Journal of Science Research International (JSRI)
ISSN: 2456-6365 (Online)

OPEN ACCESS

THE RISE OF DEEPFAKES: ANALYSING AWARENESS AND DETECTION CAPABILITIES OF AI-GENERATED MEDIA

Omkar Hari Sherkhane*, Arya Sawant,

Niraj Shinde, Aryan Mhatre and Roshni Mhatre

Department of Information Technology,

Pillai College of Arts, Commerce & Science (Autonomous), Navi Mumbai, Maharashtra, India 410206

**Corresponding author E-mail: omkarsherkhane@mes.ac.in*

Received: 25 December 2025	Revised: 19 January 2026	Accepted: 21 February 2026	Published: 28 February 2026

DOI: **https://doi.org/10.5281/zenodo.19047410**

Abstract:

Deepfake technology, powered by artificial intelligence, has transformed digital content creation by producing highly realistic synthetic images, videos, and audio. While these tools offer creative opportunities, their misuse threatens privacy, democracy, and online trust. This study investigates how individuals perceive deepfakes, their confidence in detecting manipulated media, and their views on societal risks. Based on survey responses from 51 participants, findings reveal that awareness of deepfakes is relatively high, but detection confidence remains moderate. Respondents widely acknowledge deepfakes as a serious societal issue and strongly support the development of automated detection systems. The research highlights the gap between theoretical awareness and practical detection skills, emphasizing the need for digital literacy, regulatory measures, and advanced AI-driven solutions.

Keywords: *Artificial Intelligence, Media Manipulation, Detection Confidence, Misinformation, Cybersecurity.*

Introduction

Artificial intelligence has revolutionized digital media, with deepfake technology emerging as one of its most striking applications. By using advanced techniques such as face-swapping and voice synthesis, deepfakes can generate fabricated yet convincing content. While these innovations have legitimate uses in entertainment and creative industries, their misuse raises concerns about misinformation, fraud, and reputational harm. In today's social media-driven environment, manipulated content spreads rapidly, influencing public opinion and eroding trust in authentic information. This study explores awareness levels, exposure, confidence in detection, and perceptions of societal impact, underscoring the importance of literacy and reliable detection mechanisms (1).

Literature review

Deepfakes are largely enabled by Generative Adversarial Networks (GANs), introduced by Goodfellow *et al.* (2). Scholars have examined both risks and detection strategies. Chesney and Citron (2019) discussed implications for privacy and democracy. Li and Lyu (3) proposed detection methods based on visual anomalies such as unnatural blinking. Verdoliva (4) reviewed forensic and deep learning approaches, while Westerlund (2019) emphasized the importance of public awareness. Most prior studies have focused on technical detection, leaving a gap in understanding how the public perceives and responds to deepfakes. This research addresses that gap.

Methodology

This study employed a descriptive, quantitative design. Data was collected through an online survey using Google Forms.

- **Sample size:** 51 respondents
- **Sampling method:** Convenience sampling
- **Data type:** Primary data

The survey measured awareness, exposure, confidence in detection, perceptions of societal impact, and support for AI-based detection systems.

Hypotheses

- **H1:** Most respondents are aware of deepfake technology.
- **H2:** Respondents show moderate confidence in identifying deepfakes.
- **H3:** Deepfakes are perceived as a serious societal issue.
- **H4:** There is strong support for automated detection tools.

