Face off – a semiotic technology study of software for making deepfakes

Søren Vigild Poulsen[1]

Abstract. Deepfakes, an algorithm that transposes the face of one person onto the face of another person in images and film, is a digital technology that may fundamentally alter our belief in visual modality and thus presents alarming consequences for an image-centric culture. Not only are these face-translations now so advanced that it is virtually impossible for people to tell that they are fake – this technology is also becoming accessible to laypersons who, with little or no computer skills, can use them for all kinds of purposes, including criminal intentions like revenge porn and identity theft. It is therefore timely and crucial to explore the semiotic potential of deepfakes.

This paper presents a semiotic technology perspective, i.e., the study of technology for meaning- making that is an emergent field in social semiotics, to report on findings from an ongoing study of how deepfake software is designed and used as a semiotic resource in erotic and political contexts. The paper advances the argument that the software is able to appropriate all signifiers of the face and their cultural history. Consequently, the semiotic operations of this technology prepare the ground for the problematic perspectives of synthetic facial imagery.

On this basis, the paper calls for a critical awareness of taking visual representations of current events at face value and considers how deepfake technology is embedded in unsound sharing practices of visual artefacts that tamper with the rich meaning potential of the face.

Keywords: deepfakes; visual manipulation; semiotic software; face studies; digital culture

1. Introduction

The appearance of deepfakes, images and videos in which the face of one person is transposed onto the face of another person who says or does things that the former did not say or do, is fundamentally altering our perception of visual representation. Not only are these digital face-transplantations now so advanced

---

[1]   Department of Language and Communication, University of Southern Denmark, Denmark; e-mail: vigild@sdu.dk.

that it is virtually impossible for non-specialists to tell that they are fake, but software for making deepfake images is also becoming accessible to laypersons who, with little or no computer skills, can use them for all kinds of purposes, including criminal intentions like revenge porn and identity theft. According to a report by the Dutch research group Deeptrace[2], 96% of these kinds of synthetic media are pornographic and involve women's faces being "translated" and shared in both closed and public online fora and communities without their knowledge and consent. It is therefore timely and crucial to explore the semiotic potential of deepfakes critically. Research literature on this subject matter is emerging within research fields as diverse as machine learning (e.g. Maras, Alexandrou 2019), pornography studies (e.g. Popova 2020), law studies (e.g. Chesney, Citron 2019a) and media studies (e.g. Lees, Bashford-Roger, Keppel-Palmer 2021). However, each field tends to focus on either technical aspects or use of deepfakes. This paper complements the existing research with a semiotic technology perspective, i.e., the study of technology for making meaning, which is an emergent field in social semiotics (Djonov, Van Leeuwen 2018), to link technical and cultural dimensions of deepfakes. The paper reports on preliminary findings from a study of how deepfake software is designed and used as a semiotic resource in erotic and political contexts, respectively. On this basis, the paper argues that the issue of deepfakes arises from the software's appropriation of the rich cultural history of the face, and a semiotic dimension needs to be included in the discussions of the manipulations made possible by this technology in order to evaluate the effects of synthetic media.

2. Existing research of deepfakes in machine learning and law

Deepfakes have so far primarily been studied in machine learning (in computer science) and law. I will therefore briefly place the present study in the context of deepfake studies in these two research fields. While the term 'deepfake' has been dated to an internet forum in 2017[3] (Cole 2017; 2018), studies of deepfakes go back

---

[2]   Ajder, Henry; Patrini, Giorgio; Cavalli, Francesco; Cullen, Laurence 2019. The state of deepfakes: Landscape, threats, and impact. Deeptrace. 24 January. Accessed at https://sensity. ai/reports/ on 10 January 2021.

[3]   Cole, Samantha 2017. AI-assisted fake porn is here and we're all fucked. *Vice*. December 11. Accessed at https://www.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn on 10 January 2021;

Cole, Samantha 2018. We are truly fucked: Everyone is making AI-generated fake porn now. *Vice*. Accessed at https://www.vice.com/en_us/article/bjye8a/reddit-fake-pornapp-daisy-ridley on 10 January 2021.

longer within machine learning. The area of machine learning research gradually moved its research interest from extracting human faces from visual data sets (Blanze *et al.* 2004) to face swapping in still images (Bitoul *et al.* 2008) and later videos (Dale *et al.* 2011).

Deepfakes got their name from using deep learning technology, a concept that is developed in the field of machine learning. This technology runs a neural network simulation that uses big data sets to create digital information, such as (fake) images. Neural networks are designed to mirror how the human brain operates (at least in a computational understanding of the brain). The neural network can learn to recognize patterns in a data set and to solve classification problems automatically. Furthermore, it can be given a task to do and an objective to fulfill, while monitoring itself and modifying and improving its performance. Applied to face swapping, artificial intelligence (AI) algorithms analyse a source face (e.g. the actor Nicolas Cage) from different angles to learn what it looks like; then, they transfer these features they have learned onto a target face (e.g. that of Donald Trump[4]), as if these were a mask[5] (self-evidently, in this example the result would be that Trump looks like Cage). This process can be repeated multiple times, which will result in ever better facial simulation, depending on the quality of data input and the time frame of the project.

Research literature within computer science does not consider face swapping in a semiotic sense, but addresses problems of how to replace a face of one person with that of another, while the faces are shown from different angles, in different lighting conditions and positions, and how these analyses can be automated. The latest developments focus on computational ways to discover fakes (e.g. Farid 2018; Güera, Delp 2018; Maras, Alexandrou 2019).

Running parallel to this research, the analysis of deepfakes has been taken up in law in relation to privacy research. Aside from political destabilization and interfering with democratic processes, one can easily imagine the problems that deepfakes may cause in legal cases where visual documents function as empirical basis for lawsuits (Chesney, Citron 2019a). Also, deepfakes challenge traditional forensics methods and analysis (Burke 2006; Leone 2021). Research on deepfakes in law already recognizes

---

[4]   For some reasons, Nicholas Cage and Donald Trump are often used in examples of deepfakes both in research literature and on mainstream media.

