AI IMAGING, OR, THE END OF PHOTOGRAPHY AND THE AFFORDANCES OF LATENT SPECIFICITY

Noam M. Elcott and Tim Trombley

I. INTRODUCTION

Just as artificial intelligence (AI) has irrevocably transformed photography, so too photography fundamentally changed the course of AI. Yet in its current configuration, AI imaging—that is, the generation of images through Stability AI's Stable Diffusion, OpenAI's DALL-E and ChatGPT, Google's Gemini, and other large pre-trained latent diffusion models—is distinct from photography. An AI-generated "photograph" is closer to a bar graph produced in PowerPoint than it is to a photograph produced by a digital SLR camera. (Smartphone snapshots posted to major social media sites lie somewhere in between.) Stated bluntly, from the perspective of AI imaging, "photography" is a style technically fungible with other styles, be they line drawing, cyberpunk, Studio Ghibli, or Impressionism.

Although AI imaging marks a radical rupture in the history of machine imaging, that history was and remains inextricably tied to the history of photography. This essay argues that a current and perhaps future potential for photography in the age of AI lies in the transformation of photographic

archives—including the corpora of living photographers—into latent spaces that can be sourced for the production of fully synthetic images shot through with specificities from their photographic source material.

II. THE PHOTOGRAPHIC STYLE

Even if the advancement of AI technologies slows or stalls, even if artificial general intelligence (AGI) proves to be an empty marketing gimmick, today's deep neural networks mark a rupture in the history of photography. And in so much as photography claimed an ontological link to the world, AI fundamentally alters not only our relationship to images, but also our relationship to the visual world mediated by images. So-called photographs, like all other images produced through generative AI, are data visualizations, "statistical renderings"[1] (in the apt phrase of Hito Steyerl) of highly abstracted surface patterns learned by deep neural networks. Take the following series of images (see fig. 1). At the far right is an image we would readily recognize as an early-twentieth-century portrait photograph of a woman in a forest. At the far left we recognize a computer-generated image of the same subject. But the image at the extreme right is as much a data visualization as the image at the extreme left. Indeed, they were produced with the exact same prompt: "A photo of a young woman seated in a forest by August Sander." The only difference is that the image on the left was produced using Stable Diffusion XL (SDXL), the state-of-the-art open-source generative-AI model (as of May 2024), whereas the image at the extreme right was produced with SDXL coupled with an additional model custom-trained on publicly available August Sander images. The intermediate images moving right to left mark the products of a stepwise decreased weighting of the custom model, revealing the "house style" of SDXL. We interrogate these images at greater length below. For the moment, it is essential to understand that there are no absolute distinctions among any of these images. They are all data visualizations. At no point does the image (if we pretend it is a single image) become more or less "photographic" in any substantive sense of the word.

1 Hito Steyerl, "Mean Images," *New Left Review*, 2023.

Just as Walker Evans long ago understood that there is no such thing as documentary photography, only “documentary style,”[2] today we must confront the fact that in AI imagining, there is no photography, only “photographic style.”
Or, rather, styles. As virtuosos of surface pattern, generative models can more or less successfully (depending on their underlying data and training) mimic the (digital) appearances of nearly every photographic medium, from daguerreotypes and gelatin silver through to Polaroids and smartphone snapshots (see figs. 2, 3, 4, 5, 6 and 7). From the perspective of a generative AI model, the choice between “daguerreotype” and “gelatin silver” is no different than the choice between “photograph” and “anime drawing.” AI “photography” is a suite of styles.

2 “Documentary? That’s a very sophisticated and misleading word. And not really clear. You have to have a sophisticated ear to receive that word. The term should be *documentary style*.” Evans in Leslie Katz and Walker Evans, “Interview with Walker Evans,” *Art in America* 59, no. 2 (March/April 1971): 87. See also Olivier Lugon, *Le style documentaire: d’August Sander à Walker Evans, 1920-1945* (Paris: Macula, 2001).

