

Unnatural Images: On AI-Generated Photographs

Amanda Wasielewski

In December 2018, researchers from the technology company Nvidia revealed that they had developed a machine learning architecture called StyleGAN that could automatically generate photorealistic images of human faces. Their paper was published the following spring along with the source code repository to implement it.¹ Shortly thereafter, software engineer Philip Wang created the viral website This Person Does Not Exist (thispersondoesnotexist.com), which uses StyleGAN to provide a new artificial-intelligence-generated image of a human face with each page refresh (fig. 1). According to Wang, it was designed to demonstrate the capabilities of the technology and make these developments accessible to a wider public.² In subsequent years, interest in, awareness of, and concern regarding AI-generated images have continued to grow.

In 2022 more than half a dozen text-to-image generators were released to the public, providing an easy and accessible means for anyone to generate images with AI. Commercial text-to-image generators like DALL-E, Stable

Funding for this research was provided by the Swedish Research Council (grant VR 2018-06057) and Google's Artists + Machine Intelligence (AMI) program. I would like to express my gratitude for the immensely helpful suggestions and feedback from the reviewers, editors, and editorial board of *Critical Inquiry* and to Anna Näslund for supporting the early stages of this research.

1. See Tero Karras, Samuli Laine, and Timo Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Los Alamitos, Calif., 2019), pp. 4396–405. See also Karras and Janne Hellsten, "NVlabs/Stylegan," Python library, 2019, github.com/NVLabs/stylegan

2. See Danny Paez, "This Person Does Not Exist' Creator Reveals His Site's Creepy Origin Story," *Inverse*, 21 Feb. 2019, www.inverse.com/article/53414-this-person-does-not-exist-creator-interview



FIGURE 1. AI-generated portrait, StyleGAN/thispersondoesnotexist.com, June 2022.

Diffusion, and Midjourney allow users to depict almost anything imaginable (or, more accurately, anything *describable*) with a carefully constructed string of text known as a prompt. The basis for these text-to-image generators is an innovative model called CLIP (Contrastive Language–Image Pretraining), which connects language and images.³ The publication of CLIP shook the

3. See Aditya Ramesh et al., “Hierarchical Text-Conditional Image Generation with CLIP Latents,” *arXiv*, 13 Apr. 2022, doi.org/10.48550/arXiv.2204.06125

AMANDA WASIELEWSKI is an associate senior lecturer of digital humanities and associate professor of art history in the Department of ALM (Archives, Libraries, Museums), Uppsala University, Sweden. She is the author of three monographs including, most recently, *Computational Formalism: Art History and Machine Learning* (2023) and coeditor of the anthology *Critical Digital Art History: Interface and Data Politics in the Post-Digital Era* (2024).

field of AI because it provided a powerful way to achieve multimodality. This means that generative AI can now synthesize different forms of data and thus different representational forms into intertwined conceptual pairings. Coupled with an image-creation machine learning model known as a diffusion model, many of them are able to generate images that look like photographs (fig. 2). AI-generated images that are produced using diffusion models can be enhanced using generative-adversarial-network- based tools to create a variety of photorealistic images of people, places, and things that “do not exist.” For example, GANs are regularly used to clean up, sharpen, or alter the faces in images produced by text-to-image generators, creating a more convincingly or seamless photographic appearance (fig. 3).



FIGURE 2. AI-generated portrait, DALL-E 2, (prompt: “hyperrealism, 35mm film photography, portrait of a woman, shallow depth of field”), August 2022.



FIGURE 3. AI-generated image, DALL-E 2 (prompt: “hyperrealism, 35mm film photography, club photography, woman dancing, hipster”) unedited and edited using ARC Lab Face Restoration, August 2022.

Indeed, it has become a popular practice to prompt text-to-image generators using phrases or words associated with photographic techniques. For example, AI prompts commonly specify film types, lens techniques, or camera equipment such as “35mm film,” “Sigma 85 mm f/1.4” or “Nikon D810.” Using words such as these can help replicate the visual qualities produced by different photographic techniques and equipment in an AI-generated image because photographic images in the training data have been paired with such descriptors in online text accompanying them or through textual tagging/metadata. Viewers of these AI-generated images thus often interpret them as photographic images based on their visual form.

The distinctive yet often overlooked visual qualities of photography are created through the physics of the materials and equipment used. These formal qualities are specific to different photographic techniques and thus communicate photographicness for the viewer in a variety of ways. When these visual effects are present in an AI-produced image, they give that image both a specific visual appearance and the conceptual associations attendant to it. Chief among these associations, for photography, is the perception that the subject of the image exists or existed somewhere, someplace. Culturally speaking, photographic form (even if synthetically produced) communicates as photographs do—with all that entails.

Given the vast array of potential uses for generative AI, it is unsurprising that these technologies have been met with both excitement and fear. The fact that images used to train popular text-to-image models were scraped from the web and deployed without the permission of artists or the people depicted in them has led to emerging legal disputes and concerns regarding

privacy.⁴ Meanwhile, the ability to easily create fabricated individuals or to manipulate images, audio, or video of public figures in unflattering ways (commonly referred to as deepfakes) has raised alarm, particularly in the realm of politics.⁵ AI-generated portraits are already often used as profile pictures for bots and sock puppet accounts on social media, including those that seek to foment political dissent.⁶ Analysis of generative AI in both the press and academia has, therefore, largely focused on the urgent political and social implications of its use.⁷ As the headline of a 2020 *New York Times* article on the topic suggests, these images seem to have been “designed to deceive.”⁸

The ability of certain AI-generated images to deceive the viewer comes down to the aforementioned replication of photographic form, a technical feat that further complicates the already-complicated relationship between photography and nature. The idea that photography is an objective, mechanical procedure that faithfully renders the natural world in all its incidental detail is the ur-theory of the medium. William Henry Fox Talbot, who developed one of the earliest forms of photography, famously titled his photo book *The Pencil of Nature* (1844). According to art historian John Tagg, “The value of the camera was extolled because the optical and chemical processes of photography were taken to designate a scientifically exploited but

4. See Chloe Xiang, “AI Is Probably Using Your Images and It’s Not Easy to Opt Out,” *Vice*, 26 Sept. 2022, www.vice.com/en/article/3ad58k/ai-is-probably-using-your-images-and-its-not-easy-to-opt-out. See also James Vincent, “AI Art Tools Stable Diffusion and Midjourney Targeted with Copyright Lawsuit,” *Verge*, 16 Jan. 2023, www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart and “Getty Images Is Suing the Creators of AI Art Tool Stable Diffusion for Scraping Its Content,” *Verge*, 17 Jan. 2023, www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit

5. See Tiffany Hsu, “As Deepfakes Flourish, Countries Struggle with Response,” *New York Times*, 22 Jan. 2023, www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html

6. See Mitchell Clark, “Researchers Caught a Pro-US Campaign Spreading Propaganda on Social Media,” *Verge*, 25 Aug. 2022, www.theverge.com/2022/8/25/23322214/us-government-propaganda-campaign-twitter-facebook

7. See, for example, Thomas Paterson and Lauren Hanley, “Political Warfare in the Digital Age: Cyber Subversion, Information Operations and ‘Deep Fakes,’” *Australian Journal of International Affairs* 74 (July 2020): 439–54; Chandell Gosse and Jacquelyn Burkell, “Politics and Porn: How News Media Characterizes Problems Presented by Deepfakes,” *Critical Studies in Media Communication* 37, no. 5 (2020): 497–511; Rob Cover, “Deepfake Culture: The Emergence of Audio-Video Deception as an Object of Social Anxiety and Regulation,” *Continuum* 36 (Aug. 2022): 609–21; and Soo Yun Shin and Jiyoung Lee, “The Effect of Deepfake Video on News Credibility and Corrective Influence of Cost-Based Knowledge about Deepfakes,” *Digital Journalism* 10 (Apr. 2022): 412–32.