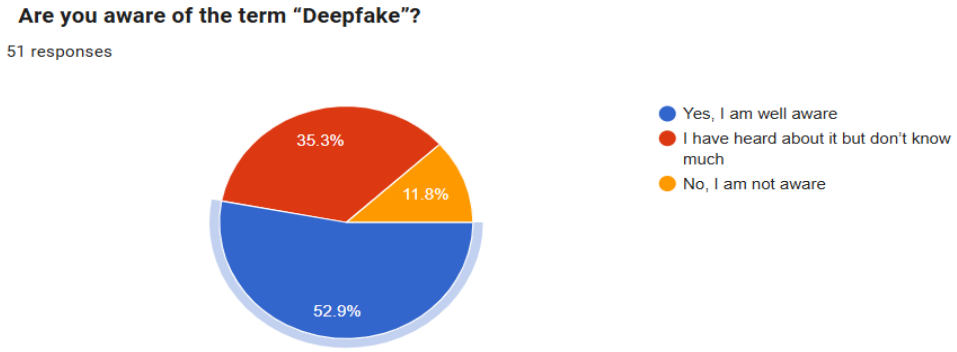
Analysis

H1: A Majority of Individuals Are Aware of Deepfake Technology

Supporting Question:

Q1: Are you aware of the term "Deepfake"?

More than half of the respondents (52.9%) stated that they are well aware of the term "Deepfake," while only a small percentage reported no awareness.



Interpretation:

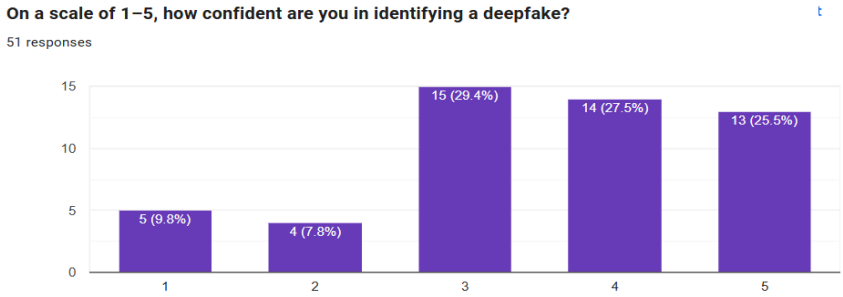
This result confirms that deepfake technology is widely recognized among the participants. Since a clear majority demonstrate awareness, Hypothesis H1 is supported. The findings indicate that deepfake technology is no longer an unfamiliar concept among the public.

H2: Individuals Have Moderate Confidence in Identifying Deepfake Content

Supporting Question:

Q5: On a scale of 1–5, how confident are you in identifying a deepfake?

Most respondents selected level 3 (moderate confidence), indicating an average level of ability to identify manipulated media.



Interpretation:

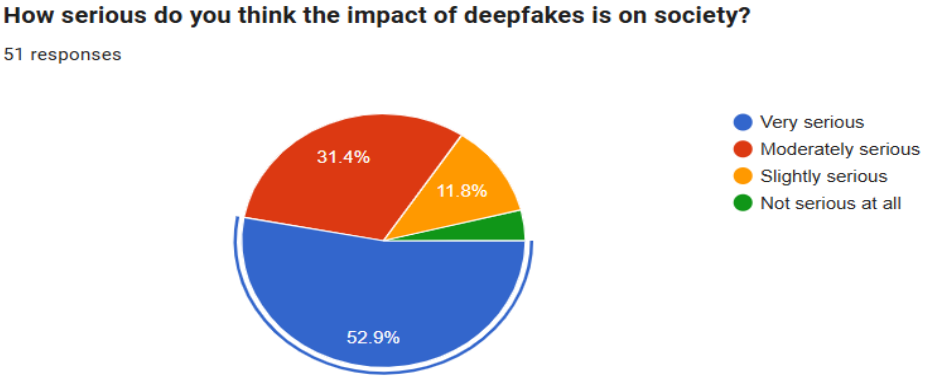
Although awareness is relatively high, confidence in detection is not strong. The majority rating themselves at a moderate level suggests limited detection capability. Therefore, Hypothesis H2 is supported, as respondents generally demonstrate moderate rather than high confidence.

H3: Deepfakes Are Perceived as a Serious Societal Issue

Supporting Question:

Q8: How serious do you think the impact of deepfakes is on society?

A majority of respondents (52.9%) rated the impact of deepfakes as very serious.