Fig. 1: Prompt: “Photo of a young woman sitting in the forest by August Sander”
Top left: image generated by SDXL. Bottom right: image generated by SDXL coupled with a custom model trained on publicly available August Sander images

Figs. 2, 3, 4, 5, 6 and 7: Prompt: “A person in a chair” + one of the following processes (from left to right): “daguerreotype”, “collodion wet plate”, “gelatin silver”, “Polaroid”, “digital SRL”, “iPhone”

Fig. 8: Francis Galton, *The Jewish Type*, circa 1885. Series of composite photographic portraits mounted on board, 29 × 22 cm. Galton collection, UCL Library Services, Special Collections, London. Inv.: Galton/2/8/1/11/2

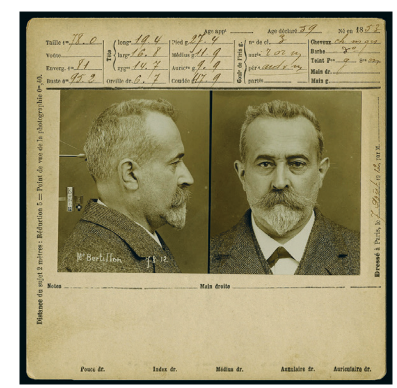

Fig. 9: Anthropometric sheet by Alphonse Bertillon, 1912. Archives of the Préfecture de Police, Paris, Bertillon 1

III. AI AND PHOTOGRAPHY

"AI photography" is as much an oxymoron as "AI Impressionist painting." Nonetheless, the histories and current configurations of AI imaging and artificial intelligence *tout court* are indissociable from the history of photography. A *longue durée* history would have to engage the advent of photography and its ever-deeper penetration into all aspects of culture through technologies such as halftone printing and digitization. At various junctures in this history, the indexical chain that tethered images to reality encountered weak links: not only recent digital manipulations but also nineteenth-century press disseminations of photographs in the form of wood engravings (xylography), barely distinguishable from handmade drawings (which they partially were) and authenticated exclusively through captions that specified "after a photograph."[3] In the years to come, unmanipulated camera-based photographs will similarly require external labeling and authentication. Unlike any of its "photographic" predecessors, however, AI-generated "photographs" are wholly untethered from reality. They are images derived from labeled images, not images of the world. The quality of the generated images is thus overwhelmingly dependent on the quality and quantity of images (and their associated labels/captions) in the underlying trained dataset. AI photorealism is impossible without training datasets containing millions of photographs. This is the superficial link between photography and AI, as it remains at the level of surface style. Other styles require other types of images, all of which—including photographs, many of which are not natively digital—must first be digitized.

The more fundamental link between AI and photography lies in a series of datasets consisting of image-text pairs, including vast numbers of photographs. For the recent AI boom, the most influential of these datasets was ImageNet (2009-).[4] Comprising over fourteen million labeled images organized into about twenty thousand categories, ImageNet aimed, in the words of its primary architect Fei-Fei Li, to "map out the entire world of objects."[5] The ImageNet Large Scale Visual Recognition Challenge, established in 2010 as an annual competition, was the testing ground for machine

3 See Bernd Weise, "Aktuelle Nachrichtenbilder 'nach Photographien' in der deutschen illustrierten Presse der zweiten Hälfte des 19. Jahrhundrets," in *Die Eroberung der Bilder: Photographie in Buch und Presse, 1816-1914*, ed. Charles Givel, André Gunthert, and Bernd Stiegler (Munich: Fink, 2003), pp. 62–101.

4 Histories and critiques of ImageNet abound. See esp. Kate Crawford and Trevor Paglen, "Excavating AI: The Politics of Images in Machine Learning Training Sets," (2019). https://excavating.ai; Emily Denton et al., "On the Genealogy of Machine Learning Datasets: A Critical History of ImageNet," *Big Data & Society* 8, no. 2 (July 2021).