8. See Kashmir Hill and Jeremy White, “Designed to Deceive: Do These People Look Real to You?,” *New York Times*, 21 Nov. 2020, www.nytimes.com/interactive/2020/11/21/science/artificial-intelligence-fake-people-faces.html

‘natural’ mechanism producing ‘natural’ images whose truth was guaranteed.”⁹ Although subsequent theory—including Tagg’s—has contested the erasure of human subjectivity and social relations that this implies, the view that photography is a direct document of the natural world has remained popular.¹⁰

Framed within long-standing debates on the ontology of the photographic image, this article seeks to answer the question: Can photorealistic AI-generated images be defined as photographs? At first glance, defining AI-generated images as photographs may seem implausible: they are created without a model/scene to photograph, without a photographer, and without a camera, lens, film/sensor, or any of the other mechanical trappings traditionally associated with photography. The name photography itself points to technique, that is, light writing, so one might understandably question how a method of image creation that *does not* directly make use of light-sensitive materials or sensors to record images might still bear this name. Despite their plurality and diversity, photographic practices prior to generative AI were joined broadly together by the act of recording light in some way. Even digitally altered photographs have identifiable origin points as singular photos, though they may be subsequently carved up, airbrushed, distorted, or collaged.

AI-generated images, on the other hand, are far further removed from the recording of light. While there are many, many digital photographs in the data used to train AI image models, the influence of these photos on any one AI-generated image is indirect. The products of generative AI are not traceable to any single data point. One cannot see a single or even a collection of images in the output because it is a statistical extrapolation of that data rather than a collage or superimposition. Nevertheless, the light reflected from existing subjects and recorded in the training photographs is still somehow present in AI-generated photographs.

The simplicity of this etymological description of photography, however, is not sufficient as a definition. As Richard Bolton writes, “The intrinsic and universal properties of the photograph have never been established with complete satisfaction. . . . The wide range of photographic applications raises the possibility that photography has no governing characteristics at all save adaptability.”¹¹ With photography’s adaptability in mind, therefore, I argue

9. John Tagg, “A Means of Surveillance: The Photograph as Evidence in Law,” in *The Burden of Representation: Essays on Photographies and Histories* (Minneapolis, 1993), p. 78.

10. For a compendium of other critical essays from the 1980s that take this tack, see *The Contest of Meaning: Critical Histories of Photography*, ed. Richard Bolton (Cambridge, Mass., 1989).

11. Bolton, “Introduction,” in *The Contest of Meaning*, pp. xi–xii.

that AI-generated photographs do indeed belong within in the wider discourse of photography, not only because they have a relationship to traditionally produced photographs in their training data, but because the form they take, their use, and their interpretation are essentially photographic. In this article, I define photographic form as the visual qualities of two-dimensional images that were originally produced by photographic techniques and technology and are therefore associated with them. I exclude the physical and material dimensions of photographic objects from this discussion because the main arena for the use, dissemination, and reception of AI-generated photography (and other types of photography today) is digital.¹²

AI-generated photographs resonate with many of the classic themes in photography, including nature and the real, the unconscious and the uncanny, and discourses of power. Each of these themes are tied to the assumed truth value of photography and thus have implications for AI-generated images. If photography is defined solely by particular methods, then photo-realistic AI-generated images are not photographs. If it is defined, however, by its form and the use and interpretation of that photographic form—as it often is—how can they be anything other?

Nature, Illusion, and Deception

From manual retouching and photomontage to Photoshop and Instagram filters, photographic manipulation has always been part of the practice of photography.¹³ Despite this, photographs continue to be interpreted—against all odds—as objective copies of nature. Supporting this assumption, computer vision researchers commonly refer to photographic images as natural images. Although Talbot described photographs using this term as early as 1844, its use in computer science does not seem to stem from the historical discourse on photography. Rather, the descriptor *natural images* is more likely to have entered computer science as an analogue or extension of the term *natural language*, meaning language that has developed through human use.¹⁴ The field of natural language processing (NLP) deals with how

12. The material and physical qualities of photographs are important to consider. However, the vast majority of photographs today are viewed in digital form, even those that originated in predigital methods. So, discussion of the physical or material qualities apart from how they are manifested in the visual appearance of the two-dimensional digital replica, is excluded from the present discussion.

13. See, in regards to digital photography, for example, Manovich's early rebuttal of Mitchell: Lev Manovich, "The Paradoxes of Digital Photography," in *Photography after Photography: Memory and Representation in the Digital Age*, trans. Pauline Cumbers, ed. Hubertus V. Amelunxen et al. (Munich, 1996), pp. 61–62.

14. William Henry Fox Talbot, "A Brief Historical Sketch of the Invention of the Art," in *Classic Essays on Photography*, ed. Alan Trachtenberg (New Haven, Conn., 1980), p. 29.

machines can be made to imitate and create language that passes for human utterances or text. The notion of naturalness in the development of human languages and the presumed naturalness of photography are not equivalent, however. One is natural in the sense that it is undirected, the other merely assumed to be.

Contemporary AI tools make use of artificial neural networks. Designed to mimic the neural structures of biological brains, artificial neural networks were developed in the late 1950s but did not gain widespread interest and support until the last decade.¹⁵ Thanks to increases in computational processing power and the availability of massive datasets, researchers once again became interested in them. The results of these efforts have produced systems that are capable of “learning” complex tasks such as language or image creation.

Computer vision—the automated interpretation of digital images—has widely adopted the use of artificial neural networks over the past decade. Recent research in this area seeks to automate pattern recognition and categorization of image data via neural networks trained on a large amount of data. Decades of work in developing categorization capabilities for computer vision have also aided the development of generative AI techniques.¹⁶ Trained to recognize certain elements or objects in an image, artificial neural networks can be engineered to create new versions of the elements or objects found in those training images.

The pattern recognition and sorting functions automated by AI operate in ways that are still unclear even to the researchers developing them. The artificial neurons of the network automatically detect the features of an image, assign them statistical weights, and ultimately aid in the pattern recognition and categorization of elements of that image. Exactly how they do so, however, is not clear from the input or output to the system. This black-box aspect of contemporary AI has led to the emergence of a research subfield called Explainable AI (xAI), which has grown in prominence over the past few years.¹⁷

15. See F. Rosenblatt, “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain,” *Psychological Review* 65, no. 6 (1958): 386–408.

16. See James E. Dobson, *The Birth of Computer Vision* (Minneapolis, 2023).

17. See A. Barredo Arrieta et al., “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI,” *Information Fusion* 58 (June 2020): 82–115; Vanessa Buhrmester, David Münch, and Michael Arens, “Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey,” *arXiv*, 27 Nov. 2019, arxiv.org/abs/1911.12116; Gabriëlle Ras et al., “Explainable Deep Learning: A Field Guide for the Uninitiated,” *arXiv*, 30 Apr. 2020, arxiv.org/abs/2004.14545v1; Nicolas Gonthier et al., “Weakly Supervised Object Detection in Artworks,” *arXiv*, 5 Oct. 2018, arxiv.org/abs/1810.02569; Jan Zacharias et al., “Designing a Feature Selection Method Based on Explainable Artificial Intelligence,”

As noted, two of the most widely used AI techniques for generating images at present are GANs and diffusion models.¹⁸ The images produced by techniques such as these imitate photography but are created via processes that do not resemble traditional photographic methods in the slightest. Simply put, GANs operate by pitting two neural networks against one another—one generates and the other judges. This increases the likelihood a generated image could successfully pass for a natural image from the training data, which in turn helps achieve convincingly photorealistic imagery. Diffusion models, on the other hand, are trained by adding noise to images and then reversing this process to be able to then create images out of noise. Once trained, diffusion models are remarkably flexible in what kind of images can be produced, hence their implementation in text-to-image generators. Even from this simple description, it is clear that these processes have very little in common with traditional digital or film-based photography.