[5]   It is important to note that machine learning *per se* is not a technology for face manipulation. Software like Faceswap applies machine learning methods. Deepfake software utilizes this technology that processes image data, enabling an algorithm to learn how to encode and decode image data in a particular computational operation, resulting in face replacement. The same method may in principle be used to swap any other body parts or objects in motion in videos.

the dangers of fake imagery to basic human rights; one issue is the use of deepfakes to discredit politicians and public figures; another issue is how to safeguard the fundamental rights, identities and safety of civilians against false images and videos. Detection software is being developed, and, as Silbey and Hartzog (2020: 961) point out, synthetic media can be used for positive intentions as well: "The potential upside of deepfakes is that they might help muster the political will to address the larger, structural problems made worse by the inability to trust what we see and hear". I sympathize with this viewpoint, but remain a bit pessimistic about the time perspective of any productive reactions and positive initiatives.

A semiotic technology perspective that I will present in the next section attempts to bridge the points of view introduced with respect to both technical and law-related aspects by exploring software for making synthetic media. Furthermore, this approach seeks to make a new contribution by focusing on the fact that it is due to the technological advances that deepfakes are widely circulated, both in terms of speed and reach of circulation. As such, the present study bears kinship with Virilio's (1986[1977]) concept of 'dromology' that he used to diagnose logics of speed in society.

3. Methodology

I want to apply a semiotic technology approach to study how deepfake software incorporates semiotic resources for making videos and images of face replacements. Semiotic technology research explores how all different types of technologies for meaning-making enable both the production and distribution of multimodal texts and artefacts as well as the performance of semiotic practices (e.g., Zhao *et al.* 2014; Poulsen 2018). This emergent research area stems from social semiotics (Halliday 1978; Hodge, Kress 1988), and has three overlapping foci: (1) to map semiotic resources in the technology of meaning-making; (2) to study how these are used in semiotic practices – and more broadly, social practices; and (3) to observe the histories, narratives and discourses of technologically mediated resources.

Using the semiotic technology approach, I will be analysing the functional features of the user interface as signifiers with meaning potential that are built into deepfake software programs. In relation to the present object of study, research questions to ask about these kinds of semiotic technologies are: what resources for altering face features are available in the software? How does this digital technology facilitate and administrate the making of synthetic faces? How is face swapping technology used in particular social settings?

To describe facial expressions and their meaning potentials, I include in the semiotic description of deepfake imagery multimodal interaction studies,

microsociology, film and media studies, as well as art and cultural history with special focus on the face. As Van Leeuwen (2005) points out, social semiotics (and thus semiotic technology) is not a matter of "pure" theory, it needs to incorporate other theories to provide an adequate framework for analysis of semiotic phenomena under observation. Thus, following the mapping of resources in the user interface of a deepfake technology, I will draw on work about faces from the above-mentioned research fields that until now have not been included in the discussion of deepfakes.

The following section presents a tentative description of the design and technical features of particular deepfake software in semiotic terms, that is, as semiotic resources and their meaning potentials for making deepfakes or synthetic imagery. This is followed by observations from two case studies on the usage of face swapping technology in an erotic context and a political context to illustrate how the resources built into the software could be used, and the problematic results that this usage contributes to.

Ideally, I would conduct a semiotic technology study of the design and use of a particular software program, such as Faceswap: I would study how this software program is designed as a semiotic device and the texts that could be made with its specific features. However, it was not possible to determine the origin of the software that was used for making the deepfake videos chosen as the objects of this semiotic analysis. Thus, the case studies only serve to illustrate general points about the problematic nature of such technology.

4. Analysis of the design of software for making deepfakes

The following analysis includes a description of the design of deepfake software as semiotic technology. The subsequent sections will focus on: (1) the process and steps of making a deepfake video; (2) the context of selecting deepfake software; (3) the design of Faceswap's user interface; (4) how the face replacement process is structured in Faceswap; and (5) the semiotic operations of the software.

4.1. The process and steps of making a deepfake video

If one breaks face replacement down into its basic elements, a deepfake production consists of four steps, regardless of the particular software one could use in the process. As Sample (2020)[6] explains:

---

[6]   Sample, Ian 2020. What are deepfakes – and how can you spot them? *The Guardian* 13 January 2020,   was accessed at https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them on 10 January 2021.

> First, you run thousands of face shots of the two people through an AI algorithm called an encoder. The encoder finds and learns similarities between the two faces, and reduces them to their shared common features, compressing the images in the process. A second AI algorithm called a decoder is then taught to recover the faces from the compressed images. Because the faces are different, you train one decoder to recover the first person's face, and another decoder to recover the second person's face. To perform the face swap, you simply feed encoded images into the "wrong" decoder. For example, a compressed image of person A's face is fed into the decoder trained on person B. The decoder then reconstructs the face of person B with the expressions and orientation of face A. For a convincing video, this has to be done on every frame.

From a semiotic technology perspective, what is interesting when analysing deepfake software is how the digital technology structures and facilitates the deepfake production process to its users who are not necessarily technically skilled. I will therefore be discussing the role and function of the software that mediates between computational operations and visual meaning-making.