5 Li, quoted in Denton et al., "On the Genealogy of Machine Learning Datasets: A Critical History of ImageNet."

6 Up to this point, the moniker "artificial intelligence" was attached to a competing paradigm. Only with the continued success of neural networks did "deep learning" usurp the mantle of "artificial intelligence." See Chris Wiggins and Matthew L. Jones, *How Data Happened: A History from the Age of Reason to the Age of Algorithms* (New York: W. W. Norton, 2023), esp. pp. 175-195.

vision in the form of image classification and object detection. In 2012, AlexNet (designed by Alex Krizhevsky in collaboration with Ilya Sutskever, who eventually co-founded OpenAI, and Geoffrey Hinton, a "godfather" of deep learning) won the competition by a wide margin and established data- and compute-intensive neural networks as the reigning paradigm for artificial intelligence.[6] Since then, ever larger sets of labeled images, with massive numbers of photographs, have enabled ever more powerful image classification, object detection, and finally image generation through natural language prompts.

IV. LATENT SPACES, OR, PHOTOGRAPHIC ARCHIVES IN THE AGE OF AI

The practice, history, and theory of photography is inextricably bound up with that of archives.[7] As much as the classification and generation of images through deep neural networks marks a rupture with prior photographic practices, there are formidable continuities to which we must attend. Allan Sekula's "The Body and the Archive" (1986), likely the most influential essay on the subject, explored the practices, ideologies, and media systems that enabled the late-nineteenth-century fusion of photography, archives and statistics—the same triumvirate that undergirds AI imaging—in the criminal identification system of Alphonse Bertillon and the composite portraiture of Francis Galton.[8] The contemporary connection to Galton's composite photographs is readily grasped and seen. In order to arrive at an ideal image of "the criminal type" or "the Jewish type"—Galton was a seminal exponent of eugenics—Galton rephotographed, say, ten different images of a given "type," each at 1/10 exposure. The resultant image, so he claimed with preposterous confidence now ludicrously echoed in certain AI quarters,[9] captured the ideal features of the underlying type (fig. 8).

The contemporary relevance of Bertillon is less obvious but more profound. Working for the Paris police, Bertillon recognized the importance of photography in identifying recidivists and the impossibility of locating police photo-

7 In addition to sources referenced below, see, for example, Rosalind Krauss, "Photography's Discursive Spaces: Landscape/View," *Art Journal* 42, no. 4 (Winter 1982): 311–19; Molly Nesbit, *Atget's Seven Albums* (New Haven: Yale University Press, 1992); Robin Kelsey, *Archive Style* (Berkeley: University of California Press, 2007); Estelle Blaschke, *Banking on Images: The Bettmann Archive and Corbis* (Leipzig: Spector Books, 2016).

8 Allan Sekula, "The Body and the Archive," *October* 39 (Winter 1986), pp. 3-64. On Galton as a precedent for AI portraits, see also Daniël de Zeeuw and Abraham Geil, "'This Person Does Not Exist': From Real Generalisation to Algorithmic Abstraction in Photographic Portraiture," in *Reconfiguring the Portrait*, ed. Abraham Geil and Tomáš Jirsa (Edinburgh: Edinburgh University Press, 2023), pp. 43–60.

graphs in the archives when criminals regularly employed aliases and changed their appearances. Accordingly, he relied on a series of eleven bodily measurements that he regarded as constants in any adult body. For Bertillon, Sekula writes, "the mastery of the criminal body necessitated a massive campaign of inscription, a transformation of the body's signs into a text, a text that pared verbal description down to a denotative shorthand, which was then linked to a numerical series"[10] (see fig. 9). In deep neural networks—as in the Bertillon system—every image must be converted into arrays of numbers organized into matrices in order to be processed. The Bertillon system, Sekula argues, "can be described as a sophisticated form of the archive. The central artifact of this system is not the camera but the filing cabinet."[11] In so far as deep neural networks can be described as sophisticated forms of the archive, the central artifact of these systems is not the camera, prompt, or photograph, but the latent spaces of deep neural networks: high-dimensional spaces in which abstract data patterns and relations are embedded.[12]

