[เดวิด นุดี หน้า 71 - 80]

การรับรู้ความสมจริงของเทคโนโลยีดีพเฟคสื่อภาพยนตร์ฮอลลีวู้ด

Audience Perceptions of Authenticity in Deepfake Hollywood Film Media

เดวิด เอสเตบา. *วิทยาลัยนิเทศศาสตร์ มหาวิทยาลัยรังสิต.*
นุดี หนูไพโรจน์. *วิทยาลัยนิเทศศาสตร์ มหาวิทยาลัยรังสิต.*
David Esteba. *College of Communication Arts, Rangsit University.*
Nudee Nupairoj. *College of Communication Arts, Rangsit University.*
Email : david.g64@rsu.ac.th, nudee.n@rsu.ac.th

Abstract

This study delves into Hollywood film audiences' perceptions of deepfakes, specifically how they influence the perceived authenticity of content featuring these manipulated visuals. Building on industry-specific authenticity components defined by Nunes et al. (2021 : 1-20), this study explores audience perceptions of deepfaked content in Hollywood films advertisements and identifies key sentiments influencing the audience's perception of authenticity in deepfaked content. Two approaches were employed. First, comment data was scraped from YouTube advertisements of six chosen films featuring deepfaked characters. Second, interviews were conducted with fifty participants. Findings from both datasets were then compared.

The research reveals that deepfakes can create a sense of authenticity, particularly when eliciting emotional responses and honoring the original actors' legacy. However, the effectiveness hinges on execution quality and individual viewer perception. Both data sources yielded consistent results regarding perceptions of Accuracy, Connectedness, Integrity, and Originality. In simpler terms, audiences exhibited tolerance for deepfakes in these areas. Interestingly, the interview data diverged from scraped comments regarding Legitimacy (authorized use of character likeness) and Proficiency (technical mastery). Legitimacy was validated by scraped data but not interviews, while neither source validated Proficiency.

This study offers valuable insights for filmmakers, visual effects studios, and audiences alike. For filmmakers, the research sheds light on audience perceptions of deepfakes, allowing them to make informed decisions about incorporating this technology. It highlights the importance of high-quality execution and respecting the legacy of the characters being portrayed. For visual effects studios, the study provides valuable feedback on audience expectations regarding deepfakes.

This can guide them in their efforts to refine and perfect the technology to achieve a more convincing level of authenticity. For audiences, the research fosters a deeper understanding of how deepfakes can influence their perceptions. It encourages a more critical viewing experience, prompting them to question the authenticity of what they see on screen. Overall, this research paves the way for a more nuanced understanding of deepfakes in the film industry. By recognizing their potential to create emotional connection while acknowledging the limitations in achieving true authenticity, all stakeholders can work towards a future where deepfakes are used responsibly and effectively in storytelling.

Keywords : Deepfake, film advertising, audience perception, authenticity, Hollywood movies

Background of the Study

Hollywood is witnessing a revolution in performance with the emergence of deepfakes. These hyper-realistic digital avatars, powered by Artificial Intelligence (AI), are blurring the lines between traditional acting and digital manipulation (Somers, 2020). Deepfakes can seamlessly alter an actor's appearance, replace dialogue, or even remove background noise, offering cost-effective solutions for filmmakers. Deepfakes leverage AI to create highly realistic alterations of videos or photos. They can replace a person's face entirely, modify an existing performance, or even craft completely synthetic actors. The control over deepfakes can vary, allowing for animator or actor control through motion capture, or even independent generation based on prompts. Deepfakes owe their rise to advancements in Generative Adversarial Networks (GANs), a type of AI that can create realistic images, as showcased in a seminal 2014 paper (Goodfellow et al., 2014). While initially requiring expensive hardware, Hollywood quickly saw the potential. One early example is the use of deepfakes (or a combination of deepfakes and CGI) to complete scenes featuring the late Paul Walker in the 2015 movie Furious 7.

