İSTANBUL UNIVERSITY PRESS

Acta Infologica,
ACIN 2025, 9 (1): 208–222

https://doi.org/10.26650/acin.1673277
Submitted: 10.04.2025
Revision Requested: 28.04.2025
Last Revision Received: 06.05.2025
Accepted: 03.06.2025
Published Online 04.06.2025

Acta Infologica

Research Article

Open Access

Investigation of User Comments on Videos Generated by Deepfake Technology

Sertaç Kaya [1]

[1] İstanbul Aydin University, Faculty of Communication, Department of Journalism, İstanbul, Türkiye

Abstract The ease of content production with developing technologies has caused deepfake-based videos to gain popularity and prevalence on social media platforms. Deepfake technology can be used for both manipulation and entertainment purposes because it can imitate people's voices and faces realistically. This study aims to evaluate user perceptions of content produced using deepfake technology. For this purpose, we captured the comments of the first three videos with the most views on the relevant theme on YouTube using the Python programming language. From each video, top-level comments that received 100 or more likes, along with 50 randomly selected comments with fewer than 100 likes, were analyzed through content analysis to ensure the inclusion of both prominent and less visible user perspectives. As a result of the analysis conducted on the data set collected from the YouTube platform, it was observed that users experienced feelings such as admiration and surprise toward the content produced with deepfake technology, as well as fear and anxiety regarding potential risks.

Keywords Deepfake · Artificial Intelligence · Synthetic Content · Content Manipulation · Social Media

Citation: Kaya, S. (2025). Investigation of user comments on videos generated by deepfake technology. *Acta Infologica*, *9*(1), 208-222. https://doi.org/10.26650/acin.1673277

Corresponding author: Sertaç Kaya sertackaya1@aydin.edu.tr

Introduction

New digital technologies and social media platforms allow users to create, share, and consume content whenever and wherever they want (Kaplan & Haenlein, 2010; Kietzmann et al., 2011). In this way, people have the opportunity to express themselves and share their ideas, access alternative content in online environments, and interact with other users. Although this situation has increased the channels for accessing information, the abundance of information causes an increase in false and fake content (Chen et al., 2015; Marin, 2021; Mintz, 2012), and therefore, it causes difficulties for people to verify and trust information (Zoone, Luoma-aho & Lievonen, 2024). Recent developments in the field of artificial intelligence have also affected the production and consumption of information in the digital environment.

In recent years, deepfake technology has come to the forefront of content production and consumption by users. This technology uses artificial intelligence to imitate an existing image or sound, creating content that appears real but is not real. Thus, it is possible to change the faces of the people in the video, shape their speech, or create completely new content (Agarwal et al., 2020; Farid, 2022). Such videos are circulated on the internet and social media platforms, gaining popularity and reaching a wide audience. This situation, along with its advantages, also carries some risks. While creative and entertaining content can be produced using deepfake technology (Mihailova, 2021; Murphy, 2023), fake and manipulative videos (Tahir et al., 2021) can also be produced. The ease of producing both types of content makes it increasingly difficult to distinguish between real and fake media content.

This study aims to reveal users' attitudes toward video content produced by deepfake technologies and their general approach to such technology. This study aims to determine whether people perceive this type of content production as a threat or opportunity. The examined video comments are used to identify clues about people's thoughts about such videos, their views on their possible effects, and whether they can distinguish between real and fake content.

Literature Review

Deepfake is a machine learning-based software tool (Albahar & Almalki, 2019). The word is derived from the combination of the words “deep learning” and “fake” (Mirsky & Lee, 2021). This technology refers to techniques that enable the production of realistic-looking audio and video recordings of people doing things they do not do or say (Johnson & Diakopoulos, 2021; Ruiter, 2021). It began to gain popularity in 2017 when an anonymous user placed celebrities on adult content and shared it on the Reddit platform (Hern, 2018; Kietzmann et al., 2020). Therefore, the use of these techniques produces digitally manipulated hyper-realistic videos (Taulli, 2019). The production of such videos is becoming easier and cheaper day by day, and with developing technologies, they are becoming more convincing and realistic (Metz, 2019; Rubenking, 2019). Today, the ease of performing operations such as face replacement or synthesis is due to reasons such as accessibility to large-scale public data and the development of deep learning techniques (Tolosana et al., 2020). This phenomenon causes people with a large digital footprint to become more vulnerable to the production of their own copies (Bohacek & Farid, 2024). Those who produce deepfakes consist of curious amateurs who use technology for entertainment and personal interest, actors who use them for political purposes for disinformation and propaganda, malicious individuals who commit crimes such as financial fraud and phishing, and legitimate organizations that use them for entertainment, advertising, or artistic projects within the legal framework (Westerlund, 2019).

In their study, Tolosana et al. (2020) explained the manipulation made via deepfake through the classifications of entire face synthesis, identity swap, attribute manipulation, and expression swap. The entire face synthesis creates the entirety of non-existent facial images. Identity swap refers to the replacement of a person's face in the video with the face of another person. Attribute manipulation emphasizes facial editing or retouching operations, such as hair or skin color, gender, and age. Expression swap, on the other hand, means changing a person's facial expression. In another study, deepfake uses in the context of human images were classified into four categories. These are reenactment, replacement, editing, and synthesis (Mirsky & Lee, 2021):

Reenactment: The person’s expression, mouth, gaze, or body are directed. In the film and video game industry, well-intentioned uses are encountered in educational media where actors’ performances are adjusted in post-production processes and historical figures are re-enacted.