Crucially, deep neural networks are not archives in any traditional sense. They are not structured for the retrieval of specific images. Quite the contrary. Whether in image classification or generation, the aim is to identify or create images that are similar but not identical to the images in their data training sets. Accordingly, latent spaces comprise abstract data patterns, not specific photographs.[13] In a phrase that conspicuously anticipates the related concept of data embedding, Sekula concludes: "Bertillon sought to *embed* the photograph in the archive. Galton sought to *embed* the archive in the photograph."[14] Sekula's formulation need only be tweaked to speak to the newfangled conditions of the archive as latent space: abstracted data patterns (analogous to Galton's composite photographs) are embedded in mathematical latent space (analogous to Bertillon's system).

9 Among many important critiques, see Jake Goldenfein, "The Profiling Potential of Computer Vision and the Challenge of Computational Empiricism" *(FAT* '19: Proceedings of the Conference on Fairness, Accountability, and Transparency*, Atlanta, GA, 2019); Luke Stark and Jevan Hutson, "Physiognomic Artificial Intelligence," *Fordham Intellectual Property, Media and Entertainment Law Journal* 32, no. 3 (2022): pp. 922-78. See also the recent project by Clément Lambelet, which aims to criticize facial recognition and policing technologies through wholly Galtonian procedures that go unnamed and unquestioned. Clément Lambelet, "The Mathematics of Regression," *foam*, 66 *Missing Mirror* (2024): pp. 47–56.

10 Sekula, "The Body and the Archive," p. 33. See also Roland Meyer, "Operative Portraits, or How Our Faces Became Big Data," in *Reconfiguring the Portrait*, ed. Abraham Geil and Tomáš Jirsa (Edinburgh: Edinburgh University Press, 2023), pp. 27–31.

11 Sekula, "The Body and the Archive," p. 16.

V. LATENT SPECIFICITY

Although Stable Diffusion and other state-of-the-art latent diffusion models can readily generate images that are indistinguishable from photographs, those images are unlikely to be mistaken for the output of specific photographers—at least not without fine-tuning the models. The process of fine-tuning seeks to emphasize particular characteristics of images generated by a model or define concepts previously omitted. This process can take many forms, from retraining the entire model to establishing a single persona through the use of comparatively few images. In pursuit of stylistic tuning, the most proven approach is the use of a Low-Rank Adaptation (LoRA). LoRAs target narrowly defined concepts within the model for retraining in order to efficiently modify parameters at the most influential portion of the neural net. Crucially, this process does not alter the underlying latent space of the model, but rather constructs a smaller, orthogonal latent space that nudges the rendering of features toward those represented in the new latent space and the new image patterns embedded therein. Although no more "authentic" or "photographic" than any other AI-generated outputs, the images produced through a well-trained LoRA exhibit the features and qualities of the original photographs to an unprecedented degree. Stated more simply: they are (nearly) indistinguishable from the images produced by the source photographers. We call this *latent specificity*.

The specificity generated by our models is latent in at least two respects. On a technical level, our LoRAs do not add more images to the SDXL training data but instead intervene in the navigation of its latent space, which contains no images, only their abstract patterns and relations. LoRAs operate exclusively in the latent space processed by neural networks, not the pixel space visible to human beings. Second, the specificity that we generate is latent in the archives and must be learned by our models. It is not a feature that we *find* but rather a feature we *produce* through the selection, labeling, and processing of images.

12 In certain neural network architectures, "latent space" has a precise technical meaning. It is also often used, as here, in a less technical sense to refer to the unseen layers of a network and the embedding of data patterns too abstract for human comprehension. For more on latent spaces, see Noam M. Elcott, "Response to Questionnaire on Art and Machine Learning," *October,* 189 (summer 2024).

13 "Memorization" (technically: "overfitting") is a problem, not a goal, for deep neural network architecture.

