

研究論文

# 眼見不一定為實：生成式人工智慧 深度偽造圖片的真實性認知研究

白龍<sup>ab</sup>、董飛飛<sup>c</sup>、馮迪拉<sup>d</sup>

<sup>a</sup> 東莞理工學院文學與傳媒學院，東莞市，中國大陸

<sup>b</sup> 南京師範大學新聞與傳播學院，南京市，中國大陸

<sup>c</sup> 馬來西亞國立大學人文與社會科學與傳播學院，萬宜新鎮，雪蘭莪州，馬來西亞

<sup>d</sup> 印尼阿赫瑪達蘭大學文化與傳播學院，日惹市，印度尼西亞

## 摘要

生成式人工智慧 (artificial intelligence, AI) 深度偽造技術創造出的智能欺騙，衝擊人類「有圖有真相」、「目知眼見」的真實性認知傳統。研究考察 229 位受訪者對深度偽造圖片的真實認知表現，發現識別 AI 深度偽造圖片的準確率僅有 39.3%。性別類型和 AI 技能經歷的不同，並不會導致甄別水準有顯著差異，但受訪者觀看文化接近性不同的視覺

白龍，東莞理工學院文學與傳媒學院副教授、南京師範大學新聞與傳播學院媒介發展與危機管理研究所研究員。研究興趣：新媒體與社會治理、智能傳播、新聞業務、媒介哲學。電郵：[bailong@dgut.edu.cn](mailto:bailong@dgut.edu.cn)

董飛飛，馬來西亞國立大學人文與社會科學與傳播學院博士候選人。研究興趣：健康傳播、資訊行為、社會資本、風險傳播、跨文化傳播。電郵：[p129446@siswa.ukm.edu.my](mailto:p129446@siswa.ukm.edu.my)

馮迪拉，印尼阿赫瑪達蘭大學文化與傳播學院傳播科學系副教授。研究興趣：社交媒體、新聞業務、政治傳播。電郵：[dani.fadillah@comm.uad.ac.id](mailto:dani.fadillah@comm.uad.ac.id)

論文投稿日期：2024 年 8 月 14 日。論文接受日期：2025 年 4 月 1 日。

《傳播與社會學刊》，(總)第75期(2026)

內容在辨識水準上有顯著差異。研究認為人工智慧生成內容 (artificial intelligence generated content, AIGC) 深度偽造技術已經超越了恐怖谷理論 (uncanny valley)<sup>1</sup> 的閾值，虛假線索資訊丟失導致受訪者陷入決策的隨機性，對合成圖像真實性認知的不確定性增強。如未來不能正視真實性危機，人類社會的真實認知將可能被顛覆，「資訊災變」和「現實冷漠」的社會風險就有可能到來。

關鍵詞：生成式人工智慧、深度偽造資訊、智能欺騙、虛假資訊

---

Research Article

## Seeing Is Not Always Believing: A Cognitive Study on the Authenticity of Deeply Falsified Images by Generative Artificial Intelligence

Long BAI<sup>ab</sup>, Feifei DONG<sup>c</sup>, Dira FENG<sup>d</sup>

<sup>a</sup> School of Literature and Media, Dongguan University of Technology, Dongguan, Mainland China

<sup>b</sup> School of Journalism and Communication, Nanjing Normal University, Nanjing, Mainland China

<sup>c</sup> Faculty of Social Sciences and Humanities, The National University of Malaysia, Bangi, Selangor, Malaysia

<sup>d</sup> Faculty of Culture and Communication, Ahmad Dahlan University, Indonesia, Kota Yogyakarta, Indonesia

---

### Abstract

Intelligent deception created by deep-fake technology in generative AI impacts the human tradition of “pictures do not lie” and “seeing is believing” perception of authenticity. The study evaluates the ability of 229 participants to discern the authenticity of deeply faked images and reveals that accuracy in

---

Long BAI (Associate Professor, Researcher). School of Literature and Media, Dongguan University of Technology, Institute of Media Development and Crisis Management, School of Journalism and Communication, Nanjing Normal University. Research interests: new media and social governance, intelligent communication, journalism practice, media philosophy. Email: bailong@dgut.edu.cn

Feifei DONG (Ph.D. Candidate). Faculty of Social Sciences and Humanities, The National University of Malaysia. Research interests: health communication, information behavior, social capital, risk communication, cross-cultural communication. Email: p129446@siswa.ukm.edu.my

Dira FENG (Associate Professor). Department of Communication Sciences, Faculty of Culture and Communication, Ahmad Dahlan University, Indonesia. Research interests: social media, journalism practice, political communication. Email: dani.fadillah@comm.uad.ac.id

Article History: Received on 14 August 2024. Accepted on 1 April 2025.

*Communication and Society*, 75 (2026)

recognizing AI-generated deepfakes stood at only 39.3%. Differences in gender and experience with AI skills did not lead to significant differences in screening levels, but respondents viewing visual content with varied cultural proximity differed significantly in recognition levels. The study concludes that generative AI's deep forgery technology has exceeded the threshold of the "uncanny valley," where the loss of information about false cues leads to randomness in decision-making and increased uncertainty about the perceived authenticity of synthesized images. If this authenticity crisis is not addressed in the future, the perception of truth in human society may be subverted, and the social risks of "information catastrophe" and "reality indifference" may become actualized.

**Keyword:** Generative artificial intelligence, deep-fake information, intelligent deception, false information