Replacement: It is a transfer process. This involves replacing one content with another. It can be used in the fashion industry to visualize a person in different clothes or to produce entertaining outputs by changing the identity of an actor to that of a famous individual. However, it is also frequently encountered in malicious uses such as slander and blackmail.

Editing: It refers to the points where some characteristics of people are added, changed, or removed. Examples such as changing the target's clothes, age, weight, and beauty. In applications such as FaceApp, people can be given an example of changing their appearance for entertainment. The same process can be used to make a sick leader appear healthy, or fake identities can be created to mislead others.

Synthesis: Synthesis involves the use of human face and body synthesis techniques without any goal. It can be used to create characters for movies and games and to create fake personalities online.

As with many technological innovations, deepfakes bring opportunities and risks. While they can be used advantageously in sectors such as film, gaming, entertainment, and fashion (Kietzmann et al., 2020), the same technology can also be used for malicious purposes, such as blackmailing individuals, discrediting individuals or institutions, and manipulating fake news (Chesney & Citron, 2019; Fletcher, 2018; Verdoliva, 2020). For example, actors or singers can be placed in erotic content (Wang, 2019), used for revenge pornography (van der Nagel, 2020), or used to smear public figures in various ways (Maddocks, 2020). The use of deepfakes in this direction can also incite political or religious tensions, deceive the public, and create turmoil in financial markets by spreading false information (Heidari, 2024). Such fake content can be successful for basic reasons, such as credibility and accessibility (Kietzmann et al., 2020). Because it is easier to access such videos on social media platforms, the possibility of confusing viewers by sabotaging and imitating the facts because of problems with misinformation and conspiracy theories increases (Hasan & Salah, 2019; Yazdinejad et al., 2020). For example, a study has shown that deepfake videos are more vivid, convincing, and reliable than fake news articles (Hwang, Ryu & Jeong, 2021). The content produced using this technique creates uncertainty in the mind of the user; therefore, trust in the news on social media may decrease (Vaccari & Chadwick, 2020).

It is becoming easier to create disinformation by imitating political figures or by presenting pornographic content with images of celebrities (Heilweil, 2020). It has been revealed that 96% of deepfakes found online are pornographic, and almost all of them belong to women (Abram, 2020). Another study demonstrated that users who encounter deepfakes depicting politicians significantly worsen their attitudes toward that politician (Dobber et al., 2021). The finding that deepfake videos increased from 7,964 to 14,678 between December 2018 and July 2019 (Duffy, 2019), the number of these videos exceeding 85,000 in December 2020

and doubling every six months since December 2018 (Petkauskas, 2023) is also important in terms of showing the speed at which such content spreads. In addition, whether the person or persons in the deepfake will object to the way it is represented, whether the deepfake deceives the audience, and the purpose of creating the deepfake cause the moral discussion of the content produced with artificial intelligence (Ruiter, 2021). Therefore, the main threat is that deepfake-based content weakens our trust in the reality of what we see and hear (Floridi, 2018) and can easily lead individuals to false beliefs (Fallis, 2021). Because people, including educated ones, have difficulty distinguishing between real and fake (Matern et al., 2019). Although people trust their own abilities, the credibility created by having seen it results in difficulty in detecting deepfakes (Köbis, Doležalová & Soraperra, 2021).

In order to understand a technology like deepfake, it requires an examination from the perspectives of media and society, media production, media representation, media audiences, gender, law, and politics. Because it shapes social perception, users are involved in content production. Artificial intelligence is conveyed in the media with entertainment and horror themes, and the ability to interact with the audience comes to the fore. In addition, the first deepfake content has a gender-based analysis due to the creation of pornographic videos of fake female celebrities and revenge porn directed at women, the emergence of initiatives to reduce possible risks, and political contexts due to the potential to increase uncertainty and reduce trust through fake political videos (Karnouskos, 2020).

In line with all these definitions, approaches, and discussions, this study aims to better understand people's feelings and thoughts by examining deepfake videos and user perceptions toward this technology. The analysis will examine how different cognitive and emotional reactions to deepfakes are reflected in user comments.

Methodology

With the rise of artificial intelligence (AI) and deep learning techniques, the circulation of fake content in digital environments is rapidly increasing. Therefore, the main purpose of the study is to reveal users' feelings and thoughts about videos produced with deepfakes. In this context, this study seeks an answer to the research question, "What are the general perceptions of users toward deepfake videos?".

This study's sample selection process was carried out based on certain criteria. In the search made on the YouTube platform, among the videos that were reached with the keyword "deepfake video" and had over 10 million views, those that were open to user comments were preferred. Thus, only content that reached a large audience and received high interaction was included in the analysis. The aim was to make more meaningful inferences about the general reactions of the public to such content. However, two of the five videos that met the specified criteria were excluded from the study; one was closed to comments, and the other was excluded because its musical structure made user comments focus on artistic content, not deepfake technology. This limited the number of samples to three videos.