4.2. Context of selecting deepfake software

Previously, making deepfakes required advanced programming, but free or open-source software that can perform the computational operations in a few simple steps is now widely available to the general public. The following observations were made during the research phase for software that will enable the creation of deepfakes:

(1) A Google search presents several options for downloading software for desk PCs as well as recommendations indicating accessibility. There is also easy access to apps for mobile phones that allows you to make short fake videos, images and audio.
(2) Serious tech media evaluate the best deepfake software on the market.
(3) Many of the software programs are marketed for laypersons: users are not required to have computer skills or special knowledge about deep learning or computer vision.
(4) The online Gethub repositories hold several communities for deepfakes.
(5) Some software programs (e.g. Machin Tube) come with pre-loaded templates of faces to make deepfakes of, e.g. politicians, singers and celebrities (like Donald Trump).
(6) While this study focuses on the visual features, I also wish to mention free technology for deepfake audios, e.g. Lyrebird AI, which has the capacity to form sentences just upon hearing a few spoken words, thus creating a digital voice. This software also enables overdubbing that "allows you to replace

recorded words and phrases with synthesized speech that is tonally blended with the surrounding audio"[7]. So, while face replacement is normally what is associated with deepfakes, the technology also includes audio, in which case only the voice of someone else is dubbed. These deepfakes are called 'voice cloning' and an example of this is a video from 2019 in which the voice of Barack Obama was dubbed by actor Jordan Peale.

In the following section, I will analyse the user interface of Faceswap. Faceswap[8] (2019-) is the leading open source face replacement technology developed by Matt Tora, Bryan Lyon, and Kyle Vrooman. It is coded in the programming language Python. The software is powered by Artificial Intelligence using so-called Convolutional Neural Networks, and it has an active forum of volunteers and users. I selected Faceswap over e.g. DeepFakeLab, which is more technical, since the latter was developed for computer science students and uses command lines. Faceswap features a graphic user interface for laypersons, and its tutorials describe the process in a great detail, avoiding technical terms.

4.3. The user interface of Faceswap

The graphic user interface of Faceswap presents four headings (Extract, Train, Convert, Tools), which suggests a workflow in the making of deepfake, starting from left to right. In contrast to, e.g. DeepfakeLap which offers the users multiple command line operations and sub functions, the process of making deepfakes is simplified (three steps) and streamlined (unless the user wants to include specific effects in the video, they will go with the default settings under each heading).

Regardless of the heading, the user interface is divided into two windows. In the left-hand frame window different features relevant to the chosen heading are provided, e.g. under Extract, file trajectories can be selected of the videos that are made into fake images or videos. In the right-hand frame, information about and results of the current stage of the process are displayed, e.g. the number of frames extracted from the videos, or a preview of the results of the process.

4.4. The process of making deepfakes in Faceswap

As mentioned above, Faceswap structures the making of deepfakes in a few simple steps. That said, the software offers the user different editing options in Faceswap user interface, depending on the purpose of the training and type of data. Faceswap

---

[7]   Descript website. Accessed at https://www.descript.com/lyrebird-ai on 16 December 2019.
[8]   The official website of Facewap is https://faceswap.dev.

mirrors the processes that were presented in Section 4.1. In the software tutorial[9], these steps are described as follows:

> Gather data: Find videos/images of the people whose faces you want to swap
> Extract: Get all faces from data, Sort, Remove incorrect faces
> Training: Use the AI to learn how to swap faces (i.e., encoding and decoding)
> Convert: Apply the AI on data (either new or already trained) to swap faces in the
>   videos

This organization of headings and the tools under each heading would, in semiotic terms, be the overall 'semiotic regime' (Van Leeuwen 2005), i.e. a social regulation of the use of semiotic resources and meaning-making processes. The design of the user interface enforces a particular way of producing face replacement videos.

Already at this stage of the analysis, the observation can be made that, on the one hand, the simplified workflow enables laypersons to make deepfakes, and, on the other hand, it has consequences for the users' understanding of what the software does. One such effect is blackboxing, which hides the functions performed by the technology. This is to some degree the case with Faceswap, compared to DeepFaceLap that lets the user follow the computational processes in more detail. Still, the blackboxing process does not run as straightforwardly as e.g. Reface, an app that allows the user to take a selfie and embed it in a five-to-ten-second Hollywood movie or music video. With this app, the face-swapping process is totally hidden, and the user is only presented with a simple step-by-step process: after taking a selfie, the image is blurred and the mobile screen displays stars tingling while the underlying AI analyses the visual information, animating the data processing as a magical process. Next, the user simply chooses between the selected movies in which (s)he wants his/her face to be inserted. This video is then generated by showing a timeline, and the final results to be saved or shared. Blackboxing seems to be a side effect of making advanced software accessible to the general public. After the deepfake video has been made, the Faceswap software offers ways to save and share the content on e.g. social media networks. Thus, the software prompts the user to distribute the deepfake imagery and participate in the circulation of the synthetic media content. This is the basic setup.

4.5. Technical features as semiotic resources

A semiotic technology study of Faceswap would describe semiotic resources of the software, i.e. the technicaly mediated signifiers and their meaning potentials. However, it is difficult to describe the deepfake software and videos separately,

---

9   Facewap 2019 was accessed at https://forum.faceswap.dev/index.php on 10 January 2021.

as the meaning potential of the user interface is dependent on the videos that are being made into deepfakes. Therefore, a full analysis would break down the semiotic features of a deepfake video into different levels: the faces (of two persons); the environment that the faces are placed in; and video editing software. Each of these layers contains several semiotic systems that contribute to the overall meaning-making of a deepfake video. In the following I will concentrate on the editing software, but include aspects from the other levels. If one follows the workflow of Faceswap, each step utilizes several semiotic resources.

