mt

Photography at a Standstill

GEOFF COX

London South Bank University, UK

Abstract

What has replaced the still photograph are dynamic and distributed image-assemblages that unsettle received notions of space-time — no longer limited to traditional representation and not necessarily even visual. When it comes to computer vision for example, the descriptor photography seems largely redundant (despite deep learning computer vision systems being trained on large datasets of photographic images), and so too the tired metaphor of the eye that once supported its theories and practices. What is at stake here, as ever, is a kind of 'seeing' (if we continue to call it that) that makes clear what is visible, sensible and knowable, and crucially also what is not. We might call this seeing algorithmically, or seeing like a dataset, or perhaps even seeing like an infrastructure, comprised of fake images that render fake history. The logic of this invokes the complex notion of 'image is dialectics at a standstill', encapsulating a constellation of possible outcomes. To what extent is the radical potential that Benjamin once foresaw in montage applicable to image-based AI given that it seems less an instrument to imagine a qualitatively different future but simply more of the same. What can be seen is not so much representational nor photographic but latent traces of material relations and infrastructures that render historical experience in compromised form.

Keywords

Image politics, standstill, latency, history, image-based AI

> It's not that what is past casts its light on what is present, or what is present its light
> on the past; rather, image is that wherein what has been comes together in a flash
> with the now to form a constellation. In other words, image is dialectics at a standstill
> (Benjamin, 1999: 462).

Introduction

The practice and theory of photography appears to be at a standstill in its most literal sense. What has replaced the still photograph are dynamic and distributed image-

assemblages that unsettle received notions of space-time — no longer limited to traditional representation and not necessarily even visual. Even in the case of computer vision, the descriptor photography seems largely redundant despite the fact that deep learning computer vision systems are trained on large datasets of photographic images. So, too, does the tired analogy of eye/lens that once supported its theories and practices seem only relevant to those already deeply invested in the maintenance of its institutions, as such seeing can no longer be thought of as singular or indexical truth or reality, as it was often assumed to be, but indicative of a wider need to manifest authority and power through distributed forms of machine seeing. It follows that we might instead refer to 'seeing' algorithmically, or like a dataset, or perhaps seeing like an infrastructure, akin to what Adrian Mackenzie and Anna Munster call "platform seeing" to describe a new way of seeing distributed through data practices and machinic assemblages (2019: 3). What is at stake here is a 'way of seeing', if we continue to call it that — or perhaps a "dialectics of seeing" (Buck-Morss, 1995) — that makes clear what is visible and sensible, and at the same time what is not, thus raising the question of how these 'invisual' relations of production can be made knowable.[1] Furthermore, how to account for the historical dimension of computational statistics and the logic of probability that has largely replaced representational forms?

The ambition of this essay is to explore these ideas with reference to Walter Benjamin's proposition "image is dialectics at a standstill" (1999: 462), introduced in unfinished essays between 1927 and 1940. It indicates a moment of transition in which "time takes a stand [*einsteht*] and has come to a standstill" (2003: 369) and draws attention to the experience of a determinate historical relation, between 'now' and 'then' to put it simply, or, in Benjamin's words, "the relation of what-has-been to the now is dialectical" (1999: 463). Time is at a standstill such that the past enters into a *constellation* with the present — "in a flash with the now to form a constellation" as described — objectively interrupting the mechanical temporal process of historicism (1999: 462, 463), performing a "dialectical leap" in the "open air of history" (2003: 395). If once the still photograph could make a claim to interrupt the flow of events and historical-narrative structures in this way, this essay asks whether it continues to do so under contemporary conditions.