The imagination of AI—if we can even call it that—is ultimately limited by the data introduced to it, so AI-generated photographs are not created *ex nihilo*. They owe their photorealistic appearance to learnings from massive datasets of traditionally produced digital photographs. This data, once processed inside the black box of the AI, produces something that no longer maintains any semblance of indexicality. As such, AI-generated photos lack a one-to-one correspondence to a preexisting referent, what Roland Barthes calls the “umbilical cord” linking “the body of the photographed thing to my gaze.”¹⁹ And yet, they could not exist without the photographic forms that feed their creation. To borrow Barthes’s metaphor, who is to say the gaze cannot nourish itself from an invented, discontinuously composite body depicted in an AI-generated photograph?

Barthes famously described photography as a “message without a code.”²⁰ Later in *Camera Lucida*, he elaborates this conception of photography with another metaphor—the photograph as a semiotic laminate. Barthes contends that the sign and the referent of a photograph are fused to each other to such an extent that they “cannot be separated without destroying them

Electronic Markets 32 (Dec. 2022): 2159–84; Dennis Grinwald et al., “Visualizing the Diversity of Representations Learned by Bayesian Neural Networks,” *arXiv*, 26 Jan. 2022, arxiv.org/abs/2201.10859; and Ian E. Nielsen et al., “Targeted Background Removal Creates Interpretable Feature Visualizations,” *arXiv*, 22 June 2023, arxiv.org/abs/2306.13178

18. See Ian Goodfellow et al., “Generative Adversarial Nets,” *Advances in Neural Information Processing Systems* 27 (2014): 2672–80, and Jascha Sohl-Dickstein et al., “Deep Unsupervised Learning Using Nonequilibrium Thermodynamics,” in *Proceedings of Machine Learning Research* 37 (2015): 2256–65, proceedings.mlr.press/v37/sohl-dickstein15.html

19. Roland Barthes, *Camera Lucida: Reflections on Photography*, trans. Richard Howard (New York, 1981), p. 81; hereafter abbreviated *CL*.

20. Barthes, “The Photographic Message,” in *Image - Music - Text*, trans. Stephen Heath (London, 1977), p. 17.

both" (CL, p. 6). Later he reiterates that "every photograph is somehow co-natural with its referent" (CL, p. 76). Nowhere is the persistence of this view on the ontological status of photography more apparent than in artificial-intelligence research and the popular responses to AI-generated photographs. The Nvidia research team that created StyleGAN claims, for instance, that their architecture works with "a new dataset of human faces," not *photographs* of human faces.²¹ This is also evident in the name of the website This Person Does Not Exist, which elides the representative medium entirely and leaps directly to the "person" depicted (who does not exist). As Barthes writes, "a photograph is always invisible: it is not it that we see" (CL, p. 6). Indeed, the object-recognition procedures that are so prevalent in machine learning and computer-vision applications conflate the thing represented with the thing itself by definition. That is, the aim of object recognition is to automatically determine the objects present in a scene, and the photograph thus acts as the machinic eye rather than a representative medium.

The assumed physical existence of a photographic referent has remained a core concern of photography, even after the digital era and Photoshop lowered the bar for manipulation. In the 1990s, the use of digital computers to alter or composite photographic elements was positioned as a seismic shift away from the documentary function of photography.²² Scholars questioned, as I and others have done with regard to AI-generated images, whether digitally altered images could even be defined as photographs.²³ While many seemed convinced photography would never be trusted again, there are moments of recognition that the *form* of photography, no matter how manipulated or composited it may be, still implies documentation.²⁴ In other words, these visual qualities of photographicness—for example, the use of shallow depth of field, motion blur, or the granularity and contrast associated with certain types of film—are themselves enough, when replicated, to continue to carry the perceptual baggage of truth to nature that photography built up in its years of existence. We are conditioned to see photographs as documents of things captured with a camera, even if they are not made that way.

21. Karras, Laine, and Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," p. 4396.

22. See William J. Mitchell, *The Reconfigured Eye: Visual Truth in the Post-Photographic Era* (Cambridge, Mass., 1992); Timothy Druckrey, "From Dada to Digital: Montage in the Twentieth Century," *Aperture* 136 (Summer 1994): 7; and Martin Lister, "Introductory Essay," in *The Photographic Image in Digital Culture*, ed. Lister (New York, 1995), p. 1.

23. See Amelunxen, Stefan Iglhaut, and Florian Rötzer, "Preliminary Remarks on the Project 'Photography after Photography,'" in *Photography after Photography*, p. 10.

24. See Jonathan Green, "Pedro Meyer's Documentary Fictions," *Aperture* 136 (Summer 1994): 34; Rötzer, "Re: Photography," in *Photography after Photography*, pp. 21–22; and Martha Rosler, "Image Simulations, Computer Manipulations: Some Considerations," in *Photography after Photography*, p. 37.

Thirty years later, it is clear that Photoshop and digital computing have not, in fact, done much to weaken photography's default assumptions.²⁵ Indeed, the twin ascendance of social media and mobile phone technology starting in the middle of the first decade of the twenty-first century ushered in a new era of photographic credibility. More people are taking more photographs, but the nature of many of these photos has shifted. People still take pictures of important things and events, but they also take many more pictures of things that are not particularly important. A lot of the photographs taken today are throwaway documents meant to communicate quickly with others. They are also a part of normalized, participatory surveillance. For example, e-scooter and package-delivery apps regularly require either the user or the driver to provide a photograph as evidence that the scooter was parked properly or that the package was delivered to the recipient's door. Abuses of power—from police brutality to war crimes to everyday racist encounters—are far more easily documented and exposed by members of the public thanks to these technological shifts.²⁶ This trust in photographic documentation persists even after the onslaught of online disinformation during the so-called post-truth era of the mid-2010s.²⁷

As all of these technological changes have occurred, scholars and artists have continued to mull the documentary facets of contemporary photography and the relationship between photographic subject, photographer, and viewer.²⁸ The documentary turn in contemporary art beginning in the '00s played precisely on the continued trust in the photographic form through subtle manipulations and fictions.²⁹ Photography seems, therefore, to be unbreakably bound to its assumed ability to provide evidence.³⁰

25. See Scott Walden, "Truth in Photography," in *Photography and Philosophy: Essays on the Pencil of Nature*, ed. Walden (Malden, Mass., 2010), pp. 108–10.

26. See *Amateur Images and Global News*, ed. Kari Andén-Papadopoulos and Mervi Pantti (Chicago, 2011).

27. See Fredrik Bjercknes, "Images of Transgressions: Visuals as Reconstructed Evidence in Digital Investigative Journalism," *Journalism Studies* 23, no. 8 (2022): 951–73, and Ryan Watson, *Radical Documentary and Global Crises: Militant Evidence in the Digital Age* (Bloomington, Ind., 2021).