14 Sekula, "The Body and the Archive," p. 55. Emphasis added.

Latent specificity must be defined in opposition to two related frameworks. First, the kitschy images produced through generic generative AI (introduced above and explicated below). The second is modernist medium specificity. In Clement Greenberg's classic formulation: "It is by virtue of its medium that each art is unique and strictly itself. To restore the identity of an art the opacity of its medium must be emphasized."[15] Greenberg opposed these self-reflexive artistic processes to what he deemed kitsch: "If the avant-garde imitates the processes of art, kitsch … imitates its effects."[16] Latent specificity makes no claims on medium specificity. Our LoRA images are no more authentic an expression of AI, neural networks, or latent space than the kitschy, generic images produced by SDXL. For Greenberg, imitation is the problem for which medium specificity is the solution. For generative AI, it is imitation all the way down. The only question is whether that imitation is specific or generic, a difference better seen than described.

15 Clement Greenberg, "Towards a Newer Laocoön," in *Clement Greenberg: The Collected Essays and Criticisms, vol. I*, ed. John O'Brian (Chicago: University of Chicago Press, 1986), p. 32.

16 Clement Greenberg, "Avant-Garde and Kitsch," in *Art and Culture* (Boston: Beacon Press, 1989), p. 15.

The necessity and potency of LoRA training can be clearly observed in the following images. We have created LoRAs for two distinct sets of photographers. The first was trained on roughly one thousand publicly available photographs of peasants by August Sander, a leading portrait photographer and avant-garde artist of Weimar, Germany. His monumental effort to photograph and classify representative types from every corner of German society, *People of the Twentieth Century* (c. 1910s-1950s), included a full volume dedicated to peasants.[17] The second was trained on thousands of photographs scraped from the Library of Congress' repository of Farm Security Administration (FSA) photographs. Between 1935 and 1944, the FSA and related offices commissioned extensive documentation of American life by many of the now-canonical documentary photographers: Walker Evans, Dorothea Lange, Arthur Rothstein, Gordon Parks and many others. We trained our two FSA LoRAs on roughly one thousand images by Lange and over fourteen hundred images by Parks. Notably, Parks was the only African American photographer employed as part of the project and his FSA photographs include his all-time most famous: Ella Watson, also known as "American Gothic" (1942).

17 August Sander, *People of the Twentieth Century* (New York: Aperture, 2022). On Sander, classification, and AI, see also Noam M. Elcott, *Photography, Identity, Status: August Sander's People of the Twentieth Century* (Chicago: University of Chicago Press, forthcoming).

As made patent in the comparisons below, the gap between the base model (SDXL) and our LoRA is vast for "August Sander" images and comparatively subtle for "FSA" images (scare quotes designate AI-generated images and should not be confused with original photographs by the named photographers.). We believe the differences are a product of the number of relevant images in the training dataset, especially LAION-5B and its English-language bias, and the relative robustness of the latent spaces encompassing August Sander and FSA photographers, respectively. Without access to LAION-5B (which is no longer open to scholars) or extensive computing power and related resources (as recently employed by Anthropic and OpenAI[18]), these remain speculations. For better and for worse, the images largely speak for themselves; we will merely provide explanatory captions.

18 See Anthropic-TK and OpenAI: https://cdn.openai.com/papers/sparse-autoencoders.pdf.

Here we contrast the output of the stock configuration of SDXL (see fig.10) against our LoRA (see fig. 11) for "August Sander." The left image is legible as a girl in the forest; but "August Sander" is reduced to a pastiche of airbrushed sepia tones. The model has some base-level awareness of Sander's timeframe but nothing approaching stylistic signifiers. What's more, the girl is a generic adorable child, better suited for an advertising label than Sander's *People of the Twentieth Century*. The image on the right, by contrast, could readily pass as a digital image of an original Sander print.