In the study, the comments of three videos on the YouTube platform were captured using the Python programming language. Since it was intended to examine directly the comments made on the videos determined as a sample, "top-level comments" were determined as the basic unit to be analyzed. In the evaluation process of the comments, those that received 100 or more likes were selected to better understand the impact of the content on large audiences. The primary reason for analyzing only "top-level comments" with over 100 likes in this study is that such comments are more likely to reflect views that have gained visibility, support, and resonance among a broader audience. In fact, in the three videos analyzed, the total number

of likes on comments with 100+ likes was 212,431, 133,555, and 87,733, respectively. These figures indicate that the analyzed comments are not merely individual expressions but also reflect opinions that have found broader social resonance. From this perspective, the selected comments serve as a strong sample to represent dominant perceptions and emotional responses among users, thereby contributing to a more meaningful and accurate interpretation of the public perception of deepfake technology. Nevertheless, in order to increase the diversity of viewpoints and ensure the inclusion of potentially underrepresented voices, an additional set of 50 randomly selected comments with fewer than 100 likes were incorporated into the analysis of each video. This approach aims to go beyond the most prominent comments and achieve a more balanced and comprehensive understanding of user perspectives, thereby increasing the representativeness of the findings. The user comments were analyzed using content analysis. Themes were created inductively on the basis of findings directly obtained from the data. The repetitive expressions in the comments and expressions that were similar to meaning were grouped under the same theme.

In order to ensure validity and reliability within the scope of content analysis, two academics independently conducted the coding process. The coding process was continued until 100% agreement was achieved among the academics. The aim of this approach is to ensure the accuracy of the coding process and the consistency of the obtained data. In this context, five themes were identified for the first two videos, and another five different themes were identified for the third video.

Results

The comments were taken from the YouTube platform on 17.03.2025. The video data corresponding to the relevant date are presented in Table 1:

Table 1
Content Information of Videos

Post Title	Number of Views	Number of Comments	Number of Likes > 100
Bill Hader impersonates Arnold Schwarzenegger [DeepFake]	22,565,151	23,170	214
Bill Hader channels Tom Cruise [DeepFake]	14,590,306	16,484	153
You Won't Believe What Obama Says In This Video! 😊	10,047,828	8,851	102

As part of the study, 214 comments from the first video, 153 from the second, and 102 from the third were analyzed based on their high number of likes. In addition, 50 randomly selected comments with fewer than 100 likes were included in each video. In total, 619 user comments were examined through content analysis. The first two videos were uploaded to the YouTube platform by the account named "Ctrl Shift Face," while the third video was uploaded by the account named "BuzzFeedVideo." This situation causes differences in the content and context of the videos; thus, the thematic distribution of user comments also differs. For this reason, the comments of the first two videos were categorized under the same themes, while the comments of the third video were categorized under different themes.

Table 2

Theme and Comment Information of the Video "Bill Hader impersonates Arnold Schwarzenegger [DeepFake]"

Theme and Comment Information of the Video "Bill Hader impersonates Arnold Schwarzenegger [DeepFake]"	Number of Comments > 100+ Likes	Number of Comments Under 100 Likes
The Distortion of Reality Perception and Creepy Effect	85	28
Admiration for and Amazement at Deepfake Technology	51	8
Entertainment and Humor	54	6
Bill Hader	5	2
Future Risks of Deepfake Technology	19	4

Table 2 presents the thematic categorization of user comments on the video, providing insight into the distribution of reactions based on their content and engagement levels. The comments on the theme of distortion of perception of reality and creepy effects focused on how users felt when they realized how realistic the deepfake technology in the video was. Commenters expressed their feelings of being disconnected from reality with sentences such as "I really questioned my sanity.", "I thought I was having a stroke.", "I was about to drive myself to the emergency room.", "My mind was blown for about 5 minutes.", "Sh*t am I drunk?". Comments such as "That might officially be the creepiest thing I've ever seen...", "Jesus christ, this technology is terrifying." and "Its creepy how seamlessly the face swap is working." stand out as examples of users' fear.

The comments on admiration for and amazement for deepfake technology emphasize that this technology has made great progress and that the level of reality is high. Users state that the face swap process looks extremely natural and express their admiration for the created content. For example, a comment such as "From the thumbnail I thought that this was Arnold's son." indicates that the user found the face swap realistic and thought that it was an image of Arnold Schwarzenegger's son. In addition, comments such as "How did u do this. This looks pretty amazing." and "Bloody hell that's an impressive DeepFake, the transitions are really subtle." express the admiration felt by those who watched the video.

The theme of entertainment and humor comments focuses on users writing about the possibilities offered by deepfake technology in an exaggerated, funny, and sometimes critical language. Comments such as "If he had the Arnold Schwarzenegger Body. He would be him 90%", "He is the Arnold who didn't choose bodybuilding as a career." and "This is what a non bodybuilding Arnold Schwarzenegger would look like." are seen as ironic reflections of an iconic appearance being more ordinary or unusual. Here, users are attempting to make funny inferences from the new perspectives revealed by deepfakes. In addition, the expressions "And they spent millions for mustache in JL..." and "They should have removed henry cavills moustache like this." refer to the removal of Henry Cavill's mustache with CGI (Computer-generated imagery) in the 2017 film Justice League. In the film, the actor's mustache was digitally removed, but the result was not very successful. This comment also humorously expresses that deepfake technology can be more successful in this type of visual effects.