4.5.1. Gathering data (representing faces)

A prerequisite of face swapping is the videos or images that function as input to the computational process. The function of uploading face images incorporates a semiotic resource that is essential to all visual communication, namely 'existential process', i.e. the act of representing something as 'being there' (Boeriis 2009). Thus, simply to be able to represent images of a source and a target face becomes the grounds for making any deepfake, and the subsequent editing processes depend on the usage of this resource. That the point of departure is an existing facial representation may seem self-evident, but it is a fundamental feature for this very reason. Highlighting the existential resource for representation also draws attention to another feature, namely that the input faces and the instantiated resources that they manifest are included in the meaning-making process that the semiotic software creates. The videos are in themselves multimodal artefacts that draw on multiple semiotic systems, and since each video typically contains extensive footage of face angles, expressions and movements, they display a complex ensemble of instantiated meanings. The resources of the face images exist independently of the deepfake technology. Nonetheless, the deepfake technology depends on the textual meaning of the faces, for such software functions as a parasite on the meaning made in the videos; the software re-semiotizes (Iedema 2001) or transduces (Kress 2010) the meaning potentials of the two videos by emerging two faces into one video.

4.5.2. Extraction (analytical process and reconfiguration of facial resources)

The next step is to extract face data from the still images. As Faceswap[10] (2019) describes in their tutorial on face replacement:

> At the highest level, extraction consists of three phases: detection, alignment and mask generation:

---

[10] Facewap 2019. [Guide] Extraction – a workflow. Accessed at https://forum.faceswap.dev/viewtopic.php?t=27z on 10 January 2021.

> Detection – The process of finding faces within a frame. The detector scans an image and selects areas of the image that it thinks are faces.
> Alignment – Finding the "landmarks" […] within a face and orienting the faces consistently. […]
> Mask Generation – Identify the parts of an aligned face that contains face and block out those areas that contain background/obstructions.
>
> […] Extracting serves 2 main purposes:
> Training: To generate a set of faces for training. These faces will also contain the alignment information and masks which are required for training the model. […]
> Converting: To generate an alignments file and mask for converting your final frames. The alignment file contains information about where each face is in each frame so that the conversion process knows where to swap faces for any given frame.

The purpose of the extraction process is to reduce the face information in an image to its core elements for an algorithm to reproduce this face with the least amount of data. The reduction process may be similar to 'abstraction' in Peirce's diagrammatical reasoning (see CP; Peirce 1976). The mentioned 'landmarks of the face' refer to key data points of a face that indicates facial position and expression in each frame (see Fig. 1). These operations are the core of the semiotic transformations in the initial face placement process.

Essentially, two semiotic ideational resources are utilized in the process: first, categorization and segmentation of facial elements incorporate an 'analytical process', i.e. a whole-parts relation (Kress, Van Leeuwen 2006); second, I understand the alignment process that tracks facial movement based on a number of landmarks of the face as a resource for reconfiguration of semiotic multiple different semiotic systems. Historically, studies of facial resources and their meaning potentials have divided features of a face into different semiotic systems or modes, e.g. the gaze (Argyle, Dean 1965; Argyle, Cook 1976; Goffman 1963; Belting 2013; Exline, Fehr 1982), head movement (Altorfer *et al.* 1992, 1998; Jossen *et al.* 2000), gesture (Kendon 1967; 1978; Schegloff, Sacks 1973; Goodwin 1986; Norris 2004), speech (Halliday, Hasan 1976). These semiotic systems are heuristic analytical categories, as facial resources co-create meaning multimodally. I would claim that this alignment is the essential part of the editing software, since it lays the ground for training algorithms to make deepfakes. At the same time, this exact operation creates the issues of synthetic media. Semiotically speaking, alignment as a resource of reconfiguration includes two elements: first, by aligning a face's diverse signifiers, the software appropriates multiple different semiotic systems that people use to make facial meanings. The software reconfigures

semiotic features of the face that carries an extremely rich cultural history, i.e., the meaning potential that people throughout history have created with each facial semiotic system (Boehm *et al.* 2015). Second, by incorporating all of these different semiotic systems into landmarks that make up a map of the face, the software offers previously unseen ways to manipulate the totality of these facial semiotic systems and their meaning potentials all at once.
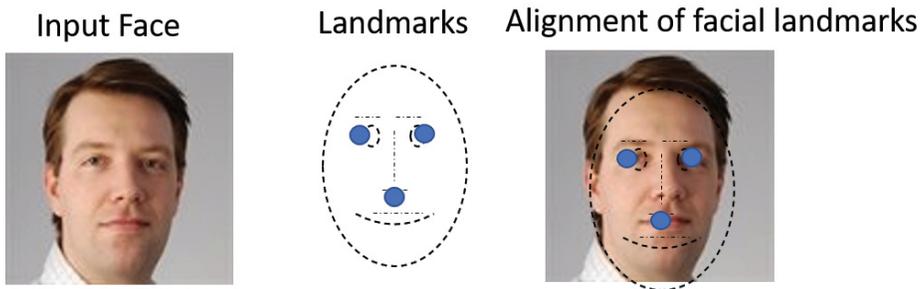

Figure 1. An example of landmarking in the process of facial alignment. Author's own production inspired by a figure in Wang *et al.*[11]

4.5.3. Training (semiotic principles for articulation and interpretation)

I consider the last two steps in the Faceswap workflow as having to do with overall principles for meaning-making. The Training function concerns the establishment of a strategy for semiosis, so the source face is interpreted as the target face. This strategy is drawn upon in the making of a face swap, as far as it can be automated. As such, it becomes a resource for interpretation and re-articulation (Kress, Van Leeuwen 2001) of a semiotic production.