SDXL's incapacity to produce something resembling a Sander photograph clarifies an architectural paradigm within the training of models. Although Stable Diffusion publicly releases its models for modification by the development community, they are indisputably products. The goal of this suite of products is to create images that will suit the commercial and aesthetic sensibilities of the broadest possible audience. Accordingly, Stability AI tunes the model before its release to avoid the production of images that could become liabilities from public relations or legal perspectives, while also engendering an identifiable "house style" that can be associated with their brand. The process naturally favors a suppression of specificity at the edges of

Fig. 10, 11: Prompt: "A photo of a girl child with a bow in her hair standing in a forest by August Sander". Left: image generated by SDXL. Right: image generated by LoRA trained on works by August Sander

Figs. 12, 13, 14: Prompt: "A photo of a farmer by Dorothea Lange". Left: image generated by a trade web interface. Center: image generated by SDXL. Right: image generated by LoRA trained on works by Dorothea Lange

Figs. 15, 16, 17, 18, 19: Prompt: "A photo of a farmer by Dorothea Lange"

the model's latent space in favor of generically acceptable results. By catering to the broadest audience, the training of models parallels the erasure of individuality in Galton's composite photos.[19] Individuality is stripped to facilitate the most economically viable, and blandly proficient, mean.

Images 12, 13, and 14 show the output of a single prompt interpreted through a commercial web interface (left), the output of SDXL (center), and our Dorothea Lange-trained LoRA (right).[20] The image on the left is a blandly attractive composition with no direct relation to Lange. It would be more at home on a mass market wine label from California than in the files of the FSA. Although the middle image demonstrates a basic conception of the world of Dorothea Lange, it is a superficial resemblance emblematic of the blending of her work within the larger and unfocused context of the model's latent space. Our LoRA image, by contrast, exhibits many of the peculiarities of Lange's images in the FSA archive: the visible edge of the film and matte, the overexposure at the bottom, the strong sunlight and deep shadow, the frayed edges of the hat, the flat dusty landscape. None of these features were named in our prompt; instead, they were embedded in the LoRA's latent space and integrated into the final image. The consistent results are images difficult to distinguish from Lange's originals. Study the images below (see figs. 15, 16, 17, 18, and 19). Can you distinguish Lange from "Lange"?[21]

Image 20 on the left is the output of SDXL in response to the prompt "a photo of a worker at a construction site by Gordon Parks." Even within a relatively well-defined latent space around the photographs of Parks, what initially appears like a compelling image can be shown to be a generic iteration of his style, which is rendered with far more specificity in the right image (our Parks LoRA) (see fig. 21).

The series of images here comprise iterations of the same prompt with no LoRA (see fig. 22) through full-strength LoRA, with 10% stepwise intensifications in between (i.e., the middle image shows the LoRA at 50% strength). Comparing the SDXL and LoRA images reveals several dimensions of Parks' distinctive style during his time with

19 Galton: "The blur of their outlines, which is never great in truly generic composites, except in unimportant details, measures the tendency of individuals to deviate from the central type." Sekula is quick to note that, although it is more conspicuous at the periphery, blurring is in fact spread over the entirety of the image. "Only an imagination that wanted to see a visual analogue of the binomial curve would make this mistake." Sekula, "The Body and the Archive," pp. 47–48.

20 The commercial web interface is: https://stablediffusionweb.com/.

21 The second and third images starting from the right are by Lange. The rest were generated by our Lange LoRA with the prompt "a photo of a farmer by Dorothea Lange."

the FSA. In place of the perfect exposure and *Life*-magazine-colored hues of SDXL's generalized embedding, we witness the emergence of black-and-white film grain texture, imperfect exposure, and, most intriguingly, a compositional shift as the LoRA reaches full application. Our Parks LoRA regularly produced subjects at a heroic elevation—a composition evident throughout the training images but unaddressed in his nearly contemporaneous reflections on photographic portraiture.[22] (Compare figs. 23, 24) When plotted along this image axis, the generated construction worker rises as we assume the position of "Parks'" camera moving from parallel with the worker to eventually foregrounding him against the skyline.