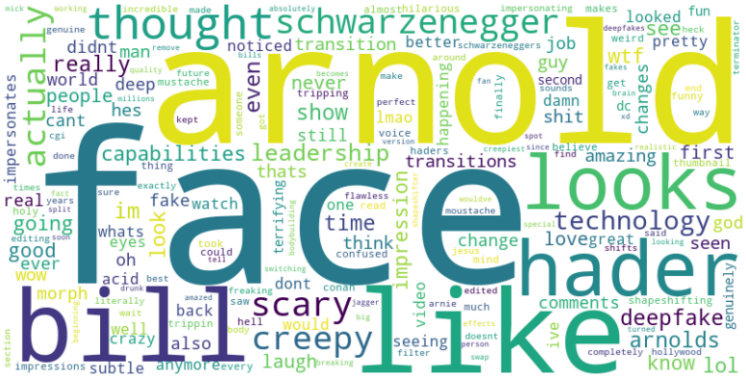
In the theme about Bill Hader, users highlight Bill Hader's ability to imitate a subject that is truly commendable, despite the impressiveness of deepfake technology. Comments such as "Huge fan of Bill Hader since Superbad." An even bigger fan now that he transformed into Arnold ;)", "I love how Bill Hader just enjoys himself. He's a real genuine guy." and "I love it when Bill giggles at himself. Always makes me laugh as much as when he's doing impressions." are prominent examples in this context. It can be seen that Hader's talent and approach, independent of deepfake technology, are praised in these comments.

Comments on the future risks of deepfake technology reflect serious concerns about its ethical, legal, and societal risks. Expressions such as “This is scary technology. Soon we won’t know if what we’re watching is real or fake.”, “Amazing and scary at the same time. It looks almost perfect. In 2-3 years max this will be beyond recognizable.”, “This is so absolutely creepy. Imagine how this will impact politics, especially now when you can’t trust anything the “media” says.” emphasize that it will cause people to question the accuracy of what they watch or hear. These expressions also indicate that we are entering a period in which the perception of reality will be shaken, trust in information will decrease, and people will be forced to question the content they watch. In addition, “The implications of possible misuse of this technology are terrifying!” “Deepfakes are going to be used to frame people one day. I think it's likely already happening.”, “Political deepfake video is what will cause world war 3.” comments like these reflect concerns about the potential misuse of this technology. These statements indicate that deepfake technologies can also be used for propaganda, false accusations, and disinformation. These increases concerns that the boundaries between real and fake will become increasingly blurred, which could have dangerous consequences for individuals and communities.

The comments that were randomly selected from the comments that received fewer likes in order to increase the variety of the comments also support the previous findings. It can be seen that users gave similar reactions in the comments. Especially in the comments under the title “Distortion of Reality Perception and Creepy Effect”, there are statements such as “Woah wtf I thought I was losing my shit...”, “What the heck did I just watch?”, “damn, it's both creepy and impressive.”, “fuckin scary...”, “This is... confusing... Stop doing that.”, “this is scary. but fun. but scary.” and “Trippy!” These statements demonstrate that users have difficulty believing what they see and the confusion this creates. The impressive aspect of the technology is emphasized with comments such as “this is the best deepfake i have seen”, and “I'm so impressed by the flawless camera trick.” in the category of “Admiration for and Amazement at Deepfake Technology”. Humorous comments and praise for Bill Hader's acting show that viewers find the content entertaining and admirable at the same time. However, some comments also raise serious ethical concerns by drawing attention to the risks that technology may pose in the future, especially manipulation and information security. Examples of this situation are comments such as “I think this technology will confuse the heck out of people rather soon..”, “Creepy shit .you could be framed for a crime you didn't do .fuck that ...”, “This was so funny but also so scary and just the thought of this tech being misused is very unsettling.”, “This is why you can't ever believe anything the news or TV tells u . It's really going to be bad for us.”. Two comments that were not directly related to the content were excluded from the analysis.

Figure 1

Word Cloud of Comments on “Bill Hader impersonates Arnold Schwarzenegger [DeepFake]” Video



As illustrated in Figure 1, the word cloud highlights the most frequently used words in user comments, reflecting viewers' emotional responses to the video. In addition, we determined that certain words reflecting the emotional reactions created by the content and deepfake technologies in the viewers were frequently used in the examined user comments. In particular, "creepy" and its derivatives (e.g. creepiest, creepiness, etc.) were used in a total of 23 comments, "scary" 18 times, "crazy" 5 times, and "terrifying" 4 times. These data demonstrate that viewers adopt emotional language when evaluating videos. In addition, these words, which are preferred to be used at a high frequency, also show that the video content and deepfake technology are perceived as disturbing, giving a sense of uncanny or unusual.

In addition, the use of words with positive value judgments, such as "Good" (9 times) and "Amazing" (7 times), demonstrates that users find content and technology esthetically and technically successful and gain their appreciation and admiration. Another striking point is the frequent use of expressions such as "Really" (11 times), "Wow" (8 times), "Wtf" (11 times) and "Confused" and its derivatives (7 times). This situation reveals that users feel astonishment, amazement, and sometimes even confusion toward the content and the technology.