4.5.4. Conversion (resource integration/blending)

Faceswap differs from face alterations or improvements in face apps like TikTok, Snapchat or Instagram which apply filters to face images because the face attributes from the target face in a deepfake are also transferred to the source face. Therefore, the meaning-making principle is not metaphorically from a source (face) to a target (face) (Lakoff, Johnson 1980), the target face does not get semanticized by the features and qualities of the source. In social semiotic terms, this process constitutes

---

[11] Wang, Junjue; Amos, Brandon; Das, Anupam; Pillai, Padmanabhan; Sadeh, Norman; Satyanarayanan, Mahadev 2017. A scalable and privacy-aware IoT service for live video analytics. *Proceedings of the 8th ACM on Multimedia Systems Conference*, 38–49; accessed on 20 January 2021.

an 'integration resource' (Bauldry, Thibault 2006), i.e. a principle for how semiotic resources may be combined in multimodal ensembles. One can describe this as a blend (Fauconnier, Turner 2002) that merges different semantic domains into an emergent meaning construction. In this analytical framework, a deepfake video would be conceptualized as a meaning construction that blends face features of a source with face expressions of a target. As a result, the deepfake video can show a person doing things that are fictional (i.e. not authentic), yet trustworthy.

5. Analysis of two contexts of deepfake usage: pornography and politics

In this section, I explore how the semiotic deepfake technology is used in diverse semiotic practices where fake faces may serve harmful purposes and help maintain unhealthy social practices. I consider two types of contexts in which deepfakes are being used. Each context analysis reports on preliminary observations.

5.1. Deepfakes in pornography

A simple Google search for deepfakes reveals that deepfakes appear as a category on major porn sites (e.g. Pornhub), as well as on sites dedicated to this kind of pornographic content (e.g. MrDeepFake). Officially, sites like Pornhub have banned deepfakes from their site, but videos can still be found if searched for on these sites[12]. I did a search on 'xvideo.com' and found multiple videos of Hollywood actresses, female singers and politicians. In the following, I present general observations from an analysis of a deepfake video of a famous American singer. Given the explicit and transgressive nature of the imagery, I will refrain from describing details in the video and only focus on general observations[13].

I observed that the face of the singer is shown in long camera shots, and that the face is displayed at the centre of the footage, while the 'body', i.e. the body of the porn actress, performs sexual acts, so that the viewer can clearly see and dwell on the fake face features. As such, the face presents itself in its completeness, which Balázs (1970[1923]), in his studies of silent films, has described as a key element

---

[12]   Ajder, Henry; Patrini, Giorgio; Cavalli, Francesco; Cullen, Laurence 2019. The state of deepfakes: Landscape, threats, and impact. Deeptrace. 24 January. Accessed at https://sensity. ai/reports/ on 10 January 2021.

[13]   I am aware of the problematic nature of analysing erotic deepfakes, since I as researcher engage in the unsound social practice that I criticize, and as such, I take part in upholding the market for these products. I have chosen not to disclose the name of the woman in the video analysed. I am also aware that the viewing of such material potentially poses a criminal act.

of the close-up. The whole video is filmed from a subjective point of view, where the woman looks directly and up close at the camera. The use of the gaze is an essential resource that intensifies the direct visual addressing of the viewer (Kress, Van Leeuwen 2006). The close-up, as Doane (2003) argues, "[…] embodies the pure fact of presentation, […] of showing – a "here it is"" (quoted in Jerslev 2017: 91.) that creates a sense of direct access to reality. This suggests the fascination of the male gaze with the face as an object; I also suspect that the depicted face and gaze of the singer bespeaks of a video producer who takes pride in the deepfake production. The face shot in these videos is "undisturbed", as it were. In this regard, the videos that swap the faces of a woman celebrity onto the body of a porn actress seem to deviate from other pornographic videos where the face of the woman is "manhandled" as it is not uncommon for the (female) face in porn to be slapped, stepped on, penetrated, or covered by semen. Perhaps the facial display in the deepfake porn is also a question of limitations to the deepfake technology, which is not able to handle changes in face shots that include other elements. There are small glitches in the videos where it becomes visible that the face of the actress has been manipulated: the pixels that mark up the face area become visible and, as a result, presumably break the illusion of the depicted fantasy.

In this context, I suggest that the use of deepfake software for making fakes videos of celebrities, and public figures advances the already existing genres of 'fake nudes' and 'celeb nudes/fakes' that include videos of lookalikes. Also, these videos are fed to a market for intimate videos produced without the consent of the depicted women whose faces are placed in the context of nude scenes. Only in a few cases, e.g. nude footages of celebrities, has the "leaking" of sex tapes been used in personal branding as a deliberate act of empowerment. I would claim that nude deepfakes primarily relate to hate porn and revenge porn that have devastating consequences for the victims without whose consent pictures or videos of them are exposed to the public at large. As pointed out by Popova (2020), the use of deepfake nudes may at first glance appear as a fantasy of having intercourse with a celebrity, yet the phenomenon is not like fan fiction, but is part of a transgressive act, and, more broadly, of a misogynistic culture centring on fantasies of violating acts by a man on a particular woman's face and body. Furthemore, while there may be a (legal) difference between a person making such manipulated imagery and viewers who watch the videos, but who themselves would not create this kind of content, the viewers of these videos, in their engagement with deepfake porn, nevertheless uphold this social practice. By viewing, commenting and further sharing of fake images and videos, they legitimize the existence of such deepfakes. As such, the production and distribution of deepfake porn supports a sharing economy using women's sexuality as a commodity.

5.2. Deepfakes in politics

Turning to politics, and looking beyond examples of political satire, one can easily find several examples of deepfakes of political figures that supposedly say or do something that would discredit their reputation, and this is a global issue. A recent instance was a deepfake video of the Indian politician Manoj Tiwari that was shared between groups on WhatsApp during the 2020 Delhi elections. In this regard, I made the preliminary observation that Tiwari was shown talking to voters in Haryanvi and English, but in the real video he speaks in Hindi. However, unlike the erotic deepfake that centres on a fabricated sexual act, this political deepfake video was not manipulated beyond a synchronization of the voice with the depicted person's limb movements. From watching this video, and others, it seems that it is especially linguistic resources that can be cloned with a deepfake voice technology, so they are used to make fake statements. The false account was accompanied by the presentation of a particular politician's face. In this context, the identity of a political figure is closely associated with his or her face, as it also appears on political posters. Like deepfake porn, the video of Tiwari utilizes the fundamental ideational resource of representing a person's face. By combining linguistic and gaze resources, politicians like Tiwari may be portrayed as making outrageous or discriminating utterances that would damage the support from their voters if these were true. As such, the purpose of political deepfake videos and audios is an attempt to destabilize political systems and to shatter people's trust in democratic processes. The very act of making a person say something that he or she has never uttered demonstrates that the use of deepfake technology in manufacturing false images and videos in a political context could cause double harm by contributing to the increasing disgust with politicians as well as disbelief in trusted news sources that play a crucial role in democratic processes (Paterson, Hanley 2020).