Traditionally, audiences craved genuine portrayals, evidenced by the acclaim for Heath Ledger's iconic Joker (Oscars, 2009). However, Hollywood's embrace of deepfakes – AI-powered digital actors – raises questions about authenticity. This shift is driven partly by deepfakes' economic benefits. The 2023 Screen Actors Guild strike exposed a deep divide. Actors protested the loss of control over their likenesses with deepfakes, highlighting a disconnect with the industry. While their concerns were valid, audience empathy remained low (Leger, 2023). This suggests studios may prioritize cost-efficiency over actors. The rise of deepfakes challenges traditional notions of authenticity in film advertising. Movies like Warner Bros.' Aquaman: The Lost Kingdom (2023) seamlessly blend deepfakes and CGI, blurring the lines between live-action and animation (Warner Brothers Ent., 2023). This raises the question: can deepfakes achieve authenticity, and how do audiences perceive synthetic elements in marketing compared to traditional content? This study aims to analyze the perceived authenticity of deepfakes in Hollywood film advertising and explore how audiences react to synthetic content alongside traditionally filmed material.

Research Objectives

This study delves into the complexities of audience experience with deepfaked film advertisements, guided by two core objectives: to explore the degree to which deepfakes replicate the elements traditionally associated with authenticity in film acting; and to investigate the key sentiments that shape audience perception of authenticity when presented with deepfaked content.

The research focuses specifically on deepfakes within Hollywood film advertising. Data was drawn from publicly available YouTube comments in North America, alongside in-depth interviews with Rangsit University alumni in Thailand. The study intentionally avoids an overly broad historical analysis of deepfakes, relying on the context established earlier. The research concentrates on the immediate and direct perceptions of audiences regarding authenticity in deepfaked Hollywood film advertisements.

Literature Review

This review explores the current understanding of deepfakes within the film industry, focusing on their impact on audience perceptions of authenticity. Several studies provide foundational knowledge about deepfakes themselves. Several studies explore the technical aspects of technology, outlining its creation processes, potential benefits, and inherent risks (Abdulreda and Obaid, 2022 : 745-755; Mahmud et al., 2023 : 13-23; Westerlund, 2019 : 40-53). These studies highlight the potential for deepfakes to be misused for malicious purposes, such as spreading misinformation or damaging reputations. Furthermore, they emphasize the need for ongoing development of detection methods and regulations to mitigate these threats.

The review of literature also found studies that delve into the psychological impact of deepfakes on audiences. Studies by Murphy et al. (2023 : 1-19) and Murphy & Flynn (2022 : 480-492) suggest that deepfakes might not be inherently more likely to distort memories compared to other forms of media, such as text descriptions. However, their findings also reveal a sense of unease among audiences regarding deepfakes. Participants expressed concerns about the technology's potential to disrupt the shared social experience of film and erode artistic integrity Murphy et al., 2023 : 1-19 ; Murphy & Flynn, 2022 : 480-492).

While significant research explores the potential for misuse of deepfakes in areas like user-generated content, malicious impersonations in pornography, and political propaganda, ethical concerns extend beyond these well-documented issues. Deepfakes also raise questions about their responsible use in educational contexts, particularly documentary filmmaking. Lees (2023 : 1-22) delves into the specific use of deepfakes within documentary filmmaking and raises critical ethical questions about the manipulation of reality in factual presentations. The ability to seamlessly alter voices or faces in documentaries using deepfakes creates tension between artistic expression and the filmmaker's obligation to represent truth. Lees' work highlights the need for a deeper conversation about the ethical boundaries surrounding deepfake usage in documentaries.

While existing literature offers valuable insights into deepfakes and their potential societal and ethical implications, a significant knowledge gap remains. Although Bode et al. (2021 : 849–854) provide a compelling discussion on the broader ramifications of deepfakes in film, and Jin et al. (2023) investigate external factors influencing the believability of deepfakes, a more nuanced understanding of how audiences specifically evaluate the authenticity and emotional engagement of deepfakes within film narratives or advertising is lacking.