Table 3

Theme and Comment Information of the Video "Bill Hader channels Tom Cruise [DeepFake]"

"Bill Hader channels Tom Cruise [DeepFake]" Video Themes	Number of Comments > 100+ Likes	Number of Comments Under 100 Likes
The Distortion of Reality Perception and Creepy Effect	39	16
Admiration for and Amazement at Deepfake Technology	51	11
Entertainment and Humor	42	8
Bill Hader	6	1
Future Risks of Deepfake Technology	15	10

Table 3 presents the thematic classification of user comments on the video 'Bill Hader channels Tom Cruise [DeepFake],' reflecting viewers' emotional and cognitive reactions, especially concerning the unsettling nature of deepfake technology. Commenting on the theme of distortion of perception of reality and creepy effect, expressions such as "Let's be real.....this can't be 'normal'", "That's crazy.", "Holy shit...this truly creepy.", "I instantly questioned my sanity." stand out. These comments demonstrate that the user's perception of reality has been shaken, and they are worried. One user's statement "If you hadn't put deep fake in the description I would be questioning reality because it is such a subtle transition." shows that they only grasped the truth through the comments of other users. This situation can be evaluated as an indicator that individual perception has reached a level that can be easily manipulated using deepfake technology. Another user's statement "Im a software engineer... im always amazed with new technology coming out...but this is scary." proves that even people with technical knowledge are worried about developing technologies.

Comments on the theme of admiration for and amazement toward deepfake technology include expressions such as "The way you seamlessly transition between Cruise and Hader's face is mind-bending, genuinely amazing stuff." and "This is without a doubt the most realistic deepfake I have seen so far." These comments can be evaluated as praise indicating how much deepfake technology has developed and how close it has come to reality. In addition, one user commented that "1:39 is such an impressive transition because it's so damn quick and still seamless. This is the best deepfake I've ever seen." indicates that the lines between real and fake are disappearing, and another user's statement "This is hands down, the

creepiest but the best and funnest and coolest shit ive seen in some time." shows both admiration for the technique and concern about its potential.

In entertaining and humorous comments, many users reference different series and movies and express their desire for this technology to be applied to such productions. Comments requesting that different actors' faces be added to productions such as Superman, Iron Man, Inglourious Basterds, and The Office highlight the entertaining aspects of deepfake technology and show that it is seen as a creative tool that can offer new and alternative versions to viewers. Another user commented, "Bill Hader has Mystique's powers, but only the face lol." Here, the user humorously emphasizes the effectiveness of deepfake technology by comparing Bill Hader's ability to skillfully change facial expressions to the powers of Mystique, the shapeshifting character in the X-Men series.

Comments on the theme about Bill Hader show that the actor is appreciated not only for the impressiveness of deepfake technology, but also for his extraordinary ability to imitate, modesty, and natural acting skills. Viewers express admiration for Hader, highlighting his talent regardless of the deepfake effects.

The theme of the future risks of deepfake technology highlights serious concerns about the potential for misuse of deepfakes and the risks they may pose in the future. "This technology should be destroyed, could ruin innocent people's lives.", "This is both amazing and creepy. And the technology is going to create a shit ton of distrust in media (or any) videos in the near future.", "We will never be able to believe anything we see on video.", "The potential uses for this technology are frightening.", "Just imagine how hard we've been already manipulated by this.", "This technology will be used to frame someone and most likely has. Scary future is in store for us.", "They're going to try to weaponize this for the 2020 election." These comments highlight the ethical and security concerns created by deepfake technology. Users believe that this technology can undermine media credibility, can be used to slander or manipulate people with fake content, and can have negative effects on political processes.

There are statements similar to the previous statements in the comments under 100 likes for this video. The comments under the title Distortion of Reality Perception and Creepy Effect illustrate the confusion, discomfort, and anxiety that technology creates in individuals. Comments such as "Yes ... Creepy", "That was creepy as hell!", "My brain is confused and hurt.", "This melted my brain", "This is so strange to watch" reflect these feelings. Comments such as "Awesome", "Incredible" and "How? How?" can be interpreted as indicators of feelings of admiration and surprise. On the other hand, the comments under the category of Future Risks of Deepfake Technology carry serious concerns about the ethical and social consequences of the technology: Comments such as "This is the beginning of the end of 'video proof'" or "This kind of technology is terrible and able to make a lot of disasters, should be forbidden" show that deepfake is perceived not only as a means of entertainment but also as a potential threat. The five comments examined in this context consist of expressions that were not directly related to the content.

Figure 2

Word Cloud of Comments on "Bill Hader channels Tom Cruise [DeepFake]" Video

Figure 2 visualizes the most frequently used words in user comments on the video, illustrating the coexistence of both positive and negative emotional responses to deepfake technology. It is also noteworthy that expressions such as "Creepy" and its derivatives (15 times), "Crazy" (5 times), and "Scary" (6 times) were used in the comments. The preference for these adjectives that express discomfort and concern indicates that users find the deepfake technology in the video creepy or damaging to their perception of reality. On the other hand, it was determined that words with positive judgments such as "Good" (9), "Amazing" (10), "Love" (8), and "Great" (6) were also used with remarkable frequency for the same content and technology. This shows that users evaluate video and deepfake technology as technically successful and esthetically problem-free.

Table 4

Theme and Comment Information of the Video "You Won't Believe What Obama Says in This Video! 😊"

"You Won't Believe What Obama Says In This Video! 😊" Video Themes	Number of Comments > 100+ Likes	Number of Comments Under 100 Likes
Ironic and Critical Comments on BuzzFeed's Credibility	32	10
Obama-Focused Commentary	8	7
Internet Culture Humor	32	11
Media Literacy and Fake Content Warnings	15	15
Unexpected and Surprising Content	15	6

Table 4 presents the thematic distribution of user comments, revealing a range of responses including ironic critiques of source credibility, humorous engagement with internet culture, and concerns related to media literacy and misinformation. The theme of ironic and critical comments regarding BuzzFeed's reliability focuses on the source sharing of the content rather than the content of the video. In their comments, users express ironic, sarcastic, and critical reactions to BuzzFeed being presented as a "reliable news source." Comments such as "Is Buzzfeed reliable? That's a funny joke.", "*The last news source ill trust is buzz feed*", "It's ironic how it's coming from buzzfeed.", "Yea on the list of non trusted news sources buzzfeed is under cnn." stand out under this theme. In the comments, the source's history and content strategy are questioned, while its reliability as a source is criticized using a humorous and sarcastic tone.