That false political videos pose a serious threat to democracy was indicated by a CIA report on worldwide threat assessment published prior to the recent US presidential election. The report states that "adversaries and strategic competitors probably will attempt to use deepfakes or similar machine-learning technologies to create convincing – but false – image, audio, and video files to augment influence campaigns directed against the United States and our allies and partners" (Coates 2019: 7).[14] It is not only political institutions that are affected by such problematic use of deepfakes. The use of deepfake technology in manufacturing false images

---

[14]   Coats, Daniel R. 2019, Worldwide threat assessment of the US intelligence community https://www.odni.gov/index.php/newsroom/congressional-testimonies/item/1947-statement-for-the-record-worldwide-threat-assessment-of-the-us-intelligence-community; accessed on 20 January 2021. Quotation refers to page 7.

and videos in a political context. This affects not only political institutions: deepfakes can also damage the reputation of businesses and organizations, which may have a negative impact on stock prices and bring along long-term consequences when customers and shareholders no longer trust enterprises and markets. The fact that deepfakes have become a phenomenon to be aware of threatens to undermine the perceived objectivity of videos featuring politicians and public figures. It should be a trademark of a healthy political climate to dispute and critically discuss decision-makers' solutions to social problems, and it is also paramount to be able to hold politicians accountable for their actions and claims. In this context, deepfakes align with fake news and distribution of misinformation (Chesney, Citron 2019b).

6. Concluding remarks

This article is a report on a preliminary study of the design and use of semiotic technology for making deepfakes. On this basis, I have argued that a semiotic approach is needed to describe why the design and use of face replacement technology may potentially cause disturbance, since the manipulation of the face tampers with semiotic resources that carry an immensely rich meaning potential. By analysing the leading face swapping software, Faceswap, as a semiotic technology, I have tried to demonstrate how the software appropriates many semiotic resources of a face on video and how it is able to manipulate these. Still, much more research on the software and its usage in different contexts needs to be carried out in this area.

When asked about their responsibility for spreading deepfakes at a tech conference in 2019[15], two of the three founders of Faceswap answered that it is not their duty to prevent this from happening. If they do not develop the technology, someone else will come along and do this. Furthermore, they argue that a better countermovement to the rise of deepfakes is to familiarize the general public with this kind of software; likewise, alongside the development of deepfake software, scientists and companies are working on software for detecting deepfakes (e.g. Nguyen *et al.* 2019). While this viewpoint may be valid, the article would claim that a joint effort by tech companies, political leaders and decision-makers, a more critical press and research is a better solution to some of the problems discussed in this article. Also, it is important to support the digital literacy of the general public, so that they do not uncritically believe in images and videos circulated on

---

[15] Tora, Matt; Lyon, Bryan; Vrooman, Kyle 2019 Faceswap 2019: Faceswap: Developing a platform for practical Deepfakes. Recorded presentation held at *Devoxx* conference, 7 November Belgium was accessed at https://devoxx.be/talk/?id=103201on 20 January 2021.

the internet, especially on social media. In addition to these initiatives, semiotic technology descriptions of deepfake software can help us understand how synthetic media is made and how society is affected when people's faces and the values and meanings we assign to them are fabricated. The face is the most valuable part of who we are and who we believe others to be.

References

Argyle, Michael; Cook, Mark 1976. *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.

Argyle, Michael; Dean, Janet 1965. Eye-contact, distance and affiliation. *Sociometry* 28(3): 289–304. https://doi.org/10.2307/2786027

Altorfer, Andreas; Goldstein, Michael J.; Miklowitz, David J.; Nuechterlein, Keith H. 1992. Stress-indicative patterns of non-verbal behaviour: Their role in family inter-action. *British Journal of Psychiatry* 161(18): 103–113. https://doi.org/10.1192/S0007125000297122

Altorfer, Andreas; Käsermann, Marie-Louise; Hirsbrunner, Hanspeter 1998. Arousal and communication: The relationship between nonverbal, behavioral, and physiological indices of the stress response. *Journal of Psychophysiology* 12: 40–59.

Balázs, Béla 1970[1923]. The close-up/The face of man. In: Balázs, Bela (ed.), *Theory of the Film: Character and Growth of a New Art*. New York: Dover, 52–89.

Baldry, Anthony; Thibault, Paul 2006. *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook*. London/Oakville: Equinox.

Belting, Hans 2013. *Faces: Eine Geschichte des Gesichts*. Munich: C. H. Beck.

Bitouk, Dmitri; Kumar, Neerah; Dhillon, Samreen; Belhumeur, Peter; Nayar, Shree K. 2008. Face swapping: Automatically replacing faces in photographs. ACM *Transactions on Graphics* 27(3): 1–8. https://doi.org/10.1145/1360612.1360638

Blanz, Volker; Scherbaum, Kristina; Vetter, Thomas; Seidel, Hans-Peter 2004. Exchanging faces in images. *Computer Graphics Forum* 23: 669–676. https://doi.org/10.1111/j.1467-8659.2004.00799.x

Boehm, Gottfried; Budelacci, Orlando; Di Monte, Maria Giuseppina; Renner, Michael (eds.) 2015. *Gesicht und Identität/Face and Identity*. Paderborn: Wilhelm Fink Verlag. https://doi.org/10.30965/9783846758151

Boeriis, Morten 2009. *Multimodal Socialsemiotik og Levende Billeder*. [*Multimodal social semiotics and moving images*. Unpublished PhD thesis.] Odense: University of Southern Denmark.

