CREATIVE POWER, DECEPTIVE POTENTIAL: HOW IMAGE GENERATORS ARE RESHAPING VISUAL TRUTH

E.-C. GROSS[1]   T. PETREAN [2]

Abstract: *The rapid evolution of generative AI image models, such as Midjourney, DALL·E, and Google Gemini, has transformed visual communication by democratizing creativity while simultaneously amplifying risks of disinformation. This paper examines the dual nature of these technologies, highlighting their capacity to enhance artistic expression, accelerate creative workflows, and foster new forms of participation, while also enabling large-scale synthetic deception with political, social, and economic consequences. Through the analysis of case studies, industry practices, and regulatory frameworks, the study outlines both opportunities and threats generated by AI image tools.*

Keywords: *Generative AI; Deepfakes; Visual Communication; Disinformation; Media literacy*

1. Introduction

The emergence of sophisticated AI image generators has fundamentally altered our relationship with visual media. In 2023, the realism of AI-based images from tools like Midjourney, DALL-E, or Stable Diffusion reached such a high level that it led to a significant wave of viral AI-generated photos, including fabricated images of Pope Francis and Donald Trump that deceived millions of viewers (Momeni, 2024; Becker & Harwell, 2023). These systems, powered by diffusion models and transformer architectures, can produce photorealistic images from text descriptions in seconds, democratizing creative capabilities once reserved for skilled artists and photographers. Midjourney describes itself as "an independent research lab exploring new mediums of thought and expanding the imaginative powers of the human species," positioning these tools as creative amplifiers rather than mere technical utilities. Indeed, research by the Marketing AI Institute (2024) found that AI adoption among creative professionals has accelerated dramatically, with many reporting they "couldn't live without AI" in their daily workflows.

[1] *Lucian Blaga University of Sibiu*, eduard.gross@ulbsibiu.ro , corresponding author
[2] Center for Inclusive Education, Mediaş

However, generative AI plays a dual role: it enables the rapid creation and targeted dissemination of synthetic content that can be weaponized for disinformation campaigns (López-Borrull & Lopezosa, 2025).

The stakes are considerable. Between 2023 and 2024, there was a fourfold increase in the number of deepfakes detected in fraud, and in 2023 alone, an estimated 500,000 deepfakes were shared on social media (Chandra et al., 2025). This proliferation occurs against a backdrop of global elections, where AI models enable malicious actors to manipulate information and disrupt electoral processes, threatening democracies (Csernatoni, 2024). The challenge is particularly acute because deepfakes are often indistinguishable from authentic media, which has led to their misuse in spreading misinformation, impersonation, and other malicious activities (Amerini et al., 2025). This paper examines three leading image generation platforms—Midjourney, OpenAI's DALL-E, and Google's Gemini (formerly codenamed "Nano Banana")—in order to understand their technical capabilities, creative applications, and potential for misuse. We analyze documented cases of both beneficial and harmful applications, evaluate current mitigation strategies, and propose a framework for balancing innovation with information integrity. Our central argument is that while these technologies offer transformative creative potential, their deployment requires to be coordinated technical, regulatory, and educational interventions to prevent erosion of visual truth in democratic societies.

2. The Promise of Creativity

The rise of generative artificial intelligence (AI) image models such as Midjourney, OpenAI's DALL·E, and Google's Gemini has sparked an unprecedented transformation in the landscape of visual communication. For centuries, creative work has been tied to the mastery of tools and techniques, from the brush to the camera. Today, the diffusion model and transformer architecture stand alongside these earlier innovations, reshaping not only how images are produced but also how creativity itself is conceptualized. At its most optimistic, this technological shift represents a democratization of creative expression. Tools that once required years of technical training are now accessible to anyone with a keyboard and an internet connection, allowing users to translate text prompts into photorealistic or stylized images in seconds (Rombach et al., 2022).