28. See Ariella Azoulay, *The Civil Contract of Photography* (Brooklyn, N.Y., 2008); Jacques Rancière, "Notes on the Photographic Image," trans. Darian Meacham, *Radical Philosophy* 156 (July–Aug. 2009): 8–15; and Barbara Savedoff, "Documentary Authority," in *Photography and Philosophy*, pp. 111–37.

29. See Mark Nash, *Experiments with Truth* (Philadelphia, 2004); Carrie Lambert-Beatty, "Make-Believe: Parafiction and Plausibility," *October* 129 (Summer 2009): 51–84; and T. J. Demos, *The Migrant Image: The Art and Politics of Documentary during Global Crisis* (Durham, N.C., 2013).

30. See Aaron Meskin and Jonathan Cohen, "Photographs as Evidence," in *Photography and Philosophy*, pp. 70–90, and Cohen and Meskin, "On the Epistemic Value of Photographs," *Journal of Aesthetics and Art Criticism* 62 (May 2004): 197–210.

The sense (and sensation) of trust in photography continues to coexist with the rational knowledge of its malleability, however. As with optical illusions, truth and falseness are in tension within the image. We know a photograph is not real, but we see it as such. Theory on pictorial illusion thus offers insights into the function of photographic form for AI-generated images. Photographs are rarely discussed in the context of pictorial illusionism in art history, but there is a body of literature on this topic in philosophy that shows computer scientists are not the only ones interested in object recognition.³¹ Beginning a decades-long debate, Kendall L. Walton asserts that the existence of a photographic subject is the key distinction in separating photographic and other forms of realism, even if that photographic depiction is of “werewolves and Martians.”³² Earlier, in a rare mention of photography, Richard Wollheim claims that “correctness” matters when viewing a photograph.³³ However, he is forced to add the caveat that this distinction becomes difficult when models are used to depict historical figures. Like Barthes, these writers elide the representative medium, choosing to focus on the existence of the referent. In defining or discussing photography in this way, AI-generated images cannot possibly pass muster, despite the presence of existing referents in the training data. In a wider sense, this focus on existence means that photography is again and again boiled down to an objective procedure that closes down theoretical possibilities for it.

Photographic illusion is only valued for its truth to nature if one is familiar with the process of photography and the peculiarity of its form. Any reading of a picture—painted or otherwise—as realistic or naturalistic is a construct created both by the maker of the picture and the cultural conditioning of the viewer.³⁴ I would argue, therefore, that photographic verisimilitude—not verisimilitude itself—is the root of the illusion (or deception) for AI-generated images.³⁵ By this I mean that there is a type of verisimilitude that

31. Regarding art history/philosophy, see “Agency and Automatism: Photography as Art since the Sixties,” a special issue of *Critical Inquiry* 38 (Summer 2012). Regarding philosophers concerned with perception, see, for example, Jessica Pepp, “On Pictorially Mediated Mind-Object Relations,” *Inquiry* 66, no. 2 (2023): 246–74. For key essays, see *Photography and Philosophy*.

32. Kendall L. Walton, “Transparent Pictures: On the Nature of Photographic Realism,” *Critical Inquiry* 11 (Dec. 1984): 250. The wider topic/debate (along other argumentary lines) was earlier sparked by Roger Scruton, “Photography and Representation,” *Critical Inquiry* 7 (Spring 1981): 577–603.

33. Richard Wollheim, “Seeing-as, Seeing-in, and Pictorial Representation,” in *Art and Its Objects: With Six Supplementary Essays* (New York, 1980), p. 207.

34. See E. H. Gombrich, *Art and Illusion: A Study in the Psychology of Pictorial Representation* (New York, 1960), and Nelson Goodman, *Languages of Art: An Approach to a Theory of Symbols* (Indianapolis, 1968).

35. This applies to photography too, as Goodman notes in a footnoted source; see Goodman, *Languages of Art*, p. 15 n. 15.

is specific to photography and bears with it perceptual baggage attached to that form. Simply put, the viewer is conditioned to interpret the AI-generated image as documentation because it looks specifically like a photograph. As mentioned, the visual form of photography is the result of the physics of the materials and tools used to create it. So, if the form can be (near) flawlessly replicated, the subject's existence is beside the point. What matters, ultimately, is the photographic form.

Barthes's idea of photography as a message without a code or a semiotic laminate not only reinforced the presumed naturalness of photography but also its objectivity. In "Rhetoric of the Image," Barthes famously analyzed the semiotics of advertising photography for Panzani, an Italian food company. The ad's impact lies in the implication that photography is a faithful rendering of nature, "the myth of photographic 'naturalness.'"³⁶ Barthes's designation of the photograph as a "perfect *analogon*" of reality but not, of course, reality itself reaches for an ideal that photography always somehow fails to fulfill.³⁷ He outlines two factors that distinguish drawing from photography: the need for "*rule-governed* transpositions" and the fact that drawing does not "*reproduce everything*."³⁸ Never mind that digital photography is literally produced through discontinuous code transformed into an image (Barthes's writing on photography appeared just on the cusp of the development of digital photography).³⁹ All photographs do, in fact, have a code. This is the trick of photography: to disguise its transpositional process with the particularities of its form.

The transpositions deployed by lens and chemicals in traditional photography produce an image that is subject to the rules of lens distortion and light-sensitive reactivity.⁴⁰ The historic bias latent in photographic materials (for example, color-film stock or digital-camera design that is tailored to the depiction of light-skinned people and thus excludes those with darker skin tones) is a case in point.⁴¹ Photography does not capture nature equally and without transformation. Moreover, a photograph does not necessarily need to reproduce everything. For example, photographic techniques like shallow depth of field strategically blur unwanted background material. In fact, one of the hallmarks of StyleGAN portraits is the use of shallow depth of field.

36. Barthes, "Rhetoric of the Image," in *Image – Music – Text*, p. 44.

37. Barthes, "The Photographic Message," p. 17.

38. Barthes, "Rhetoric of the Image," p. 43.

39. On this theme, see also Vilém Flusser, *Into the Universe of Technical Images*, trans. Nancy Ann Roth (Minneapolis, 2011).

40. See Joel Snyder, "Picturing Vision," *Critical Inquiry* 6 (Spring 1980): 499–526.

41. See Lorna Roth, "Looking at Shirley, the Ultimate Norm: Color Balance, Image Technologies, and Cognitive Equity," *Canadian Journal of Communication* 34, no. 1 (2009): 111–36.

This seems to serve a dual purpose: it creates a photographic gestalt and avoids the need to generate convincing background material.

For Barthes, the important aspect of the Panzani image is not that the scene it captures actually existed but that it *appears* to have existed. By virtue of its form, an image can be read as a photograph without having an automatically produced relationship to a natural or real referent. The idea that photographs are uncoded and objective is easily disputed, but it is an idea that nonetheless demands further discussion because it is still such a prevalent assumption. The “myth of photographic ‘naturalness’” not only underpins the impact of the photographic advertisement but also the impact of AI-generated photography. Both types of image deceive in different ways. And both deceptions are built on photographic form.