Much contemporary art has operated self-reflexively in this manner to negate naturalized time flows, reveal what has been repressed in the historical past, and thereby draw attention to a present that is contingent with historical-material conditions. An example of this tendency is Jeff Wall's staged large-scale photographs which appear to be part of a larger event, as in the case of *A Sudden Gust of Wind (after Hokusai)* (1993) that creates the illusion of the *decisive moment* and reworks (art) history (specifically Hokusai's woodcut *Travellers Caught in a Sudden Breeze at Ejiri* (c.1832)). The aesthetic register of the composite image encapsulates this sense of nonlinear time, part of a wider narrative that we don't have access to directly and suggestive of a historical reality that is able to be recompiled through montage techniques. But what of this now? In times of political stagnation, to what extent does the logic of the image at a standstill still hold? If networked images have replaced the single photograph as cultural form (Dewdney and Sluis, 2023), set in motion through their inherent reproducibility and liberated from aura and hierarchies of value — akin to what Hito Steyerl has named the "poor image" (2009) and, in turn, "mean image" (2023) — then, what qualities of stillness remain operative, if at all?

Fig. 1: Image generated by the prompt "A Sudden Gust of Wind (after Jeff Wall)", using Stable Diffusion and photographic filter.

Latency

Time cannot be seen or experienced directly; it is perceived through the various ways it is represented, measured, calculated, only made knowable through various techniques, media, infrastructures, and aesthetic regimes. In this sense there is an inherent relation between standstill and latency through which time is produced and

made experienceable. Latency stands as a figure through which something has not yet appeared, that which "connects a past event and its future manifestation with the tendencies of a still blind present" (Schwarte, 2019: 95). As such it has motivated the political imagination.

More pragmatically, the latency involved in any real-time technical process draws attention to the parallels between developmental progress on the one hand and slowing things down to the point of stillness (or degrowth) on the other. There is an echo of the 'halting problem' of computation here — the problem of determining whether a computer program will finish running or continue to run forever, and that establishes the idea of the Turing machine on which computing is based. In such cases, Wolfgang Ernst describes the "delayed present" (2016) in which it seems that the present can never quite manage to settle in time. Delay takes place in the operations of machine memory, which is held in a state of latency, as evident in the case of buffering large video files (indicated by the spinning of the on-screen 'wheel of doom' or irregularity of the progress bar). This parallels the ways in which human cognition is always slightly behind the present moment in spontaneous actions or premonitions; in both cases, human and machine, this is a storage and processing issue. With contemporary developments in real-time computation, and where the dominant tendency is to speed up rather than slow down, this is a "time-critical" issue, and for Ernst, it is as if "the present no longer has time to take place" (2016: 36).

Concerning photography and latency, the temporal register of Roland Barthes's *Camera Lucida*, published in 1980, shifts from an engagement with the "past ('this has been') to the deferred present ('this has just been')" (Ernst, 2016: 40). Although Barthes describes the photograph as derived from an "immobilization of time" (1984: 91), it is clear that it has always moved across temporal registers, and its perceived characteristics such as an indexical relation to reality and ability to freeze time were always imaginary — just as film stills are part of a larger narrative, or screenshots are decontextualized from the inner complexity of the computer. Yet many of these founding myths associated with photography, evident in a contradictory term like 'digital photography', are repeated if not amplified in image-based AI that builds upon photographic imaginaries. Along with the integration of AI into consumer and professional software alike — from OpenAI's DALL-E to MUSE AI's Midjourney

and Stability AI's Stable Diffusion — it would be tempting to conclude that the use of the term photography, as well as post-photography, is inadequate to capture the complexity of algorithmic operations. Human and machine seeing, as well as image synthesis, have been transformed by AI, as we shift from indexicality and truth-value to statistics and probability distribution.