In creative industries, this transformation has been embraced with remarkable enthusiasm. Adobe's *Creative Trends Report 2024* reported that creative professionals had integrated AI into their workflows, reporting time savings in early concept development stages (Adobe, 2024). This acceleration allows for a new mode of creativity characterized by iteration and abundance. Where traditional brainstorming required sketching, storyboarding, or hiring specialized artists, today's art directors and designers can generate dozens of concepts instantaneously, testing ideas against client expectations or audience preferences. The rise of "prompt engineering" as a creative skill—crafting the right textual inputs to coax a system into producing desired results—demonstrates that

linguistic precision and conceptual clarity now function as creative capital (Oppenlaender et al., 2024). Communities such as Midjourney's Discord server, with over 19 million members, have become laboratories of aesthetic experimentation, where users collectively develop what could be called a "synthetic visual culture."
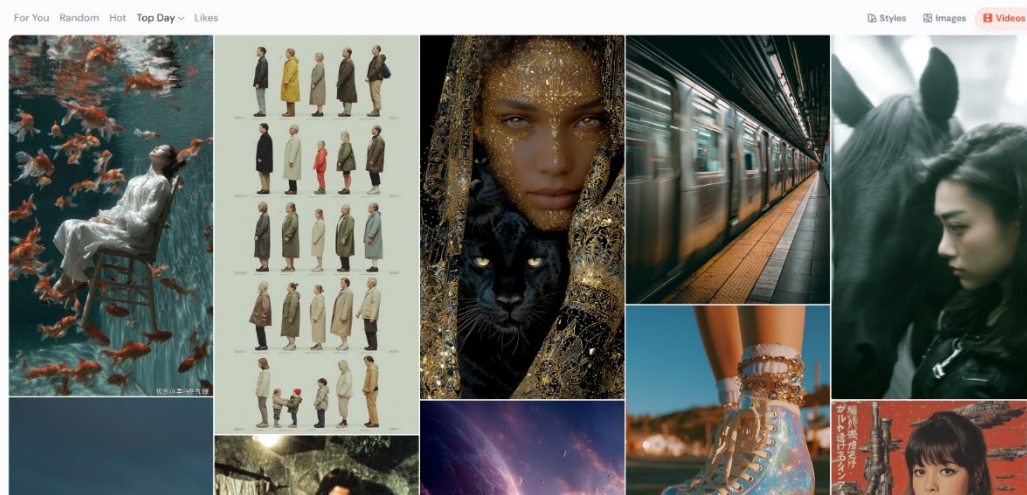

Fig. 1. *A screenshot from Midjourney pool of AI generated visuals*

The advertising industry offers some of the clearest illustrations of this shift. Heinz's "AI Ketchup" campaign in 2023 capitalized on the peculiar fact that when prompted with the concept of "ketchup," DALL·E generated images resembling Heinz bottles, reflecting brand dominance in the cultural imagination (Gross, 2023). Coca-Cola's "Create Real Magic" initiative invited consumers to generate AI art with official brand assets, turning customers into co-creators and harnessing the participatory ethos of digital culture (Swant, 2023). Such campaigns demonstrate that AI is not merely a tool for cost-cutting but a medium for new forms of engagement. The personalization capacity of generative AI—its ability to produce countless variants of an image tailored to different demographics—enables advertisers to iterate rapidly and target specific audiences with unprecedented precision (Gartner, 2024). The implications extend beyond advertising and design. Artists have used AI tools for style exploration, hybridization, and conceptual prototyping. The role of the artist is thus being reframed: not as a sole producer of forms but as a curator of algorithms, a negotiator between machine capacity and human intention. Critics point out that this shift challenges long-standing notions of originality and authorship, raising questions about whether works produced with AI assistance qualify as authentic or derivative (Hartmann et al., 2025). Intellectual property disputes, such as those brought by Getty Images and artist collectives against AI developers for using copyrighted training data without consent, underscore the contested legal terrain (Romero-Moreno, 2025). In parallel, aesthetic concerns are raised by the tendency of models to replicate dominant visual tropes and biases embedded in training datasets.