Research Objectives

1. To investigate whether deepfake technology can replicate the impact of authentic media.

2. To identify key sentiments influencing the audience's perception of authenticity in deepfaked content.

Research methods

The research used a two-step approach.

In the first step, 100 public comments were collected and analyzed from YouTube trailers featuring deepfakes in five films: Central Intelligence (2016), Guardians of the Galaxy Vol. 2 (2017), Rogue One: A Star Wars Story (2016) featuring deepfakes of both Carrie Fisher and Peter Cushing, and The Irishman (2019). The Dark Knight (2008) was added to serve as a control (no deepfakes) and was included as a standalone reference of human authenticity in Hollywood film trailers to allow a comparison of results; therefore, its 100 users and 65 participants were not counted in the overall totals or used for sentiment analysis.

The 100 primary comments from users located specifically in North America (by utilizing a Virtual Private Network (VPN) to alter the researchers' location and gain access to YouTube's North American region) from each of the selected film trailer to gather audience sentiment data. Usernames were anonymized and their comments categorized based on their perception of the deepfakes' authenticity using a modified authenticity framework. Comments were then coded with "+" for validation, "-" for invalidation, and left blank for those not directly related to the deepfakes.

The researchers employed a specific coding method to analyze the data. This method, called Initial Coding, is a core principle of Constructivist Grounded Theory (CGT). It involves breaking down the comments into smaller parts and assigning codes based on their content. Each new piece of data is examined for potential new codes, allowing the coding scheme to evolve as the analysis progresses. Specifically, the researchers identified keywords that expressed either validating or invalidating opinions towards the deepfakes. They then calculated the proportion of comments falling into each category. This initial coding ensured a close connection between the raw data and the analysis, laying the groundwork for further exploration. It's important to note that comments could be coded into multiple categories, but comments that lacked any explicit validation or invalidation of the deepfakes were excluded. For instance, comments praising the use of deepfakes without mentioning reasons were not categorized. This resulted in a final set of 289 comments out of the initial 500 scraped comments.

The next step of the research involved in-depth interviews. Researchers interviewed 50 alumni students from Rangsit University, specifically targeting those who had watched at least 3 Hollywood films. These interviews aimed to validate or challenge the same authenticity aspects identified in the comment analysis. To gain deeper insights, participants answered a series of Yes/No questions followed by open-ended discussions to explore their thought processes and reasoning. Interestingly, to ensure participant comfort, the researchers opted for typed summaries of interview responses instead of recordings. This choice reflects research by Naunheim et al. (2023) suggesting many people dislike being recorded. To maintain accuracy, these summaries were read back to participants for verification and correction. Then the interview data was compared with the comment analysis. The researchers were particularly interested in uncovering any discrepancies in audience perception between viewing authentic and deepfaked content.Specifically, they wanted to see if the audience's validation or invalidation of specific authenticity elements changed based on their awareness of the deepfakes. Additionally, they compared how these authenticity measurements differed between audiences who believed they were watching real actors and those who knew the actors were deepfakes.

Research Findings

1.   The data scraping and analysis results

The results of the anonymized user comments were marked with '+' for validation, '-' for invalidation, and left blank for non-participants and were categorized into six authenticity components (Nunes et al., 2021 : 1-20):

Accuracy: Perception of Truthful Representation - Both humans and deepfakes need to be perceived as accurately portraying the character.

Connectedness: Engagement and Transformation - Audiences need to feel engaged and emotionally invested with the character, regardless of whether it's a human or a deepfake.

Integrity: Motivations and Quality - For humans, audiences need to believe the actor was the right choice for the role. For deepfakes, audiences shouldn't perceive financial gain compromising the performance quality.

Legitimacy: Alignment with Expectations - Both humans and deepfakes need to fit within the audience's expectations for the type of film and acting style.