Burke, Peter 2006. *Eyewitnessing the Uses of Images as Historical Evidence*. London: Reaktion Books.

Chesney, Robert; Citron, Danielle K. 2019a. Deepfakes: A looming challenge for privacy, democracy, and national security. *California Law Review* 107(6): 1753–1820.

Chesney, Robert; Citron, Danielle K. 2019b. Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs* 98(1): 147–155.

CP = Peirce, Charles S. 1931–1958. *Collected Papers of Charles Sanders Peirce.* Cambridge: Harvard University Press. (Vols. 1–6, Hartshorne, Charles; Weiss, Paul, eds., 1931–1935; vols. 7–8, Burks, Arthur W., ed., 1958.) [In-text references are to CP, followed by volume and paragraph numbers.]

Dale, Kevin; Sunkavalli, Kalyan; Johnson, Micah K.; Vlasic, Daniel; Matusik, Wojceich; Pfister, Hanspeter 2011. Video face replacement. *ACM Transactions on Graphics* (SIGGRAPH), 30. https://doi.org/10.1145/2070781.2024164

Djonov, Emilia; Van Leeuwen, Theo 2018. The power of semiotic software: A critical multimodal perspective. In: Flowerdew, John; Richardson, John E. (eds.), *The Routledge Handbook of Critical Discourse Studies.* New York: Taylor & Francis, 566–581. https://doi.org/10.4324/9781315739342-39

Doane, Mary Ann 2003. The close-up: Scale and detail in the cinema. *Differences: A Journal of Feminist Cultural Studies* 14(3): 89–111. https://doi.org/10.1215/10407391-14-3-89

Exline, Ralph. V.; Fehr, Barbara J. 1982. The assessment of gaze and mutual gaze. In: Scherer, Klaus R.; Ekman, Paul (eds.), *Handbook of Methods in Nonverbal Behavior Research*. Cambridge: Cambridge University Press, 91–135.

Farid, Hany 2018. Digital forensics in a post-truth age. *Forensic Science International* 289: 268–269.

Fauconnier, Gilles; Turner, Mark 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Goffman, Erving 1963. *Behavior in Public Places*. New York: Free Press.

Goodwin, Chuck 1986. Gestures as a resource for the organization of mutual interaction. *Semiotica* 62(1/2): 29–49. https://doi.org/10.1515/semi.1986.62.1-2.29

Güera, David; Delp, Edward J. 2018. Deepfake video detection using recurrent neural networks. *15th IEEE International Conference on Advanced Video and Signal Based Surveillance* (AVSS), 1–6. https://doi.org/10.1109/AVSS.2018.8639163

Halliday, Michael A. K. 1978. *Language as Social Semiotic*. London: Edward Arnold.

Halliday, Michael A. K; Hasan, Ruqaiya 1976. *Cohesion in English*. London: Longman.

Hodge, Robert; Kress, Gunther 1988. *Social Semiotics*. Cambridge: Polity Press.

Iedema, Rick 2001. Resemiotization. *Semiotica* 137(1/4): 23–39. https://doi.org/10.1515/semi.2001.106

Jerslev, Anne 2017. SKAMs 'lige her' og 'lige nu': Om SKAM og nærvær. ['Right here' and 'right now' in SKAM: On SKAM and presence.] *Nordisk Tidsskrift for Informationsvidenskab Og Kulturformidling* 6(2): 75–81. https://doi.org/10.7146/ntik.v6i2.99080

Jossen, Stefan; Käsermann, Marie-Louise; Altorfer, Andreas; Foppa, Klaus; Zimmermann, Henrich; Hirsbrunner, Hans-Peter 2000. The study of emotional processes in communication: Peripheral blood flow as an indicator of emotionalization. *Behavior Research Methods, Instruments, & Computers* 32: 47–55. https://doi.org/10.3758/BF03200787

Kendon, Adam 1967. Some functions of gaze-direction in social interction. *Acta Psychologica* 26: 22–63. https://doi.org/10.1016/0001-6918(67)90005-4

Kendon, Adam 1978. Looking in conversation and the regulation of turns at talk: A comment on the papers of G. Beattie and D. R. Rutter et al. *British Journal of Social and Clinical Psychology* 17: 23–24. https://doi.org/10.1111/j.2044-8260.1978.tb00891.x

Korshunova, Iryna; Shi, Wenzhe; Dambre, Joni.; Theis, Lucas 2017. Fast face-swap using convolutional neural networks. *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice, 3697–3705. https://doi.org/10.1109/ICCV.2017.397

Kress, Gunther 2010. *Multimodality – A Social Semiotic Approach to Contemporary Communication*. London, New York: Routledge.

Kress, Gunther; Van Leeuwen, Theo 2006. *Reading Images – The Grammar of Visual Design*. (2nd ed.). London: Routledge. https://doi.org/10.4324/9780203619728

Lakoff, George; Johnson, Mark 1980. *Metaphors We Live by*. Chicago, London: University of Chicago Press.