Yet despite these challenges, the promise of creativity remains one of interest. Generative AI allows for experimentation at scale, empowering those who might lack traditional artistic training to participate in visual culture. It lowers the cost of entry for small businesses and independent creators, who can now compete with large studios in producing compelling visuals. It fosters hybrid practices where AI-generated imagery becomes raw material for human refinement, much like photography once supplemented painting rather than replacing it. From this perspective, AI models can be seen as amplifiers of imagination, co-processors that extend human faculties into new domains of possibility (Kumar, 2025).

3. Deepfakes, Disinformation, and the Collapse of Visual Trust

The very qualities that make generative AI a boon for creativity—photorealism, personalization, and scalability—also make it a formidable tool for deception. The past two years have seen a surge in incidents that illustrate the disruptive power of synthetic media in political, social, and economic contexts. In 2023, AI-generated images of Pope Francis in a white Balenciaga-style puffer coat circulated widely, convincing millions of viewers of their authenticity (Becker & Harwell, 2023). That same year, fabricated photos of Donald Trump's arrest briefly dominated social media feeds, while a fake image of a Pentagon explosion triggered a temporary dip in stock markets (Bond, 2023). Analyzed as revelatory case studies (Yin, 2003), these incidents reveal how visual plausibility alone can generate large-scale credibility attribution, exposing a structural vulnerability in contemporary visual epistemology. Across the examined case studies, scholars describe this phenomenon as part of a larger 'post-truth' ecosystem, in which visual plausibility is decoupled from indexicality, and shared standards of evidence collapse (Momeni, 2024; Chesney & Citron, 2019). The "liar's dividend"—where the existence of deepfakes enables malicious actors to dismiss authentic evidence as fake—further complicates accountability. Politicians and corporations can deny documented events, claiming they are fabrications, while genuine forgeries spread unchecked.

Detection technologies, though advancing, remain insufficient. Traditional forensic methods based on identifying anatomical inconsistencies or lighting mismatches are increasingly ineffective against state-of-the-art models. Research in computer vision has turned toward multimodal detection, analyzing spatiotemporal signals and cross-modal coherence, but adversaries continually adapt (Khan et al., 2025). Benchmarks such as *Deepfake-Eval-2024* emphasize that detectors perform poorly on in-the-wild samples affected by compression, resizing, and platform-specific artifacts (Chandra et al., 2025). This arms race between generation and detection indicates that relying solely on forensic solutions is untenable (Amerini et al., 2025).

From a comparative case study perspective, platform governance reveals significant vulnerabilities resulting from divergent moderation policies across generative AI systems. While OpenAI enforces strict filters preventing DALL·E from generating images of public figures, Midjourney initially adopted a more permissive approach, only introducing

political figure bans after widespread misuse (Becker & Harwell, 2023). Google's Gemini embeds watermarks in its outputs, but such signals can be easily lost or manipulated through simple edits (U.S. GAO, 2024). The heterogeneity of policies across platforms allows malicious actors to gravitate toward the least restrictive systems, exploiting regulatory arbitrage.

At the regulatory level, responses vary widely. The European Union's *AI Act*, set to take effect in 2026, mandates machine-readable identification of AI-generated content and explicit labeling of political deepfakes. China imposes mandatory labeling of "deep synthesis" media, with penalties for non-compliance. In contrast, the United States regulation remains fragmented, with California's SB 942 requiring watermarking for large-scale generators while federal efforts rely largely on voluntary commitments (Romero-Moreno, 2025). These disparities reflect broader cultural and political differences but also highlight the absence of a harmonized international approach. Without cross-border standards, enforcement is easily circumvented in the global digital ecosystem.

The societal stakes of synthetic deception are significant. Democracies depend on the circulation of credible information and the maintenance of trust in institutions. As synthetic media blurs the line between real and fabricated, citizens' capacity to distinguish truth from fiction erodes. Research indicates that concerns about AI-driven misinformation are often heightened more by media coverage of AI than by direct exposure to AI tools themselves, creating a paradox in which heightened fear may itself contribute to vulnerability (Yan et al., 2025). In this sense, synthetic deception is not merely a technical problem but a socio-political one: it destabilizes the conditions of trust upon which public life depends.