Originality: Novelty and Uniqueness - For deepfakes, the audience should perceive a sense of novelty or a unique approach. For humans, originality relates to the performance itself, not the actor.

Proficiency: Skill and Expertise - Both humans and deepfakes need to appear believable and well-crafted. For audiences with no prior knowledge of deepfakes, the performance should be seamless.

According to Table 1, the sentiments extracted from the scraped comments data reveal that among the six constructs of authenticity, Connectedness received the highest validation, followed by Accuracy, Legitimacy, Proficiency, Originality, and Integrity, in descending order.

Table 1 The categorization of users' sentiments

Film Advert featuring deepfakes	Number of Uses	Components of authenticity											
		Accuracy		Connectedness		Integrity		Legitimacy		Originality		Proficiency	
Central Intelli.	13	+4 30.8%										+2 1.1.1%	
Ghostbusters of r.	45			+35 30%				+9 20%					
Rogue One: C.	55	+37 67.3%						−46 53.8%				−5 12.7%	+15 26.2%
Rogue One: P.	80	+35 41.7%	17 2.5%	+51 63.8%				−31 38.9%		+15 8.1%			+22 27.0%
The Irishman.	91	+5 20.8%		+55 55.7%		+1 1.1%		+3 3.3%				+2 2.2%	
Total	289	+111 38.4%	−17 5.6%	+150 42.2%	−6 6%	+1 0.3%	−6 2%	+84 28.6%	−8 4%	+15 5.2%	−6 4%	+11 3.8%	−37 12.8%
Control clip													
The Dark Kni.	83	+125 37.5%	−8 8%	+171 55.3%	−6 6%	+18 50%	−7 5.0%	N/A	N/A	+114 34.4%	−6 4%	+20 50.5%	−6 8%

Legend: [+] Validation. [−] Invalidation.
The control clip does not feature deepfakes nor count towards totals. It is a standalone reference of human authenticity for comparison.

2. The interview results

The interview results indicate that the participants validated the Originality of the deepfaked advertising of Hollywood films the most, followed by Connectedness, Integrity, Accuracy, Legitimacy, and Proficiency respectively (See Table 2).

Table 2 The interview results from the initial coding of 50 participant's sentiments

	Components of authenticity											
	Accuracy		Connectedness		Integrity		Legitimacy		Originality		Proficiency	
Validation	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
number	32	18	36	14	34	16	19	31	42	8	13	37
Percentage	64%	36%	72%	28%	68%	32%	38%	62%	84%	16%	26%	74%

In this final step, the data derived from the scraping and coding step was compared with the interview data to identify fulfilled components of authenticity and unveil underlying themes and connections. Intriguingly, the interview group exhibited a minimal 3.1 percent invalidation in Integrity, whereas scraped data revealed 5.7 percent and 12.4 percent invalidation in Accuracy and Proficiency, respectively. Interviews further elucidated varying invalidation rates, ranging from 16 percent for Originality to a substantial 74 percent for Proficiency.

The overall comparison results indicate that deepfakes were perceived as more authentic than not, aligning with the initial coding of scraped comments. The element of nostalgia, which is a common theme in scraped data, was not prominently reflected in the interview results, while opinions about the mixed quality of deepfakes and stronger emotional responses to deepfakes portraying deceased actors persisted.

Conclusion

The study uncovered nuanced insights that shed light on the complex relationship between audiences, deepfakes, and genuine actors. Firstly, deepfakes possess the capability to instill a sense of authenticity, particularly when they evoke emotional responses and pay homage to the legacies of original actors. However, the degree of effectiveness is contingent upon the quality of execution and the subjective perceptions of the audience. Interestingly, two aspects – Legitimacy and Proficiency – weren't validated by the interviewees, suggesting a more nuanced view of these elements. However, audiences seemed willing to overlook these concerns in favor of a generally positive experience, though perhaps not quite as authentic as with real actors.