But perhaps it's not quite as clearcut as this. As previously mentioned, computer vision systems are trained on large datasets of photographic images and founded on "the epistemic and aesthetic affordances of photographic practice" (Chávez Heras and Blanke, 2020). Daniel Chávez Heras and Thomas Blanke argue that image-based AI is materially and conceptually connected to optical regimes of visibility (the eye/lens), and that photography remains a latent force in computer vision and its representational logic of visuality. This position is underpinned by the assumption that photography has always been based on calculation, as in the case of the numbers and metrics involved in focal length, exposure, aperture and so on, and consequently that photographic theory is necessary for a deeper understanding of computer vision systems and their particular ways of (machine) seeing an external reality. Similarities can also be drawn from the use of AI in the latest version of Photoshop and mobile phone applications which establishes further parallels between datasets and the place of the archive in the history of photography. Whilst largely sympathetic to these ideas, photography here serves more as a trope to supplement a politics of the image delimited by what is considered to be optical or representational. This is where it seems especially important to understand how the reality that appears to be depicted by such systems unsettles received notions of space-time. What is latent here is not photographic realism or the optical unconscious so much as the spatial-temporal structures through which something is made visible, sensible, and knowable. It would appear that the inherent ambiguity of the image has been accelerated by generative computational processes that mimic the representational image (inasmuch as it looks like a photograph) and yet reproduce an external reality that is more evidently synthetic (what we might call computational realism). Moreover, and to be clear, image-based AI does not represent the world but further reproduces its representations in compromised form.

Latent space

What is referred to as *latent space* offers further insight into this synthetic reality. It refers to a mathematical space which maps (or embeds) what a neural network has learnt from training images and is thus able to position items that resemble each other in close proximity to one another. In his essay 'Latent Deep Space', Fabian Offert explains how "latent space sampled by a generative adversarial network could be described as an analogical space where the produced literal images are also analogical 'images' which, *as a set*, constitute an analogy of the machine's perspective on the world" (2021). This worldview, or ideology if preferred, follows the prescribed logic of what the machine knows, based on the given particularities of a specific image dataset that informs its learning, and as such departs widely from indexical reality. Offert goes on to explain how GANs (generative adversarial networks) — a type of neural network that approximates the probability distribution that defines a set of images by means of an interplay between two deep convolutional neural networks — are not only opaque systems on a technical level but also in terms of knowledge production. Given the 'adversarial' interplay between the neural nets — the generator and discriminator — it would be tempting to draw analogies here between the operations of GANs and the inherent contradiction between two opposing forces associated with dialectical materialism.[2] Yet the operations of a GAN are clearly not progressive in this sense, nor are neural networks more broadly, as they pass inputs through, and correlate across, hidden layers and future training epochs. Any predicted output from such processes becomes correlative with its past, formed out of the data on which it was trained in the first place, producing its own particular model of historical time based on its source data.

Just as the determinism of clock-time was once symptomatic of industrial production, the ways in which different kinds of time exist contemporaneously across different space-times as part of global capitalism now reproduce an experience of time compressed into a successive form to provide even more standardized subjectivities (The Invisible Committee, 2017: 21). Images distributed across global networks construct a different kind of historical subject and are part of a conception of history which is based on machine-time (so-called Unix-time that counts time in seconds since January 1st, 1970 at 00:00:00 UTC). This example demonstrates how the control of

time continues to play a central role in the control of planetary networks and infrastructures, whether centralized or not, as with blockchain technology, for example, in which data is stored in blocks that are linked together in a chain. Each block in the network typically contains a cryptographic hash of the previous block, such that new transactions are timestamped to verify their validity, and so on, as part of the chain (of events), rendering historical subjects fungible (as were slaves). More to the point, behind this operational logic is the principle that no one can change the blockchain or make changes to any data in the past once it is recorded in a block and assigned a cryptographic hash (although they are not entirely immune to attacks). This is important as, although a distinct technology, blockchain is often linked together with AI to improve levels of trust, transparency, and security — through the use of NFTs and smart contracts, wallets, and ledgers, for example, and used within DAOs (decentralized autonomous organizations) to improve decision-making processes and probabilistic structures. In general, though, under the immutable logic of the blockchain, there seems to be one true if not totalitarian version of history (Bowden, 2018), despite the radical decentralized claims associated with the technology (and notwithstanding decentralization as a model of power in itself). History is rendered immutable and is *effectively depoliticized* along with the political imaginary associated with latency. By extension all images produced in this way are evidently 'fake' (so-called 'deep fakes') even if they look (indexically) 'real', and although GANs and diffusion models seem to be open to latent possibilities, what is generated in effect is historically determined. To put it simply, these are fake images that render fake history.