4. Safeguarding Visual Truth: Governance, Standards, and Literacy

The comparative analysis of the examined cases indicates that reconciling creative potential with the risk of deception requires coordinated technical, regulatory, and educational interventions. This requires moving beyond technical fixes to a multi-stakeholder framework encompassing standards, governance, and literacy. At the technical level, watermarking and content provenance frameworks such as Google's SynthID and the Coalition for Content Provenance and Authenticity (C2PA) represent promising avenues. By embedding imperceptible signals or cryptographic chains of custody into digital media, such systems aim to provide verifiable records of creation and modification (Nadimpalli, & Rattani 2024a). Yet these measures are vulnerable to adversarial attacks such as cropping, compression, or screenshotting, raising doubts about their robustness (Nadimpalli & Rattani 2024b). To be effective, provenance standards must become universal defaults rather than optional add-ons, ensuring interoperability across platforms and accessibility to journalists, regulators, and end-users (U.S. GAO, 2024).

Regulatory frameworks must complement these technical measures. The EU's risk-

based approach offers a comprehensive model, but its success will depend on enforcement and cross-border recognition. In the absence of global coordination, fragmented regimes risk leaving loopholes for malicious actors. A proportionate and adaptive legal strategy is essential: one that distinguishes between artistic experimentation, commercial advertising, and political communication, imposing stricter disclosure and provenance requirements on high-risk uses while preserving space for creative freedom.

Equally important is the cultivation of social resilience through education and media literacy. Traditional literacy frameworks, focused on evaluating sources and cross-referencing claims, must evolve to account for the new reality in which any image could plausibly be synthetic. Journalists, educators, and policymakers require training in "visual forensics," combining technical verification tools with contextual analysis. For the public, media literacy must shift from "spotting flaws" to adopting a more skeptical orientation toward all visual media, encouraging verification habits such as reverse image searches and provenance checks. By embedding these practices into educational curricula and professional standards, societies can foster critical engagement without succumbing to cynicism.

Ultimately, balancing innovation and integrity demands a coordinated effort. Creative industries should codify ethical standards for AI use, including disclosure of synthetic imagery, consent mechanisms for likeness replication, and bias auditing. Policymakers should incentivize adoption of provenance standards through procurement policies and liability frameworks. Technology companies must prioritize safety-by-design, implementing auditable logs, tiered access to high-risk capabilities, and robust red-teaming. Civil society organizations and journalists must hold platforms accountable and advocate transparency. Only through such collective action can societies safeguard informational integrity while preserving the immense creative potential of generative AI.

The duality of generative AI is unlikely to disappear. On one hand, it empowers unprecedented creative abundance, opening avenues for experimentation, participation, and efficiency. On the other, it destabilizes the very epistemic foundations of democracy by eroding visual truth. The lesson of recent years is that these two dimensions are inseparable: the same features that enable one also to enable the other. As such, any sustainable response must treat creativity and deception not as separate phenomena but as twin outcomes of the same technological affordances. Technical detection will always be a step ahead of the generative approach, but standards, governance, and literacy can raise the costs of deception and lower the friction of verification.

In conclusion, generative AI represents both the most powerful creative instrument of the digital era and the most insidious threat to informational integrity. The question is not whether societies can embrace one without the other but whether they can manage both simultaneously. The window for effective governance narrows daily as adoption accelerates and capabilities advance. The future of visual truth—and by extension, the resilience of democratic societies—depends on the choices made now to balance creative power with protective safeguards.

5. Conclusion

Generative AI image technologies represent one of the most disruptive innovations of the digital era, simultaneously expanding creative horizons and challenging the foundations of visual truth. On the positive side, platforms such as Midjourney, DALL·E, and Google Gemini democratize access to creative tools, reduce production costs, and stimulate artistic experimentation and cultural participation. These systems enable professionals and non-specialists alike to engage in new forms of visual expression and accelerate workflows across industries from advertising to education. At the same time, however, their potential for harm is considerable: AI-generated images can be weaponized for disinformation campaigns, impersonation, electoral manipulation, and the erosion of trust in visual media.