In the Obama-focused comments theme, users express admiration for Obama's personal characteristics or general image. They are seen to express these feelings with comments such as "Please come back as a president!", "He still looks cool.", "We want you back Obama! We want you back Obama and I know i'm

honestly not alone ;) ♡♡." In addition, critical and humorous statements are also included among these comments and directly refer to Obama.

The comments on internet culture humor carry humorous, viral, and popular culture effects that reflect the typical characteristics of internet culture. In comments such as “Plot twist: That wasn't Jordan Peele either. It was actually Bernie Sanders.” and “And behind Jordan Peele was Kanye West, who was actually just Donald Trump, who was actually Beyonce doing an impression, who was just Putin, which was fake and it was just John McCain who was actually just played...” the distortion of the perception of reality is addressed in a humorous style. In addition, many users repeated a swear word from the video in their comments, creating a viral effect. In addition, the comment “Remember guys an apple is never anything but an apple - cnn” mocks the media’s oversimplified analyses. Another striking point is that users create absurd humor by adding historical figures to today's digital problems using fake quotes. Sentences such as "'Most Quotes On The internet Are Fake' Winston Churchill", "'My death was a deepfake' -John F. Kennedy" and "'Don't believe everything you see on the internet.'- Sun Tzu The Art of War" reveal this situation.

The comments on the theme of media literacy and fake content warnings focused on concerns about the reliability of information created by deepfake technology. Comments such as “People need to be more aware of what they see on the internet especially on Facebook, be more cautious about what they read and watch.”, “This is great. Thanks buzzfeed. This is important to show how this happens and what fake can look like.”, “As hilarious as this video is, it's actually very educational. It teaches us that we can’t believe anything we see or hear on the internet, or TV, (especially the news) even if the words appear to be spoken straight from the person's mouth.” also indicate concerns. These comments are important in terms of demonstrating how deepfakes and artificial intelligence can erode social trust and that digital reality is becoming increasingly uncertain.

Comments on the theme of unexpected and surprising content express users' admiration and surprise despite technology and video. Comments such as “Well that was unexpected lol.”, “Damn. that confused me.”, “I did NOT expect this.”, “Whahahahaaaat?! I'm Shook.”, “I am dazed and confused.” reveal users' emotional states.

User reactions with comments with less than 100 likes reflect an atmosphere of distrust focused on the platform’s media reputation and political positioning. Users appear to directly target BuzzFeed and adopt a sarcastic and critical tone toward its content. For example, the phrases “ya trusted news sources like BuzzFeed” and “Trusted news source? Definitely not BuzzFeed nor CNN.” ironically turns the discourse of trustworthiness upside down. The comments about Obama, on the other hand, show that he is a figure that has been targeted by both intense criticism and conspiracy theories. Expressions such as “One screwed up dude. Obama really fucked things up.”, and “insane”, “fake Muslim” reflect these thoughts. In addition, there are comments in which the discourses in the video are evaluated using an entertaining approach. Here, popular movie and series references and comments from the video stand out. In comments such as “OMGG”, “I love this”, “How did you do this?”, users’ appreciation and surprise are observed. There were also thoughts about how advanced AI tools like deepfake technology could manipulate content. While some users criticized the way in which they recognized deepfake technology and analyzed such content in these comments, others pointed out that this technology could be more dangerous in the future. Among the comments, one comment was out of context.

Figure 3

Word Cloud of Comments on "You Won't Believe What Obama Says in This Video! 😊" Video

Figure 3 illustrates the prominent word frequencies in user comments, highlighting a predominantly critical discourse centered on the credibility of the source sharing the video rather than the video content itself. The word frequencies in the comments of this video show that users are dominated by a critical discourse that focuses on BuzzFeed, which shared the video rather than the content and deepfake technology. Words such as "buzzfeed" (35), "news" (37), "trusted" (26), "sources" (17), and "source" (9) that occur at high frequencies in the comments demonstrate that users frequently emphasize the issue of trust in the institution that shared the video content. The word preferences of "fake" (15), "lol" (8), and "ironic" (6) support that this questioning is made in a sarcastic and skeptical tone.

Discussion and Conclusion

The creation and modification of visual and audio media content has become extremely easy with the use of artificial intelligence. The deepfake process represents the newest way to create misleading media content. For this reason, such technologies have paved the way for creating fake voices, videos, and photographs. Since social media platforms offer a significant photo archive, some programmers can produce the content they want from these images without the knowledge and consent of the other party. The creation of video content that includes imitations of many celebrities and politicians is important in terms of identifying possible risks. This situation can cause harm to the people in the videos on the one hand, and on the other hand, it can cause serious social consequences by damaging the public's trust in the content that is believed to be true.