Lees, Dominic; Bashford-Roger, Tom; Keppel-Palmer, Marcus 2021. The digital resurrection of Margaret Thatcher: Creative, technological and legal dilemmas in the use of deepfakes in screen drama. *Convergence: The International Journal of Research into New Media Technologies* 27(4): 954–973. https://doi.org/10.1177/13548565211030452

Leone, Massimo 2021. From fingers to faces: Visual semiotics and digital forensics. *Int. Journal of the Semiotics of Law* 34: 579–599. https://doi.org/10.1007/s11196-020-09766-x

Maras, Marie-Helen; Alexandrou, Alex 2019. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. *The International Journal of Evidence* 23(3): 255–262. https://doi.org/10.1177/1365712718807226

Nguyen, Thanh. T.; Nguyen, Cuong M.; Nguyen, Dung T.; Nguyen, Duc T.; Nahavandi, Saeid 2019. Deep learning for deepfakes creation and detection: A survey. *arXiv* preprint *arXiv:* 1909.11573: 1–16.

Norris, Sigrid 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. New York, London: Routledge. https://doi.org/10.4324/9780203379493

Paterson, Thomas; Hanley, Lauren 2020. Political warfare in the digital age: Cyber subversion, information operations and 'deepfakes'. *Australian Journal of International Affairs* 74(4): 439–454. https://doi.org/10.1080/10357718.2020.1734772

Peirce, Charles S. 1976. *The New Elements of Mathematics by C. S. Peirce* (Eisele, Carolyn, ed.). The Hague: Walter De Gruyter.

Popova, Milena 2020. Reading out of context: Pornographic deepfakes, celebrity and intimacy. *Porn Studies* 7(4): 367–381. https://doi.org/10.1080/23268743.2019.1675090

Poulsen, Søren Vigild 2018. Becoming a semiotic technology – a historical study of Instagram's tools for making and sharing photos and videos. *Internet Histories* 2(1/2): 121–139. https://doi.org/10.1080/24701475.2018.1459350

Schegloff, Emanuel; Sacks, Harvey 1973. Opening up closings. *Semiotica* 8: 289–327. https://doi.org/10.1515/semi.1973.8.4.289

Silbey, Jessica; Hartzog, Woodrow 2019 The upside of deepfakes. *Maryland Law Review* 78(4): 960–966.

Stöckl, Hartmut; Caple, Helen; Pflaeging, Jana (eds.) 2019. *Shifts towards Image-Centricity in Contemporary Multimodal Practices*. London: Routledge. https://doi.org/10.4324/9780429487965

Van Leeuwen, Theo 2005. *Introducing Social Semiotics*. London: Routledge. https://doi.org/10.4324/9780203647028

Virilio, Paul 1986[1977] *Speed and Politics: An Essay on Dromology*. New York: Semiotext(e).
Zhao, Sumin; Djonov, Emilia; Van Leeuwen, Theo 2014. Semiotic technology and practice: A multimodal social semiotic approach to PowerPoint. *Text & Talk* 34(3): 349–375. https://doi.org/10.1515/text-2014-0005

Face off – семиотическое технологическое исследование программного обеспечения для создания подделок

Deepfakes, алгоритм, который накладывает лицо одного человека на лицо другого в изображениях и видео – цифровая технология, которая может коренным образом изменить нашу веру в визуальную модальность и представляет серьезную опасность для нашей культуры, ориентированной на изображения. Эти технологии не просто настолько продвинуты, что практически не позволяют усомниться в их подлинности, но также всё более доступы непрофессионалам, которые могут использовать их в самых разных целях, включая преступные намерения вроде порнографии в целях мести и кражи личных данных. В связи с этим исследование семиотического потенциала таких подделок является своевременным и крайне необходимым. В статье представлена семиотическая технологическая перспектива – новое направление социосемиотики, анализирующее, каким образом программное обеспечение *deepfake* разрабатывается и используется в качестве семиотического ресурса в эротических и политических контекстах. Выдвигается предположение, что такое программное обеспечение способно подчинить все знаки лица и их культурную историю. Следовательно, семиотические операции этой технологии подготавливают почву для синтетического изображения лица.

Face off – süvavõltsingute tegemiseks kasutatava tarkvara semiootiline tehnoloogiauuring

Süvavõltsing, algoritm, mis asetab kujutistel ja filmides ühe isiku näole teise isiku näo, on digitehnoloogia, mis võib fundamentaalselt muuta meie usku visuaalsusmodaalsusesse ja millega kuvandikesket kultuuri silmas pidades seega kaasnevad ohtlikud tagajärjed. Sellised näotõlked pole mitte ainult niivõrd edasijõudnud tasemel, et inimestel on praktiliselt võimatu otsustada, kas tegemist on võltsinguga, vaid see tehnoloogia on muutumas kättesaadavaks mitteerialainimestele, kes vähestest või puuduvatest arvutioskustest hoolimata võivad neid kasutada igasugustel eesmärkidel, sealhulgas kuritegelikel nagu seda on kättemaksuporno ja identiteedivargus. Seega on süvavõltsingute semiootilise potentsiaali vaatlemine õigeaegne ning oluline.

Artiklis pakutakse välja semiootilise tehnoloogia vaatenurk, s.t sotsiosemiootikas valdkonnana esile kerkiva tähendusloometehnoloogia vaatlus, esitamaks käimasoleva uuringu tulemusi selle kohta, kuidas on disainitud süvavõltsingutarkvara ning kuidas seda erootilistes ja poliitilistes kontekstides semiootilise ressursina kasutatakse. Artikkel väidab, et selline tarkvara on suuteline anastama kõik tähistajad ning nende kultuuriloo. Järelikult valmistavad selle tehnoloogia semiootilised operatsioonid ette pinnast sünteetiliste näokujutuste problemaatilistele perspektiividele.

Sellest lähtuvalt kutsutakse artiklis üles kriitilisusele päevasündmuste visuaalsete representatsioonide automaatselt tõe pähe võtmise suhtes ning tuletatakse meelde, et tehnoloogia on kaasatud visuaalsete artefaktide ebasoovitavatesse jagamispraktikatesse, mis kahjustavad näo rikkalikku tähenduspotentsiaali.