Despite advances in detection technologies and watermarking solutions, the arms race between generation and verification continues to tilt in favor of synthetic content. Fragmented regulations and inconsistent platform governance create exploitable loopholes, while public awareness and media literacy often lag the speed of technological adoption. This paradox underscores the necessity of a holistic response that balances innovation with responsibility, combining technical safeguards, harmonized legal frameworks, ethical guidelines for creative industries, and critical education at all levels of society.

Nevertheless, this study has certain limitations. First, it relies primarily on documented cases and secondary literature, which may not capture the full spectrum of emerging applications and threats. Second, the analysis is limited to three major platforms, while numerous other AI tools, some open-source and less regulated, play a significant role in the diffusion of synthetic media. Third, the rapidly evolving nature of the field means that regulatory and technical countermeasures discussed here may become outdated in a short time frame. Finally, cultural and regional variations in the adoption of generative AI are only partially addressed, leaving room for more localized research.

Future research should explore several directions. Empirical studies are needed to measure the societal impact of AI-generated imagery on trust, perception, and behavior, particularly in high-stakes contexts such as elections, crises, and legal proceedings. Comparative analyses of regulatory frameworks across different jurisdictions could illuminate best practices and pathways toward international harmonization. Interdisciplinary collaboration between computer science, law, communication, and psychology will be essential to develop robust detection methods and effective educational interventions. Additionally, research into the ethical dimensions of authorship, originality, and bias in AI-generated art could further clarify the cultural and legal implications of this technological shift.

In sum, generative AI is both a powerful creative amplifier and a destabilizing force for informational integrity. The challenge is not to choose between these dimensions but to acknowledge their interdependence and to design governance systems that

preserve the benefits of creativity while minimizing the risks of deception. The sustainability of visual truth, and by extension the resilience of democratic societies, depends on the capacity to act swiftly and collaboratively in the face of this technological transformation.

References

Adobe. (2024). *Creative trends report 2024: The impact of AI on creative workflows*. Adobe Inc. https://www.adobe.com/content/dam/cct/creativecloud/creative-trends-lp/2024_creativetrends/FY23Q4_CC_Stock_Stock_XY_EN_2024CreativeTrendsGuide_Reduced.pdf.

Amerini, I., Barni, M., Battiato, S., Bestagini, P., Boato, G., Bruni, V., Caldelli, R., De Natale, F., De Nicola, R., Guarnera, L., Mandelli, S., Majid, T., Marcialis, G. L., Micheletto, M., Montibeller, A., Orrù, G., Ortis, A., Perazzo, P., Puglisi, G., … Vitulano, D. (2025). Deepfake Media Forensics: Status and future challenges. *Journal of Imaging*, *11*(3), 73. https://doi.org/10.3390/jimaging11030073

Amerini, I., Barni, M., Battiato, S., Bestagini, P., Boato, G., Bruni, V., Caldelli, R., De Natale, F., De Nicola, R., Guarnera, L., Mandelli, S., Majid, T., Marcialis, G. L., Micheletto, M., Montibeller, A., Orrù, G., Ortis, A., Perazzo, P., Puglisi, G., … Vitulano, D. (2025). Deepfake Media Forensics: Status and future challenges. *Journal of Imaging*, *11*(3), 73. https://doi.org/10.3390/jimaging11030073

Becker, I. S., & Harwell, D. (2023). *Midjourney is making fake images go mainstream - The Washington Post*. Washington Post. https://www.washingtonpost.com/technology/2023/03/30/midjourney-ai-image-generation-rules/

Bond, S. (2023, May 22). *Fake viral images of an explosion at the Pentagon were probably created by ai*. NPR. https://www.npr.org/2023/05/22/1177590231/fake-viral-images-of-an-explosion-at-the-pentagon-were-probably-created-by-ai

Chandra, N. A., Murtfeldt, R., Qiu, L., Karmakar, A., Lee, H., Tanumihardja, E., ... & Etzioni, O. (2025). Deepfake-eval-2024: A multi-modal in-the-wild benchmark of deepfakes circulated in 2024. *arXiv preprint arXiv:2503.02857*.