The study shows that users often react to deepfake videos with a mix of surprise and fear, especially due to the high level of realism. Some find the technology so convincing that it makes them question their perception of reality. The ability of deepfakes to closely imitate real people and events leads to significant viewers' unease. Alongside these concerns, many users express admiration for the technology's sophistication. The realism achieved by deepfakes is seen as both impressive and creatively powerful, reflecting the rapid progress made in artificial intelligence and media generation tools. However, users also highlight the serious risks of misuse. They voice strong concerns about the social, ethical, and legal challenges deepfakes may cause, particularly the spread of false information. Many fears that such content could erode public trust and blur the line between reality and fabrication, making it increasingly difficult to verify the truth.

These findings highlight a broader issue: the blurring of boundaries between reality and simulation in the digital age. The increasing difficulty of distinguishing fake from real images threatens public trust, especially on social platforms where content is rapidly consumed and shared. This raises urgent questions about how

societies can adapt to a media environment in which perception can be manipulated so convincingly. To address these risks, several steps should be taken. Media literacy programs must be strengthened to help users critically evaluate digital content. Policymakers and technology developers must work together to establish transparent standards, improve detection tools and regulate the misuse of synthetic media. As Johnson and Diakopoulos (2021) suggested, limiting access to deepfake training data and designing technical verification mechanisms are key to mitigation. Similarly, as noted by Maras and Alexandrou (2019), counter-technologies capable of identifying fake content must evolve in parallel.

The European Union's General Data Protection Regulation (GDPR) and Turkey's Law No. 6698 on the Protection of Personal Data (KVKK) provide important frameworks for safeguarding individuals' visual and audio data; however, they may fall short in addressing AI-based threats such as deepfakes. Therefore, there is a need for additional legal provisions targeting biometric data manipulation, and supporting content verification systems are required. Moreover, the development of labeling mechanisms by social media platforms to indicate content authenticity can play a significant role in mitigating the negative impacts of deepfake technology.

Developing artificial intelligence technologies will cause more fake content to exist in the digital environment in the future due to the ease of access and use. Although fake content and disinformation are not new problems, deepfake and similar technologies significantly increase the difficulty of finding the truth and distinguishing lies.

Peer Review Externally peer-reviewed.
Conflict of Interest The author has no conflict of interest to declare.
Grant Support The author declared that this study has received no financial support.

Author Details Sertaç Kaya
[1] İstanbul Aydin University, Faculty of Communication, Department of Journalism, İstanbul, Türkiye
0000-0003-3483-572X sertackaya1@aydin.edu.tr

References

Abram, C. (2020). The most urgent threat of deepfakes isn't politics. It's porn. Retrieved from https://www.vox.com/2020/6/8/21284005/urgent-threat-deepfakes-politics-porn-kristen-bell

Agarwal, S., Farid, H., Fried, O., & Agrawala, M. (2020). Detecting deep-fake videos from phoneme-viseme mismatches. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 660-661).

Albahar, M., & Almalki, J. (2019). Deepfakes: Threats and countermeasures systematic review. *Journal of Theoretical and Applied Information Technology*, *97*(22), 3242-3250.

Bohacek, M., & Farid, H. (2024). The making of an AI news anchor—and its implications. *Proceedings of the National Academy of Sciences*, *121*(1), e2315678121. https://doi.org/10.1073/pnas.2315678121

Chen, X., Sin, S. C. J., Theng, Y. L., & Lee, C. S. (2015). Why students share misinformation on social media: Motivation, gender, and study-level differences. *The Journal of Academic Librarianship*, *41*(5), 583-592. https://doi.org/10.1016/j.acalib.2015.07.003

Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. Retrieved from https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakesand-new-disinformation-war

De Ruiter, A. (2021). The distinct wrong of deepfakes. *Philosophy & Technology*, *34*(4), 1311-1332. https://doi.org/10.1007/s13347-021-00459-2

Dobber, T., Metoui, N., Trilling, D., Helberger, N., & De Vreese, C. (2021). Do (microtargeted) deepfakes have real effects on political attitudes? *The International Journal of Press/Politics*, *26*(1), 69-91. https://doi.org/10.1177/1940161220944364

Duffy, R. (2019). New research finds 96% of deepfakes are pornographic. Retrieved from https://www.emergingtechbrew.com/stories/2019/10/09/new-research-finds-96-deepfakes-pornographic

Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, *34*(4), 623-643. https://doi.org/10.1007/s13347-020-00419-2

Farid, H. (2022). Creating, using, misusing, and detecting deep fakes. *Journal of Online Trust and Safety*, *1*(4). https://doi.org/10.54501/jots.v1i4.56

Fletcher, J. (2018). Deepfakes, artificial intelligence, and some kind of dystopia: The new faces of online post-fact performance. *Theatre Journal*, *70*(4), 455-471. https://doi.org/10.1353/tj.2018.0097

Floridi, L. (2018). Artificial intelligence, deepfakes and a future of ectypes. *Philosophy and Technology*, *31*, 317-321. https://doi.org/10.1007/s13347-018-0325-3

Hasan, H. R., & Salah, K. (2019). Combating deepfake videos using blockchain and smart contracts. *Ieee Access*, *7*, 41596-41606. https://doi.org/10.1109/ACCESS.2019.2905689

Heidari, A., Jafari Navimipour, N., Dag, H., & Unal, M. (2024). Deepfake detection using deep learning methods: A systematic and comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(2), e1520. https://doi.org/10.1002/widm.1520

Heilweil, R. (2020). How deepfakes could actually do some good. Retrieved from https://www.vox.com/recode/2020/6/29/21303588/deepfakes-anonymous-artificial-intelligence-welcome-to-chechnya