Standstill

In the opening passage of Benjamin's (final) essay 'On the Concept of History' [*Über den Begriff der Geschichte*], written (in 1940) in the midst of fascism, historical materialism is introduced as the chess-playing automaton ('The Turk') that wins every time (2003: 389).[3] The relative success of machines at playing games is invoked, especially playing chess, a recurring motif in a history of AI to draw analogies between machine intelligence and cognitive development in humans or claim the superiority of machine intelligence over human intelligence (at least for certain tasks like pattern recognition). The term 'machine learning' itself was coined by Arthur Samuel in 1959 during his

game development research at IBM (1959: 210-229) to demonstrate this ability of 'learning' from experience and predicting outcomes based on statistical methods that identify patterns and relationships in data. Concerning the figure of the chess-playing automaton, the perceived autonomy of the machine is revealed to be fake, as the chess pieces are revealed to be guided by a small person hidden in the mechanism. The parallel to machine learning, where the shadow labour of people remains mostly hidden, is made explicit in the naming of Amazon's Mechanical Turk, a crowdsourcing website to hire 'crowdworkers' (or Turkers) to perform discrete on-demand tasks that computers are unable to do as well, such as the labelling of large image datasets. The fuller knowledge of the apparatus — and more broadly the ways in which the "hidden layers of neural networks also hide the reality of human labour, as well as the absurdity of the tasks performed" (Steyerl, 2023) — provides access to consciousness of conditions or what Hito Steyerl has called the "the means of mean production" (2023). The success of the chess-playing automaton or Turker is contingent on the ability to gain control of the technology, otherwise they remain passively integrated into their circuits of precarity — providing an alienated precarious workforce for the production of precarious dataset images. In Benjamin's allegory, the dynamic of history is revealed to be fake, and the unfulfilled time of the present only be activated by social struggle. For Benjamin it is the dialectical image that is able to activate this political consciousness and action.

More to the point for this essay, the politics of standstill remains a significant issue for how we perceive and experience time in both human and machine operations. Even standing-still can be a form of action, let's not forget, as demonstrated in the tactics of social movements and protests in the form of nonviolent direct action. A pertinent example here is the image of so-called 'Tank Man' standing in Tiananmen Square in June 1989, arm raised in defiance of the army tanks moving to clear crowds of students and workers who had gathered for weeks of protest to advocate for social reform. In an apparently simple act, standing-still takes on symbolic power, as opposed to the physical violence of the moving objects (akin to the 'action-less action' of Daoism[4]). That this image has also been so widely repeated, reproduced, and re-enacted adds further emphasis on the performative power of being dynamically still (with its corollary in the blank sheet of paper) and its potential to be animated at any time.[5]

The image of/at a standstill seems to mark this dialectical relation with movement and social movements, to further demonstrate the latent potential for the transformation of social reality which is today marked by a crisis of political imagination and image politics alike (Lund, 2019). This perceived incapacity to break from existing contradictions seems to be mirrored in the inability of photography to ignite change or reveal historical conditions.

Nowness, then and now

Fig. 2: Image produced from the "angel of history" description from 'On the Concept of History' as a prompt for Stable Diffusion, with photographic filter. See footnote 6 for full quote.

