_by_uderpfakes/, (accessed 12 April 2020).

Wikipedia, [website], https://en.wikipedia.org/wiki/Deepfake, (accessed 9 March 2020).

Press kit

MIT Center for Advanced Virtuality, *IEOMD-press-release-2019.pdf*, Retrieved: with the filmmaker's permission, Cambridge, Massachusetts, 22 November 2019, (accessed 11 April 2020).
In text: (MIT Center for Advanced Virtuality, 2019).

Videos

Baby Elon Musk Montage Deepfake, [online video], 8 May 2019, https://youtu.be/WHwQeetjLwk, (accessed 10 May 2020).

Film Club: 'This Video May Not Be Real' Can we trust what we see and hear?, [online video], 17 April 2020, https://www.nytimes.com/2020/04/17/learning/film-club-this-video-may-not-be-real.html, (accessed 9 May 2020).

Nick Cage DeepFakes Movie Compilation, [online video]. 2018, https://youtu.be/BU9YAHigNx8?t=39, (accessed 12 April 2020).

President Nixon speaking with astronauts Armstrong and Aldrin on the Moon, [online video], 30 March 2011, https://www.youtube.com/watch?v=1Ai_HCBDQIQ, (accessed 12 May 2020).

Podcast

'Breaking Bongo', *Radiolab*, [Podcast], WNYC Studios, 17 November 2019, https://www.wnycstudios.org/podcasts/radiolab/articles/breaking-bongo, (accessed 27 April 2020).

Report

Ajder, H., et al., 'The State of Deepfakes: Landscape, Threats, and Impact', *Deeptrace*, 2019, https://deeptracelabs.com/mapping-the-deepfake-landscape/, (accessed 05 March 2020).

Facebook, Inc., 'Facebook Reports Fourth Quarter and Full Year 2019 Results', *Menlo Park*, 2020, https://s21.q4cdn.com/399680738/files/doc_financials/2019/q4/FB-12.31.2019-Exhibit-99.1-r61_final.pdf, (accessed 27 April 2020).