In the absence of race and gender specifications, most models, SDXL included, have a tendency to default to white, male subjects for otherwise "neutral" subject matter—as seen, for example, in the images at the start of this essay, which were generated by SDXL from the prompt "a *person* sitting in a chair" (see figs. 2, 3, 4, 5, 6 and 7). Bias remains a major unsolved problem for generative AI, except for those companies that have abandoned the struggle altogether. Our LoRA models demonstrate that the problem is solvable, at least at smaller scales. As Kate Crawford has argued, "improvements in AI will require putting a lot more care and thought into how data is collected and curated."[23] FSA photographs represent one such carefully and thoughtfully curated datasets. For all its inadequacies, the FSA aimed "to introduce America to Americans,"[24] an ambition that included systematic efforts to photograph a diverse and inclusive range of its inhabitants. These efforts are especially pronounced in the photographs of Parks and Lange. The five Lange/"Lange" images (see figs. 15, 16, 17, 18 and 19) —which were not cherry-picked—include two photographs by Lange and three generated with our Lange LoRA. Our purposefully vague prompt was "a photo of a farmer by Dorothea Lange." The results include compelling, historically consistent representations of women and African Americans to a degree rarely achieved through standard Stability AI, OpenAI, Google or other models.

22 See Gordon Parks, *Camera Portraits: The Techniques and Principles of Documentary Portraiture* (New York: Franklin Watt, 1948). The volume includes a photograph of Roy Stryker, the director of the documentary photography program of the FSA.

23 "9 Ways to See a Dataset: What's at Stake in Examining Datasets?," 2023, accessed June 20, 2024, https://knowingmachines.org/publications/9-ways-to-see/essays/9-ways-to-see-a-dataset.

24 Roy Emerson Stryker and Nancy Wood (eds.), *In This Proud Land: America 1935-1943 as Seen in the FSA Photographs* (New York: Galahad Books, 1973), p. 9. But see also Lisa H. Kaplan, "'Introducing America to Americans': FSA Photography and the Construction of Racialized and Gendered Citizens" (PhD Bowling Green State University, 2015).

Figs. 20, 21: Prompt: "Photo of a workman on a building site by Gordon Parks". Left: image generated by SDXL. Right: image generated by LoRA trained on works by Gordon Parks

Fig. 22: Prompt: "A photo of a worker at a construction site by Gordon Parks". Left to right: iterations of the same prompt with a gradual 10% intensification of the LoRA

Fig. 23: Gordon Parks, *New York, New York. Policeman no. 19687*, 1943. Image digitized from the original negative. Library of Congress Prints and Photographs Division Washington, D. C. Inv. n°: 2017851528

Fig. 24: Gordon Parks, *Daytona Beach, Florida. This boy, who is a NYA (National Youth Administration) student, receives thirty-five cents an hour for chopping wood. By doing this he is able to pay his tuition fees*, 1943. Image digitized from the original negative. Library of Congress Prints and Photographs Division Washington, D. C. Inv. n°: 2017845058

VI. CONCLUSION: THE AI WORLD PICTURE

Among the reasons generic AI images look so "real" is because the "reality" proffered by mass and social media has become so generic. A parallel logic haunts our LoRA images. We can mimic past photographers with ever greater precision; yet each synthetic image is a monument to what will be lost in the age of AI imaging.

It is possible that future AI models will be trained on synthetic data. For the moment, AI imaging remains reliant on photography, its data, its archives, its styles, and the afterglow of its fragile ontology. The capacity to synthesize "photographic" images with newfound specificity may prove to be the *coup de grâce* for the medium: why invest the enormous time and effort required of good photographs when a good-enough "photograph" is just a prompt away? At this moment, however, latent specificity opens new avenues into archives presumed dead and into imaging systems presumed closed.

If "photography" is merely a suite of styles within generative AI, AI imaging is the sublation of photographic style: the surface patterns of billions of photographs; the sprawling technological apparatuses feeding off the surface patterns of billions of photographs; and ultimately the visible worlds constructed according to these sprawling technological apparatuses. Artificial intelligence has brought about the abolition and consummation of photography as world picture.