Interpretation:

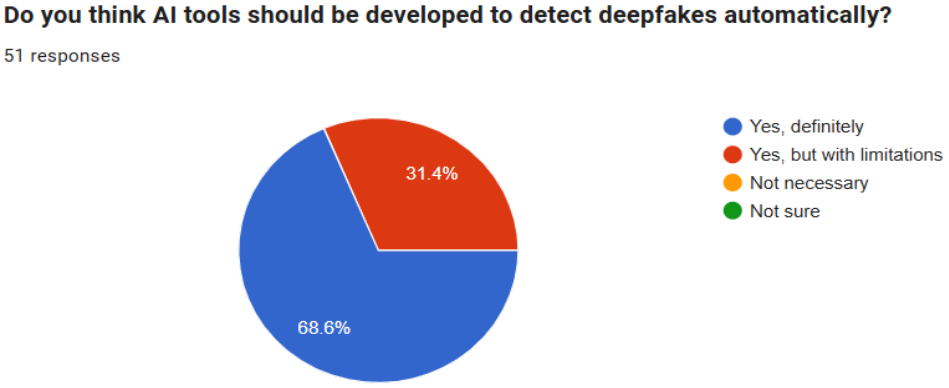
This indicates that participants recognize the significant risks associated with deepfake technology, including misinformation and cybercrime. Hence, **Hypothesis H3 is supported**, as deepfakes are widely perceived as a serious societal threat.

H4: There Is Strong Public Support for AI-Based Deepfake Detection Systems

Supporting Question:

Q10: Do you think AI tools should be developed to detect deepfakes automatically?

A strong majority (68.6%) responded positively, supporting the development of AI-based detection tools.

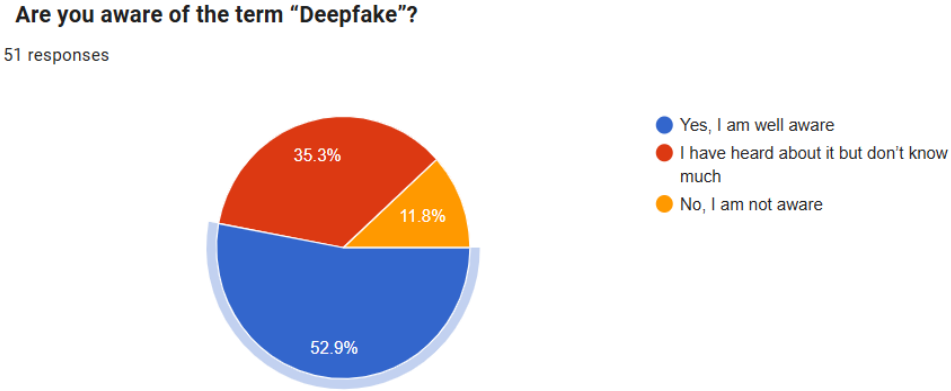


Interpretation:

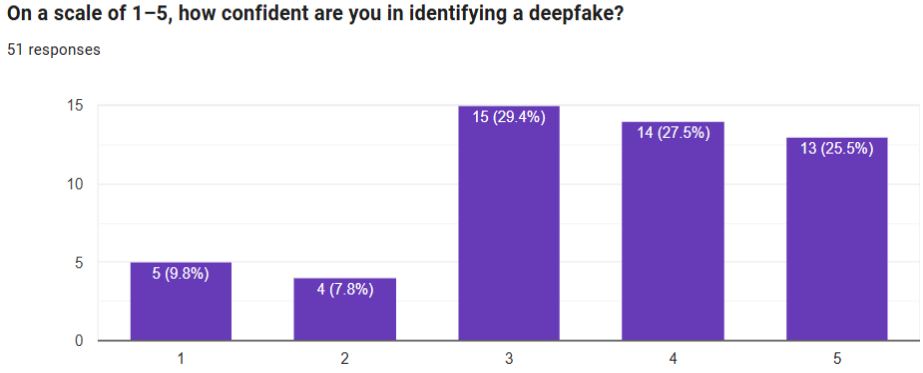
The high percentage of agreement demonstrates clear public demand for technological solutions to combat deepfake misuse. Therefore, Hypothesis H4 is strongly supported.