Hern, A. (2018). My May-Thatcher deepfake won't fool you but its tech may change the World. Retrieved from https://www.theguardian.com/technology/2018/mar/12/may-thatcher-deepfake-face-swap-tech-change-world

Hwang, Y., Ryu, J. Y., & Jeong, S. H. (2021). Effects of disinformation using deepfake: The protective effect of media literacy education. *Cyberpsychology, Behavior, and Social Networking*, *24*(3), 188-193. https://doi.org/10.1089/cyber.2020.0174

Johnson, D. G., & Diakopoulos, N. (2021). What to do about deepfakes. *Communications of the ACM*, *64*(3), 33-35. https://doi.org/10.1145/3447255

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, *53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Karnouskos, S. (2020). Artificial intelligence in digital media: The era of deepfakes. *IEEE Transactions on Technology and Society*, *1*(3), 138-147. https://doi.org/10.1109/TTS.2020.3001312

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241-251. https://doi.org/10.1016/j.bushor.2011.01.005

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, *63*(2), 135-146. https://doi.org/10.1016/j.bushor.2019.11.006

Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *Iscience*, *24*(11), 1-17. https://doi.org/10.1016/%20j.isci.2021.103364

Maddocks, S. (2020). 'A deepfake porn plot intended to silence me': Exploring continuities between pornographic and 'political' deep fakes. *Porn Studies*, *7*(4), 415-423. https://doi.org/10.1080/23268743.2020.1757499

Maras, M. H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence & Proof*, *23*(3), 255-262. https://doi.org/10.1177/1365712718807226

Marin, L. (2021). Sharing (mis) information on social networking sites. An exploration of the norms for distributing content authored by others. *Ethics and Information Technology*, *23*(3), 363-372. https://doi.org/10.1007/s10676-021-09578-y

Matern, F., Riess, C., & Stamminger, M. (2019). Exploiting visual artifacts to expose deepfakes and face manipulations. In *2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, (pp. 83-92). https://doi.org/10.1109/WACVW.2019.00020

Metz, R. (2019). The fight to stay ahead of deepfake videos before the 2020 US election. Retrieved from https://edition.cnn.com/2019/06/12/tech/deepfake-2020-detection/index.html

Mihailova, M. (2021). To dally with Dalí: Deepfake (Inter) faces in the art museum. *Convergence*, *27*(4), 882-898. https://doi.org/10.1177/13548565211029401

Mintz, A. P. (Ed.). (2012). *Web of deceit: Misinformation and manipulation in the age of social media*. Medford, N.J: CyberAge Books.

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys (CSUR)*, *54*(1), 1-41. https://doi.org/10.1145/3425780

Murphy, G., Ching, D., Twomey, J., & Linehan, C. (2023). Face/Off: Changing the face of movies with deepfakes. *Plos One*, *18*(7), e0287503. https://doi.org/10.1371/journal.pone.0287503

Nirkin, Y., Wolf, L., Keller, Y., & Hassner, T. (2021). Deepfake detection based on discrepancies between faces and their context. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10), 6111-6121. https://doi.org/10.1109/TPAMI.2021.3093446

Petkauskas, V. (2023). Report: number of expert-crafted video deepfakes double every six months. Retrieved from https://cybernews.com/privacy/report-number-of-expert-crafted-video-deepfakes-double-every-six-months/

Rubenking, N. J. (2019). How lab mice are helping detect deepfakes. Retrieved from https://www.pcmag.com/news/how-lab-mice-are-helping-detect-deepfakes

Tahir, R., Batool, B., Jamshed, H., Jameel, M., Anwar, M., Ahmed, F., ... & Zaffar, M. F. (2021). Seeing is believing: Exploring perceptual differences in deepfake videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). https://doi.org/10.1145/3411764.3445699

Taulli, T. (2019). Deepfake: What you need to know. Retrieved from https://www.forbes.com/sites/tomtaulli/2019/06/15/deepfake-what-you-need-to-know/?sh=d9c8c23704da

Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, *64*, 131-148. https://doi.org/10.1016/j.inffus.2020.06.014

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media+ Society*, *6*(1), 1-13. https://doi.org/10.1177/2056305120903408

Van der Nagel, E. (2020). Verifying images: Deepfakes, control, and consent. *Porn Studies*, *7*(4), 424-429. https://doi.org/10.1080/23268743.2020.1741434

Van Zoonen, W., Luoma-aho, V., & Lievonen, M. (2024). Trust but verify? Examining the role of trust in institutions in the spread of unverified information on social media. *Computers in Human Behavior*, *150*, 107992. https://doi.org/10.1016/j.chb.2023.107992

Verdoliva, L. (2020). Media forensics and deepfakes: an overview. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 910-932. https://doi.org/10.1109/JSTSP.2020.3002101

Wang, C. (2019). Deepfakes, revenge porn and the impact on women. Retrieved from https://www.forbes.com/sites/chenxiwang/2019/11/01/deepfakes-revenge-porn-and-the-impact-on-women/

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, *9*(11), 39-52. http://doi.org/10.22215/timreview/1282

Yazdinejad, A., Parizi, R. M., Srivastava, G., & Dehghantanha, A. (2020). Making sense of blockchain for ai deepfakes technology. In *2020 IEEE Globecom Workshops (GC Wkshps)*, 1-6. http://doi.org/10.1109/GCWkshps50303.2020.9367545