Many theorists have tried (and some would say failed) to define photography by formal parameters. One of the most notorious formalists in photography theory was curator and artist John Szarkowski. In the 1960s, Szarkowski defined photography in terms that explicitly compared it to and set it apart from painting.⁴² His concern with doing so has been attributed to the unavoidable influence of Greenbergian modernism in New York at the time.⁴³ He outlines five key facets of photography, including the thing-in-itself, writing, “The first thing that the photographer learned was that photography dealt with the actual. . . . But he learned also that the factuality of his pictures, no matter how convincing or unarguable, was a different thing than the reality itself.”⁴⁴ In order to assert the photographer’s rightful place as an artist, Szarkowski needed to deny that “what the camera saw was the truth.” However, the constructed truth of the photograph was still inextricably tied to something external: “Paintings were *made*—constructed from a storehouse of traditional schemes and skills and attitudes—but photographs . . . were *taken*.”⁴⁵ Given what we know about how they are produced and their lack of existing referent, one might conclude that AI-generated photographs are undeniably made rather than taken. They are signifiers without signifieds, in other words, empty signifiers, in the most basic sense. The making of AI-generated photos, however, is automatic. In the case of text-to-image generators or GANs created around custom parameters,

42. See John Szarkowski, *The Photographer’s Eye* (New York, 1966), p. 6.

43. See Gene Thornton, “The Place of Photography in the Western Pictorial Tradition: Heinrich Schwarz, Peter Galassi and John Szarkowski,” *History of Photography* 10, no. 2 (1986): 88; Geoffrey Batchen, “John Tagg (1949–),” in *Fifty Key Writers on Photography*, ed. Mark Durden (New York, 2013), pp. 235–36; and Clement Greenberg, “Post Painterly Abstraction,” *Art International* 8 (Summer 1964): 63–64.

44. Szarkowski, *The Photographer’s Eye*, p. 8.

45. *Ibid.*, pp. 8, 6.

the human directs the automatic process in a way not entirely dissimilar to photography. As such, we might still say that AI-generated photographs are taken as composite learnings from the data as a whole, but they are also made by a human in a specific way.

The process by which a photograph is made may determine its form, but it does not define it. The mere existence of photography (and its long-standing association with actuality) is now grounds for *any* image to connote naturalness provided it is sufficiently photographic. Thus form, not process, is paramount. In using the term *form*, I do not mean to imply that photography has a static or universal visual appearance. Social, political, and historical factors contribute to a viewer's understanding of photographic form through its use, experience, and reception. Tagg famously stated:

What alone unites the diversity of sites in which photography operates is the social formation itself: the specific historical spaces for representation and practice which it constitutes. Photography as such has no identity. . . . Its function as a mode of cultural production is tied to definite conditions of existence, and its products are meaningful and legible only within the particular currencies they have.⁴⁶

Following Tagg, one could easily define AI-generated photographs within the discourse of photography.

The position Tagg takes in the above statement, however, is emblematic of the reaction in the 1980s to formalist theory in the field. Geoffrey Batchen argues that theory such as Tagg's merely exchanges one extreme position (formal) for another (cultural). He writes, "The photograph still has an essence, but now it is found in the mutability of culture rather than in its presumed other—an immutable nature."⁴⁷ He then argues against essentializing *either* nature or culture in relation to photography.⁴⁸ Given this, it may seem strange to use the term *form* and in a way that leans toward Tagg's conception of photography. In doing so, however, I aim to show that the naturalistic form of AI-generated photographs is what connects them to the discourse of photography, even if the reading of that form as natural is culturally conditioned.

In order for AI-generated photographs to be photographs, however, the process(es) of photograph(y/ies) must be separated from our understanding

46. Tagg, "Evidence, Truth, and Order: Photographic Records and the Growth of the State," *The Burden of Representation*, p. 63.

47. Batchen, *Burning with Desire: The Conception of Photography* (Cambridge, Mass., 1999), p. 20.

48. See *ibid.*, pp. 20–21.

of the photograph. Once this is done, form (in the expanded sense noted above) is what remains. Incidentally, Gerhard Richter also reached this conclusion: that process must be separated from our understanding of photography, in relation to the large photorealistic works he created from the 1960s onward. He explains his thoughts in an interview from 1972:

I'm not trying to imitate a photograph; I'm trying to make one. And if I disregard the assumption that a photograph is a piece of paper exposed to light, then I am practising photography by other means: I'm not producing paintings that remind you of a photograph but producing photographs. And, seen in this way, those of my paintings that have no photographic source (the abstracts, etc.) are also photographs.⁴⁹

Richter sees his paintings as a reflection of the imprecision or uncertainty of visual perception and the attendant desire to arrest images. Photography is therefore defined by a visual philosophy (or impulse) rather than a particular technical apparatus, meaning that its praxis expands indeterminately. There is certainly a Duchampian tendency in Richter's statements: his works are photographs because he says they are. However, their visual form still plays a large role in the reception of these works as photographic.

Essential to the cultural understanding of photography—and relatively consistent throughout its history—is the perceived connection to a naturally occurring source, no matter how reworked or constructed the image is. The idea that photographs are naturally produced images is essential to the way those that developed the first photographic processes understood their inventions. As a means of achieving what Lorraine Daston and Peter Galison call “mechanical objectivity,” photography was used by scientists in the nineteenth century to automatically and mechanically capture nature in an image. They write, “The mechanical, objective photograph had allegedly been traced by ‘nature’s pencil’ alone, and nature was entirely artless.”⁵⁰ Although people at that time were aware photographs could be manipulated, the automatic process eased scientists’ considerable anxiety around unwanted interpretation influencing the appearance of their images.⁵¹ AI-generated photographs are also produced automatically. Rather than alleviating anxiety around subjectivity, however, they have been greeted with justifiable suspicion. Their naturalism and the ease by which they are produced have, as noted, raised fears

49. Gerhard Richter, “Interview with Rolf Schön, 1972,” in *The Daily Practice of Painting: Writings 1962–1993*, trans. David Britt, ed. Hans-Ulrich Obrist (Cambridge, Mass., 1995), p. 73.

50. Lorraine Daston and Peter Galison, “Mechanical Objectivity,” *Objectivity* (New York, 2007), p. 133.

51. See *ibid.*, pp. 131–38.

around their deceptive potential. Whether the places, people, and things in AI-generated images are real or not, they have real-world effects.

AI-generated photographs are enmeshed in a web of alternate realities, both those they produce through their existence and those their existence depends on. The next section addresses how one might characterize the realities evoked by AI-generated photographs.

The Surreal, the Real, and the Hyperreal

AI-generated images are often glibly described as surrealist.⁵² The formal hallmarks they share with the visual art of the surrealist movement from the 1920s and '30s are now familiar: incongruous pairings of objects, curved and bulbous shapes, repeated or distorted patterns, and the appearance of grotesque or amputated bodies. AI-generated photographs are based on learnings from large datasets of natural photographs and aim to imitate or simulate the photographic form. Their connection to the surreal is therefore not only the result of superficial distortion or glitches but something that is found in the form itself, in its connection to the real.

Photography was an important medium for the surrealists. They saw it as a means to access and create realities via automation, tapping into psychoanalytical concepts like the unconscious and uncanny in the process.⁵³ Like early photographers who idealized the medium as a means of creating objective images, the surrealists wanted to rid their images of conscious influence. In doing so, they did not seek objectivity, however, but another plane of subjectivity. Susan Sontag claims that photography is “natively surreal” because of “its irrefutable pathos as a message from time past.”⁵⁴ Discussions of realism or naturalism in the theory of photography are based around a formal conceit of photography: its connection to the real. This connection to the real—or the perceived real—makes the photograph an *objet trouvé*, and, like the bric-a-brac the surrealists collected at flea markets, the referent of a photograph is found and not made. The strange and unexpected juxtapositions in AI-generated photographs can also be characterized as found, as accidents of automation to be discovered.