Results

H1: Awareness of Deepfakes Over half of respondents (52.9%) reported familiarity with the term "deepfake." *Interpretation:* Awareness of deepfake technology is widespread, supporting H1.



H2: Confidence in Detection Most participants rated their confidence at level 3 (moderate). *Interpretation:* While awareness is high, detection skills remain limited, supporting H2.



H3: Perception of Societal Impact A majority (52.9%) considered deepfakes a very serious issue. *Interpretation:* Respondents recognize the risks of misinformation and fraud, supporting H3.

H4: Support for Detection Tools 68.6% favoured the development of AI-based detection systems. *Interpretation:* Strong public demand exists for technological solutions, supporting H4.

All four hypotheses were supported: awareness is high, confidence is moderate, societal risks are acknowledged, and support for detection tools is strong.

Limitations

- The sample size of 51 respondents is relatively small, limiting the generalizability of findings.
- Convenience sampling may have introduced bias, as participants were not randomly selected.
- Self-reported confidence may not accurately reflect actual detection ability; experimental testing is needed.
- The study focused on awareness and perception, rather than measuring detection accuracy through controlled experiments.
- Cultural and demographic diversity was limited, which may affect how results apply to broader populations.

Discussion

The results of this study highlight a paradox: while awareness of deepfakes is relatively high, the ability to detect them remains modest. This gap is significant because awareness without detection skills may lead to misplaced confidence, leaving individuals vulnerable to manipulation. The findings align with prior research emphasizing the risks of misinformation and fraud, but they also reveal strong public demand for technological safeguards. Interestingly, respondents placed considerable trust in AI-based detection tools, suggesting that society views technology both as the source of the problem and as the solution.

Another important insight is the recognition of deepfakes as a serious societal issue. Participants expressed concern about their impact on politics, reputation, and cybersecurity. This perception reflects growing public anxiety about the erosion of trust in digital media. However, the reliance on automated detection raises questions about accountability: who should control these tools — governments, private companies, or independent watchdogs? Addressing this question is crucial for ensuring transparency and fairness in combating deepfake misuse.

Conclusion

This study demonstrates that while deepfake awareness is growing, detection skills remain limited. Respondents recognize deepfakes as a serious societal threat, particularly in spreading misinformation and enabling cybercrime, and they strongly support the development of AI-based detection systems. The findings highlight a critical gap between theoretical knowledge and practical detection ability. Bridging this gap requires a multi-pronged approach: advancing AI-driven detection tools, strengthening digital literacy, and implementing regulatory frameworks to ensure accountability. Future research should expand on these findings by testing detection accuracy experimentally and exploring global perspectives. By combining education, policy, and technology, society can better safeguard against the risks posed by deepfakes.

Future research directions

- Conduct large-scale studies with diverse participants to improve generalizability.
- Use experimental designs where respondents attempt to identify real vs. fake media clips, measuring actual detection accuracy.

- Explore cross-cultural differences in awareness and perception, as attitudes toward media manipulation may vary globally.
- Investigate the effectiveness of digital literacy programs in improving detection skills.
- Examine the ethical and regulatory implications of deploying AI-based detection tools, including issues of privacy, accountability, and bias.

Acknowledgement

The authors thank all survey participants for their valuable input. Appreciation is also extended to mentors and peers who provided guidance during the research process.

References

1. Chesney, R., & Citron, D. K. (2019). *Deep fakes: A looming challenge for privacy, democracy, and national security*. California Law Review, 107(6), 1753–1820.
2. Li, Y., & Lyu, S. (2018). *Exposing deepfake videos by detecting face warping artifacts*. Proceedings of the IEEE CVPR Workshops, 46–52.
3. Verdoliva, L. (2020). *Media forensics and deepfake detection: An overview*. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910–932.
4. Westerlund, M. (2019). *The emergence of deepfake technology: A review*. Technology Innovation Management Review.