Both scraped data and interviews affirmed the fulfillment of Accuracy and Connectedness. Although deepfaked characters aligned with storyline expectations, emphasizing their identifiability akin to human counterparts, the invalidating group expressed concerns about stiffness, animation, and lack of authentic emotion, suggesting potential advancements in deepfake technology could reduce this group. In terms of Integrity, the data comparison findings affirmed that deepfakes met this authenticity criterion. Audiences favored digital representations of deceased or de-aged actors, with concerns about proficiency not prominently recognized. The invalidating group's sentiment was largely a general dislike for deepfakes, indicating potential persistence unless Hollywood addresses proficiency and disclosure concerns.

Meanwhile, the interview group did not perceive deepfakes as legitimate, but there could be a temporal shift in this perception as deepfake technology becomes more commonplace globally. Invalidating sentiments in Legitimacy often described the usual, weird, or uncanny nature of deepfakes. Audiences also perceived deepfakes to be original, as expressed with the sentiments of novelty, uniqueness, and innovation among validating members. However, the sustainability of this component in ensuring authenticity may hinge on continuous innovation. Invalidating sentiments often expressed encountering similar phenomena previously. Finally, both scraped data and interviews concurred that deepfakes failed to meet the Proficiency criterion due to quality issues and conspicuous artificial elements, indicating potential immediate reduction in the invalidating group. However, the persistence of this group is speculated, attributing it to Hollywood's reluctance to adopt the latest deepfake technology despite access and budgetary capabilities. While genuine actors consistently provided overwhelmingly authentic experiences, deepfakes presented a more authentic-than-not encounter.

Despite concerns related to technical flaws or ethical considerations surrounding deepfake technology, audiences consistently prioritized their connection to the characters, irrespective of the medium used for portrayal. The findings underscore a noteworthy phenomenon—audiences form deep connections with characters, regardless of whether they are portrayed by deepfakes or authentic actors.

Discussion

Based on the results of this research, there are two key aspects of deepfakes in film that the researchers would like to discuss: Respecting Legacy and Audience Connection, and Balancing Innovation and Ethics.

(1) Respecting Legacy and Audience Connection

The study explored the complex relationship between deepfakes and audience perception in film. While a positive correlation emerged between deepfakes' effectiveness and respect for the original actor's legacy (Facia, 2023), it's important to note this doesn't imply causation. The importance of clear disclosure was paramount, especially when portraying deceased actors. This finding underscores the need for respectful handling of an actor's legacy. Additionally, the study suggests audiences may prioritize their emotional connection to characters over the technology used to portray them. This was evident in the positive audience reaction to the late Paul Walker's digitally recreated scenes in Furious 7 (2015), despite the use of AI and CGI. However, ethical concerns remain. The use of deepfakes for posthumous performances raises questions about image ownership and consent. While audience opinion on deepfakes is divided, there's a general concern about misuse and societal disruption.

(2) Balancing Innovation and Ethics

Our study emphasizes the need for responsible and thoughtful integration of deepfakes in film. Striking a delicate balance between innovation and ethical considerations is crucial. The findings suggest audiences may prioritize character connection over differentiating between deepfakes and human actors. Despite varying opinions on deepfake realism, a strong emotional connection to characters was consistently reported. This could lead to further ethical considerations when using deepfakes to recreate deceased actors, potentially evoking stronger emotional responses. However, a study found that our society is unprepared for the emergence of deepfakes at any level and the significant consequences have yet to materialize owing to the nascent stage of development in deepfake technology (Karnouskos, 2020 : 138-147). By maintaining a balance between technological advancements and ethical considerations, we can harness the potential of deepfakes while mitigating potential risks and safeguarding the integrity of cinematic storytelling.

Limitations

This study acknowledges some limitations that warrant further exploration. The data which primarily came from Rangsit University alumni and YouTube comments in North America may limit the applicability of findings to a broader, more diverse audience. Cultural contexts and values can significantly influence perceptions of authenticity, so audience interpretations might differ across regions.