In *Camera Lucida*, Barthes introduces the concept of the punctum, a subjectively designated element of the photograph that “pricks” the viewer and

52. See for example, Caitlin Schneider, “The Surreal Artwork of Artificial Intelligence,” *Mental Floss*, 24 June 2015, www.mentalfloss.com/article/65458/surreal-artwork-artificial-intelligence, and Sophia June, “What Is DALL·E? The AI Machine Creating Instant Surrealist Art,” *Nylon*, 23 June 2022, www.nylon.com/life/dall-e-ai-art-machine-explained

53. See Rosalind Krauss, “The Photographic Conditions of Surrealism,” *October* 19 (Winter 1981): 3–34.

54. Susan Sontag, *On Photography* (New York, 1977), pp. 51, 54; hereafter abbreviated *OP*.

draws them out of the image, outside the frame to a reality beyond (*CL*, p. 43). There have been near-endless debates, theories, and countertheories about this concept since the time Barthes's book was published.⁵⁵ I do not wish to rehash this but rather address a simple question: Do AI-generated photographs have a punctum? In other words, is the realness of a referent necessary to isolate a punctum or is it something else? Barthes writes, "The *punctum*, then, is a kind of subtle *beyond*—as if the image launched desire beyond what it permits us to see" (*CL*, p. 59). From an etymological perspective, the *sur-* prefix in *surreal* literally means "beyond" or "above," so one might say that the punctum is the locus of surrealism in photography. As Adam Lowenstein argues, for Barthes (and film critic André Bazin), the "commitment to the realism of the photographic image, so often misunderstood as a naïvely literalist stance, is much more accurately described as an investment in surrealism."⁵⁶ In the case of the AI-generated photograph, the punctum is usually tied to its specificity. It is an accidental something created from everything and nothing that allows us to see beyond the image in question.

In the opening pages of the book, Barthes describes his experience looking at a photograph of Napoleon Bonaparte's brother and thinking, "'I am looking at eyes that looked at the Emperor'" (*CL*, p. 3). Barthes draws a direct connection between himself and the photographic subject, as if a gaze is a physical force transferring energy across space and time. However, his experience is more a matter of fantasy than physics. Barthes imagines a reality of this person's existence, a reality sparked by the verisimilitude of the representation. So lifelike is the depiction that Barthes cannot help but feel connected to the milieu that the figure is imagined to have existed within—the "blind field" springing up beyond the photograph (*CL*, p. 55). Not only was Barthes looking into the eyes that looked at the emperor, but he was looking at a photo of a person the emperor himself had seen. Barthes writes, "Photography's inimitable feature (its *noème*) is that someone has seen the referent (even if it is a matter of objects) *in flesh and blood*, or again *in person*" (*CL*, p. 79). With the advent of AI-generated photographs, this feature is in fact imitable. AI-generated photographs accomplish the same effect: the appearance of the photographic rendering of a preexisting subject matter. This is why the revelation that "this person does not exist" may be disturbing. The person depicted is not alive, was not once alive, and is not long dead. This

55. See, for example, *Photography Degree Zero: Reflections on Roland Barthes's "Camera Lucida"* ed. Batchen (Cambridge, Mass., 2009).

56. Adam Lowenstein, "The Surrealism of the Photographic Image: Bazin, Barthes, and the Digital Sweet Hereafter," *Cinema Journal* 46 (Spring 2007): 54.

person never lived at all, despite appearances that would suggest otherwise. AI-generated photographs thus expose the experience of reality in all photographs as fantasy.

In some ways, Barthes's theory of the punctum marks an endpoint in a trajectory of theory that sought to tie reality—rather than nature—to the experience of the photographic form. Fifty years earlier, Walter Benjamin explored the relationship between surrealism and photography more explicitly. Benjamin was writing in the 1930s at the height of the movement's influence, so it is perhaps unsurprising he took an interest in the relationship between surrealist concepts and the medium of photography. In "Little History of Photography," he describes a facet of photography that sparks or ignites a sense of the reality beyond that resonates with Barthes's much later concept of the punctum:

No matter how artful the photographer, no matter how carefully posed his subject, the beholder feels an irresistible compulsion to search such a picture for the tiny spark of contingency, the here and now, with which reality has, so to speak, seared through the image-character of the photograph.⁵⁷

Benjamin coined the concept of the "optical unconscious" in this essay and mentions the term again in "The Work of Art in the Age of Mechanical Reproducibility."⁵⁸ Although ultimately left underdeveloped, the concept of the optical unconscious relates to the facet of photography where referents are depicted unintentionally or in a way that is not immediately available to the conscious mind composing the picture.

In discussing and defining the optical unconscious anew, Shawn Michelle Smith and Sharon Sliwinski argue that "the pivot point is not photography's so-called indexical relationship to reality, but rather its proximity to fantasy."⁵⁹ Thus, in their view, Benjamin positions photography as a means of tapping into deeper expressions of the collective psyche in mass-media society: desires, fears, points of destabilization. Deep-learning models have something akin to their own form of the unconscious in what is called latent

57. Walter Benjamin, "Little History of Photography," trans. Edmund Jephcott and Kingsley Shorter, in *The Work of Art in The Age of Its Technological Reproducibility, and Other Writings on Media*, trans. Jephcott et al., ed. Michael William Jennings, Brigid Doherty, and Thomas Y. Levin (Cambridge, Mass., 2008), p. 276. See also Benjamin, "The Work of Art in the Age of Its Technological Reproducibility," trans. Jephcott and Harry Zohn, in *The Work of Art in The Age of Its Technological Reproducibility, and Other Writings on Media*, p. 37.

58. Benjamin, "Little History of Photography," p. 278. See also Benjamin, "The Work of Art in the Age of Its Technological Reproducibility."

59. Shawn Michelle Smith and Sharon Sliwinski, "Introduction," in *Photography and the Optical Unconscious*, ed. Smith and Sliwinski (Durham, N.C., 2017), p. 9.

space. This is where connections are formed and relationships constructed among the data, proceeding in ways that are not predetermined and therefore unpredictable and difficult to map. To an even greater degree than traditional photography, the complex products of deep-learning systems tap into fantasy and the uncanny.

AI image generators allow users to quickly and easily create an image of their fantasies—a Pygmalion impulse, one might say, to create something perfect rather than real. Instead of superior draftsmanship or other hand-craft skills, one needs the appropriate sequence of words or parameters. This is, in fact, a crucial distinction. Unlike the mythical artist Pygmalion, photographers—creating both traditional and AI-generated photographs—do so with the help of automation. Photographs seem to appear, whether in the developing fluid, on a digital camera screen, or after clicking a generate button. One does not perceive how these images are assembled, piece by piece. The photograph appears as an *objet trouvé*. Automation helps create happenstance, ceding control from the creator. Even though automation is not a force of nature per se, it often feels like one. This is why, historically, photography has been understood as a means to access unconscious desire, the spirit realm, or other hidden layers of reality. “Photographs depict realities that already exist,” Sontag writes, “though only the camera can disclose them” (*OP*, p. 122). Photography vis-à-vis automation is understood as a decoder technology, revealing the latent, objective world that has not been created by the photographer. AI-generated photographs are no different. They can be said to reveal the hidden facets of the data used to produce them—the object world of the dataset.

As seductive as this idea is, however, photography is not a decoder but a transformer. Both traditionally produced photographs and AI-generated photographs create the semblance of reality through representation.⁶⁰ The argument that photography is not representational hinges exclusively on the assumed *existence* of the photographic subject rather than its photographic appearance and form.⁶¹ The data that is used to produce AI-generated photos is, of course, also itself a representation. And so, photography has become a regress of representations. Indeed, scholars have long understood photography as a form of representation rather than a manifestation of the real. Sarah Kember, responding to fears around manipulation of photographs with computer technology in the 1990s, wrote, “How can we panic about the loss of the real when we know (tacitly or otherwise) that the real is always

60. See Snyder and Neil Walsh Allen, “Photography, Vision, and Representation,” *Critical Inquiry* 2 (Autumn 1975): 143–69.