Chesney, Robert, & Keats Citron, D. ((July 14, 2018). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *107 California Law Review 1753* (2019), *U of Texas Law, Public Law Research Paper No. 692*, *U of Maryland Legal Studies Research Paper No. 2018-21.* Available at SSRN: https://ssrn.com/abstract=3213954 or http://dx.doi.org/10.2139/ssrn.3213954

Csernatoni, R. (2024). *Can democracy survive the disruptive power of AI?*. Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2024/12/can-democracy-survive-the-disruptive-power-of-ai

Gartner. (2024). *Generative AI 2024 planning survey*. Gartner Inc. https://www.gartner.com/peer-community/oneminuteinsights/omi-2024-generative-ai-planning-how-it-organizations-preparing-zxm

Gross, E. (March 05, 2023). *When Heinz asked A.I. for a Ketchup Commercial.* Available at SSRN: https://ssrn.com/abstract=5376538 or http://dx.doi.org/10.2139/ssrn.5376538

Hartmann, J., Exner, Y., & Domdey, S. (2025). The power of Generative Marketing: Can generative AI create Superhuman Visual Marketing Content? *International Journal of Research in Marketing*, *42*(1), 13–31. https://doi.org/10.1016/j.ijresmar.2024.09.002

Khan, A. A., Laghari, A. A., Inam, S. A., Ullah, S., Shahzad, M., & Syed, D. (2025). A survey on multimedia-enabled Deepfake Detection: State-of-the-art tools and techniques, emerging trends, current challenges & Limitations, and Future Directions. *Discover Computing*, *28*(1). https://doi.org/10.1007/s10791-025-09550-0

Kumar, S. A. (2025). *The democratization of AI Image Generation*. USENIX. https://www.usenix.org/publications/loginonline/democratization-ai-image-generation

López-Borrull, A., & Lopezosa, C. (2025). Mapping the impact of Generative AI on disinformation: Insights from a scoping review. *Publications*, *13*(3), 33. https://doi.org/10.3390/publications13030033

Marketing AI Institute. (2024). *2024 state of marketing AI report*. Marketing AI Institute. https://www.marketingaiinstitute.com/2024-state-of-marketing-ai-report

Miclea, M., Momeni, M. (2024). Artificial Intelligence and Political Deepfakes: Shaping Citizen Perceptions Through Misinformation. *Journal of Creative Communications*, *20*(1), 41-56. https://doi.org/10.1177/09732586241277335 (Original work published 2025)

Nadimpalli, A. V., & Rattani, A. (2024a). Proactive deepfake detection using gan-based visible watermarking. *ACM Transactions on Multimedia Computing, Communications and Applications*, *20*(11), 1-27.

Nadimpalli, A. V., & Rattani, A. (2024b). Social media authentication and combating deepfakes using semi-fragile invisible image watermarking. *Digital Threats: Research and Practice*, *5*(4), 1-30.

Oppenlaender, J., Linder, R., & Silvennoinen, J. (2024). Prompting AI art: An investigation into the creative skill of prompt engineering. *International Journal of Human–Computer Interaction*, *41*(16), 10207–10229. https://doi.org/10.1080/10447318.2024.2431761

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684-10695).

Romero-Moreno, F. (2025). Deepfake detection in generative AI: A Legal Framework proposal to protect human rights. *Computer Law & Security Review*, *58*, 106162. https://doi.org/10.1016/j.clsr.2025.106162

Swant, M. (2023, December 18). *AI briefing: Creative Ways Companies used AI in advertising in 2023*. Digiday. https://digiday.com/media-buying/ai-briefing-creative-ways-companies-used-ai-in-advertising-in-2023/

U.S. Government Accountability Office. (2024). *Science & Tech Spotlight: Combating deepfakes* (GAO-24-107292). https://www.gao.gov/products/gao-24-107292

Yan, H. Y., Morrow, G., Yang, K.-C., & Wihbey, J. (2025)*. The origin of public concerns over AI supercharging misinformation in the 2024 U.S. presidential election. Harvard Kennedy School (HKS*). *Misinformation Review.* https://doi.org/10.37016/mr-2020-171

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Sage Publications.

Other information may be obtained from the address: eduard.gross@ulbsibiu.ro