There is also a limitation in technological evolution. Focusing on current deepfake technology may not capture future advancements or variations in quality. This could potentially affect the accuracy of the long-term implications.

In addition, the study's focus on Hollywood film advertising, rather than the entire film viewing experience, might limit the generalizability of results, while the specific films chosen may not represent the full spectrum of genres or approaches to deepfakes within the film industry.

Suggestions

Future Directions : this study lays the groundwork for further exploration. Expanding data collection beyond Rangsit University alumni to a wider demographic and geographic range would provide a more comprehensive picture of audience perception. Additionally, revisiting this study in the future could reveal how evolving deepfake technology shapes audience reactions. Future research could also examine how deepfakes influence viewers within the context of a full film experience and across different genres. Furthermore, the ethical implications of deepfakes warrant urgent investigation by researchers and policymakers.

Impact and Conclusion : this research empowers viewers by fostering a deeper understanding of how deepfakes can influence their perception of authenticity. It encourages a more critical viewing experience, prompting viewers to question what they see on screen. Recognizing both the potential for emotional connection and the limitations in achieving true authenticity is crucial. This awareness empowers all stakeholders, including creators and consumers, to work towards a future where deepfakes are used responsibly and effectively in storytelling.

<div align="center">

References

</div>

Abdulreda, A., & Obaid, A. (2022). A landscape view of deepfake techniques and detection methods. *International Journal of Nonlinear Analysis and Applications*, 13(1), 745-755. https://doi.org/10.22075/ijnaa.2022.5580

Bode, L., Lees, D., & Golding, D. (2021). The digital face and deepfakes on screen. *Convergence,* 27(4), 849–854. https://doi.org/10.1177/13548565211034044

Facia. (2023). *How did Paul Walker appear in Fast 7 after his death?* .https://www.linkedin.com/pulse/how-did-paul-walker-appear-fast-7-after-his-death-faciaai-cdqrf/

Goodfellow,I.J.,Pouget-Abadie,J.Milza,M.,Xu,B.,Waede-Farley,D.,Ozair,S.,Courville, A.and Bengio,Y.(2014). *Generative Adversarial Nets*.https://arxiv.org/pdf/1406.2661

Jin, X., Zhang, Z., Gao, B., Gao, S., Zhou, W., Yu, N., & Wang, G. (2023). *Assessing the perceived credibility of deepfakes: The impact of system-generated cues and video characteristics*. New Media & Society. Advance online publication. https://doi.org/10.1177/14614448231199664

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

Köbis, N., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience*, 24(11), 103364. https://doi.org/10.1016/j.isci.2021.103364

Lees, D. (2023). Deepfakes in documentary film production: Images of deception in the representation of the real. *Documentary Film*, 1–22. https://doi.org/10.1080/17503280.2023.2284680

Mahmud, B. U., & Sharmin, A. (2023). Deep insights of deepfake technology: A review (v2, revised January 2023). *DUJASE,* 5(1 & 2), 13-23. (January & July 2020). https://doi.org/10.48550/arXiv.2105.00192

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

Murphy, G., & Flynn, E. (2022). Deepfake false memories. *Memory,* 30(4), 480-492. https://doi.org/10.1080/09658211.2021.1919715

Nunes, J. C., Ordanini, A., & Giambastiani, G. (2021). The concept of authenticity: What it means to consumers. *Journal of Marketing,* 85(4), 1–20. https://doi.org/10.1177/0022242921997081

Oscars. (2009, February 26). *Heath Ledger wins best supporting actor for The Joker in "The Dark Knight" | 81st Oscars (2009)* [Video]. YouTube. https://www.youtube.com/watch?v=Jrt2xoy5UHo

Somers, M. (2020, July 21). *Deepfakes, explained.* MIT Sloan: https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

Warner Brothers Entertainment. (2023). *Aquaman and the Lost Kingdom*. https://www.aquamanmovie.net/

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review,* 9(11), 40-53. http://doi.org/10.22215/timreview/1282