61. See Scruton, “Photography and Representation.”

already lost in the act of representation? Any representation, even a photographic one only constructs an image-idea of the real.”⁶² If the real is already lost, hidden or not, then we are dealing with the idea of simulation.

Jean Baudrillard famously defined simulation as “the generation by models of a real without origin or reality.”⁶³ This also serves fairly well as a definition for AI photography: a real generated by machine learning models that has no origin or reality. Baudrillard’s choice of the words *generation* and *models* is fitting if not almost spooky.⁶⁴ Although his writings on simulacra both directly and obliquely referenced the emerging information age and the ascendance of personal computing in the early 1980s, his theory is still relevant to discussions of digital technology today. To speak of them in Baudrillard’s terms, then, AI-generated photographs are quintessentially hyperreal.⁶⁵

In addition to being a theorist, Baudrillard was also a photographer. In an interview from 1991, he describes his practice:

I take photographs, but not in terms of the photographic subject, or any vision of the world, etc. For me, a photograph is an expression of the pure object—it is the object which appears. I’ve no wish to impose myself as a subject in the world of images. I want the object to appear—not in a particularly aesthetic way but, on the contrary, perhaps even in a non-aesthetic way, as a purely objective object.⁶⁶

For Baudrillard, photography is a process of objectification. He is not the first to see it this way. Sontag writes that photography “turns people into objects” (*OP*, p. 14). Barthes echoes this: “Photography transformed subject into object” (*CL*, p. 13). The rise of machine readability reinforces this view of photography as objectifier. In order to perform object-recognition tasks via machine learning, everything in the world must first be transformed into objects to be identified. Images become image data. Objects are identified. Trained on such data, image generators cause objects to appear.

This objectified world, which fuels the creation of AI-generated photographs, is a perceptual realm apart from human experience. A number of scholars have addressed this machine-readable mirror world. For example,

62. Sarah Kember, “‘The Shadow of the Object’: Photography and Realism,” *Textual Practice* 10, no. 1 (1996): 145.

63. Jean Baudrillard, *Simulacra and Simulation*, trans. Sheila Faria Glaser (Ann Arbor, Mich., 2014), p. 1.

64. The words are the same in the original French.

65. Fitting given that both *surreal* and *hyperreal* have the same etymological root meaning.

66. Baudrillard, “Fractal Theory,” interview by Nicholas Zurbrugg, in *Baudrillard Live: Selected Interviews*, trans. Zurbrugg, ed. Mike Gane (New York, 1993), p. 168.

the products of image generators clearly fall within Joanna Zylińska's definition of "nonhuman photography": images that are, in a variety of ways, not "*of*" "*by*" and/or "*for*" human use or viewership.⁶⁷ In a similar vein, Harun Farocki theorized the concept of "'operative images'" as images that are meant to be read or interpreted by machines rather than humans.⁶⁸ Thanks to machine learning, advanced object-recognition tools have made it so that any and all images can and are used as operative images. This harkens back to Baudrillard, who states that, "[The real] is no longer anything but operational."⁶⁹ Thus not only are the images operational but the real itself: datafied, quantified, categorized, and machine interpretable. If the real is operational, then it must also already be objectified. Following from this, AI-generated photographs depict shadows or iterations—composites—of already-objectified subjects. The clues were already planted long ago in the idea that photography turns subject into object—it is machine readability *avant la lettre*.

Exquisite Corpse

Perhaps it is these operational roots that make AI-generated photographs unsettling or uncanny for humans. Generative models often have blind spots or difficulties with certain tasks that can lead to glitches and artifacts. For instance, StyleGAN has a hard time creating symmetry (for example, two different earrings or two sides of a person's glasses frames may not match) and, although it is not designed to produce images of multiple people, it sometimes adds distorted or grotesque figures next to the central portrait (fig. 4). Many text-to-image generators, on the other hand, struggle to depict human hands with the correct proportions and number of digits (fig. 5). Unlike GAN, text-to-image generators can also struggle with articulating convincing facial features. For example, entering the prompt "cute babies sitting in flower pots, Anne Geddes, studio photography" into DALL-E 2 produces a number of somewhat disturbing images of babies half-buried in dirt or with missing and distorted features (fig. 6). Likewise, the experience of refreshing thispersondoesnotexist.com can create an uncanny sensation due to artifacts in the generation process such as subtle differences in lighting on facial features, hair that does not sit quite right on the head, or parts of the face that do not quite match one another in scale or proportion.

67. Joanna Zylińska, *Nonhuman Photography* (Cambridge, Mass., 2017), p. 5.

68. Harun Farocki, "Phantom Images," trans. Brian Poole, *Public* 29 (2004): 17, public.journals.yorku.ca/index.php/public/article/view/30354

69. Baudrillard, *Simulacra and Simulation*, p. 2.



FIGURE 4. AI-generated portrait, StyleGAN/thispersondoesnotexist.com, June 2022.

Glitches such as these do not fully account for the uncanny quality of AI-generated images, however. On a deeper level, these images can be thought of as collages or composites—not of preexisting parts but of generated parts. They are like the surrealist exquisite-corpse game, where a body is formed in pieces, given limited information. Just as the common machine learning task of object recognition ignores the representational function of the image as a whole in favor of identifying its parts, AI-generated images constitute a whole based purely on parts pieced together. Even though the bodies and spaces that appear in each image may be seamlessly constructed, they remain discontinuous and particular. Each element of a face, for example, is constituted individually: it is a collection of eye plus nose plus lips in the right configuration rather than an indivisible and continuous whole. The space in the image is pieced together from both the words in the prompt and whatever connections are produced in the model's latent space.



FIGURE 5. AI-generated image, Stable Diffusion (prompt: “childrens’ hands reaching for candy”), August 2022.

This difficult to pinpoint uncanny quality of machine-produced images ties back to Barthes’s concept of the punctum. In *Camera Lucida*, he claims over and over again that the punctum cannot be explained or rationalized. This leads to an explicit indecisiveness that plays out in the text itself, as Barthes repeatedly shifts his identification of the punctum of each image. It seems that he willfully sees only what he wants to see.⁷⁰ Barthes asserts, “What I can name cannot really prick me. The incapacity to name is a good symptom of disturbance” (*CL*, p. 51). In other words, the emotional response stems from categorical uncertainty.

This is not something unique to AI-generated photography, however. It is a facet of all subjects made object that are represented in the photographic

70. See Margaret Olin, “Touching Photographs: Roland Barthes’s ‘Mistaken’ Identification,” in *Photography Degree Zero*, pp. 75–89.



FIGURE 6. AI-generated image, DALL-E 2 (prompt: “cute babies sitting in flower pots, Anne Geddes, studio photography”), August 2022.

form. Death is a powerful and reoccurring theme in the theory of photography because of its treatment of the body. An individual pictured in a photograph can be understood as both a body that is perpetually alive and a perfectly embalmed corpse. Sontag writes, “All photographs are *memento mori*. To take a photograph is to participate in another person’s (or thing’s) mortality, vulnerability, mutability. Precisely by slicing out this moment and freezing it, all photographs testify to time’s relentless melt” (*OP*, p. 15). No human—no subject—can be expected to remain in such a state.

As noted, Barthes sees photography as a visceral, fleshy connection between the viewer’s gaze and the object/subject. *Camera Lucida* is, after all, a meditation on photography and the death of his mother. Where Barthes sees a bodily connection, however, some see an amputation or decapitation. Theorist Philippe Dubois, for example, proposes that a photograph does not

connect but rather cuts off contact with the referent, making it unreachable except from a certain visual perspective.⁷¹ In other words, the shutter has the precision of the guillotine's blade, performing an execution.⁷² The photograph is expected to identify, to document. However, both traditionally produced and AI-generated photographs continually defy this mandate.

The primary aim of this article has been to understand, from a formal and theoretical perspective, the ways in which AI-generated photography fits within the context of the history and theory of photography more generally. In closing, I would like to tie some of the above theoretical reflections to some of the concerns that I addressed in the opening paragraphs, namely the political implications of AI-generated photography. Indeed, there is one additional and absolutely crucial way that AI-generated media resonates with photo history and theory: by distilling and refracting the mechanisms of power and control in contemporary society.

In the 1980s, theorists of photography such as Tagg and Allan Sekula began to reflect on how the practice of photography was instrumentalized in the service of structures of control and power, inspired by the work of Michel Foucault. In doing so, they reframed discussions around subject and objecthood in photography in a more explicitly sociopolitical and historically grounded way. For example, Tagg describes the standardized format of early mugshot photography as "the product of the disciplinary method: the body made object; divided and studied."⁷³ Sekula, on the other hand, describes the relationship between the uptake of photography and the popularity of eugenics-based practices of phrenology in the nineteenth century as "quite coincident."⁷⁴

The use of machine learning techniques to sort, analyze, and produce images echoes uses of photography in such pseudoscientific endeavors. Kate Crawford and Trevor Paglen's project ImageNet Roulette and accompanying essay, "Excavating AI," demonstrate this point elegantly. In the project, they expose the ways that ImageNet, a popular image training set for machine learning applications, and other image-based training sets perpetuate prejudicial classification of people. They write, "ImageNet is an object lesson, if you will, in what happens when people are categorized like objects."⁷⁵

71. See Philippe Dubois, "Palimpsestes: La Photographie comme appareil psychique (principe de distance et art de la mémoire)," in *L'acte photographique et autres essais* (Nathan, 1990), pp. 265–66.

72. See Dubois, "Le Coup de la Coupe: La question de l'espace et du temps dans l'acte photographique," in *L'acte photographique et autres essais*, p. 155.

73. Tagg, "A Means of Surveillance," p. 76.

74. Allan Sekula, "The Body and the Archive," *October* 39 (Winter 1986): 12.

75. Kate Crawford and Trevor Paglen, "Excavating AI: The Politics of Images in Machine Learning Training Sets," AI Now Institute, 19 Sept. 2019, www.excavating.ai/

If people in training sets become objects, the reverse process of generating images in, for example, diffusion model-based tools like DALL-E, turns objects into people. The question is, what kind of people are created? I would argue that they are average people, that is, statistical samplings masquerading as individuals. They are, in other words, persons who do not exist.

In “The Body and the Archive,” Sekula discusses the use of composite photography by eugenicists such as Francis Galton to identify different ethnic or criminal types. He argues that “this merger of optics and statistics was fundamental to a broader integration of the discourses of visual representation and those of the social sciences in the nineteenth century.”⁷⁶ The idea of the *average man*—the imaginary middle and its twin concept, the deviant—were thus born during this era. The successor of nineteenth-century statistics is contemporary artificial intelligence. AI, particularly deep learning, has often been described as “statistics on steroids.”⁷⁷ Thus, AI-generated photographs can be characterized as complex statistical composites, the product of mathematically described relationships charted among a multitude of deconstructed parts. This process is remarkably expedient, if conceptually suspect.

Each image in a training set is constrained to its segmentation, its categorization, and the objectification of what it represents. These qualities are then artificially divided into parts, siloed and boxed away from one another (digitized), and then united again through reconstitution of statistically occurring relationships between these supposed parts. They then make what a human might perceive as a whole (an average) or, when things go wrong, what a human might perceive as grotesque (a mistake or deviation). Ultimately, “this person” does not exist, just as the average man of the nineteenth century does not exist. They are both statistical constructions.

The overarching question that frames this inquiry is: Can photorealistic AI-generated images be defined as photographs? In the preceding text, I have aimed to demonstrate that, yes, they should be defined as such, given their formal reception and their theoretical resonance with key concerns in the field: nature and the real, the unconscious and the uncanny, and discourses of power. While a traditional photograph may be seen to epitomize specificity, AI-generated photographs contain multitudes. Seen a different way, however, AI-generated photographs effectively imitate photography, so they are an artificial or constructed instantiation of the specificity they

76. Sekula, “The Body and the Archive,” p. 18.

77. Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World* (Cambridge, Mass., 2018), p. 32. It is difficult to pinpoint the origin(s) of this phrase.

actually lack. Does it matter, though? If the only concern is defining or delimiting such images, yes of course it matters, but if one looks at how these images are used, circulated, and received—certainly not.

One of the first AI-generated photographs to go viral on social media was an image of Pope Francis wearing an oversized white puffer jacket and a blingy cross around his neck (fig. 7). It may seem unbelievable, reading this description, that so many people saw the image and thought that



FIGURE 7. AI-generated image, Midjourney. “The Pope Drip,” Reddit, 24 March 2023, www.reddit.com/r/midjourney/comments/12ovhdc/the_pope

Pope Francis had actually donned a trendy streetwear look. However, the effect of the image has as much to do with how photographs circulate online as it does the formal qualities of the image itself. It was first posted on Reddit on 24 March 2023 in a forum devoted to Midjourney, so the AI source of the photo was not initially hidden.⁷⁸ However, it was soon posted on Twitter and other social media sites with no reference to its AI origins. This case shows that even an imperfect exemplar of AI photography can spread quickly if it is widely interpreted as documentary. On more than a cursory glance at the image, one can easily see that the pope's right hand is distorted—a standard tell that the image is AI-generated.⁷⁹ Social media use, however, consists largely of quick glances on small phone-sized screens. If this image is accepted as a photograph so easily, there is nothing stopping a less obviously distorted image from enjoying the same level of credulity.

AI-generated photographs are already beginning to mix among so-called natural images. The online circulation of AI-generated photographs not only has implications for how these images affect politics and society—how our perception of photographic form may or may not change based on these technologies—but also for how they affect the machine learning models themselves and their operation. Soon images will be recycled back into training sets, fueling the generation of more images. Some computer scientists have predicted that this will lead to “model collapse,” where the model effectively forgets its training and can no longer function.⁸⁰ Online commentators have begun to call this Habsburg AI after the European royal family who intermarried and thus amplified the occurrence of certain genetic traits and conditions. It remains to be seen what effect this will have on how we understand and interpret the photographic form. It is clear, however, that over the next few years AI-generated photographs and so-called natural images will likely become harder to tell apart. As they do, the complicated relationship between photography and nature grows ever more complex.

78. See “The Pope Drip,” Reddit, 24 Mar. 2023, www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/

79. See Amanda Wasielewski, “‘Midjourney Can’t Count’: Questions of Representation and Meaning for Text-to-Image Generators,” *IMAGE* 37, no. 1 (2023): 71–82.

80. See Ilia Shumailov et al., “The Curse of Recursion: Training on Generated Data Makes Models Forget,” *arXiv*, 31 May 2023, arxiv.org/pdf/2305.17493.pdf