The infamous angel of history, looking backward while being driven forward by the storm of progress into the future (Benjamin, 2003: 392),[6] has its own history of reception and reinvention. Kodwo Eshun writes how Afrofuturism has woven together Black 'postslavery' experience and the estrangement of science fiction (2003) to decolonize the historical present. In *The Last Angel of History*, a film released in 1995 by The Black Audio Film Collective, a team of African archaeologists from the future excavate a museum from their past containing ruined documents and other materials, assembling counter-histories that contest the colonial archive from the position of the racial subject who has been locked out of official histories. Ariella Aïsha Azoulay's idea of "potential history" resonates with this, in recognizing that the imperial foundations of knowledge are formed out of the constituent violence of sovereign regimes (2019: 57) — referring to Benjamin but in this case his essay 'Critique of Violence'.[7] The photographic archive, to Azoulay, is less a way to preserve the past than to share the

present and allow others to continue to interact with it, further qualified in a footnote in which Benjamin states the past "carries with it a secret index by which it is referred to redemption" (Benjamin, 2003: 390; Azoulay, 2019: 234). In other words, it is the historical present that allows for reparation and reconciliation from within the bounds of its own inner constitution.

To Benjamin only dialectical images can be genuinely historical in this way as they contain their own temporality which he calls *jetztzeit* — commonly translated as *nowness*, or "time filled by the presence of the now" (Benjamin, 2003: 395). Nowness is thus the recognition of the potential for a new beginning in the present moment, a negation of progressive time to disrupt its flow, "to blast open the continuum of history" (2003: 396) and generate something truly emancipatory. These remain evocative references but can be easily misconstrued or fetishized in ways that overlook their political potency, according to Peter Osborne. He is referring to "often-repeated citations of Benjamin, for example, ('seizing hold of memory as it flashes up in a moment of danger' […])" that run the risk of "tragic intellectuality, a leftism marooned by history, instead of enactments of that latently dynamic, action-generating stasis to which their imagistic form aspires" (2019: 127). In other words, there is a tendency to fall back into historicism (regarding a history of Benjamin's writing) rather than maintain a critical interruption of its motor of progress — which if unchecked leads to totalitarianism. In connection to the image at a standstill, as Osborne has remarked, it seems almost impossible to say anything new that hasn't already been said, or even if there was something new to say, then for this not to fall into the ideology of newness which in itself represses the critical potential of the historical present. Osborne goes onto explain how "theoretical forms and motifs of *image* and *action* […] are in no way exempt from the problems posed by the temporal structures that they theorise" (2019: 125).

This problem associated with academic novelty is not simple of course, as it is not possible to differentiate from the temporality of capitalism itself and the valorization of all things new in the name of progress, with the hype around AI a case in point. Nowness now seems to collapse into an inert (non-dialectical) presentism rather than offer emancipatory potential. This perceived inertia is characterized by the anonymous radical leftist collective The Invisible Committee, in a passage that echoes Benjamin,

who warn against thinking about the future, and consider this to be a way of nullifying the ability to act in the present: "The current disaster is like a monstrous accumulation of all the deferrals of the past, to which are added those of each day and each moment, in a continuous time slide." To them, "life is always decided now, and now, and now" (2017: 17). To reiterate, nowness is not a transition between past and future but a distinct temporality in itself that allows for a re-composition of history and the release of its latent potential. So, the question is how to conceive of nowness now, in the context of perceived non-delayed correspondence between actions and their effects, between incoming data and its output. To what extent is the radical potential that Benjamin foresaw in the image at a standstill applicable to image-based AI given that it seems more ready to *implode* than explode, not an instrument to imagine an open horizon of possibilities but simply more of the same? This includes the realization of the climate crisis and impended planetary catastrophe to which AI contributes, for example.

On the once radical idea of the image at a standstill, Osborne thinks what is missing from much commentary is a stronger account of the structures that render historical experience. He points to the effects of what he calls "the distributed image" (2018: 139) which produces new spatial and temporal forms and new social relations. Even early photography falls into this category by way of its reproducibility, yet clearly computational image forms and the influence of network culture extend this distributed-ness or networked-ness even further, on a planetary scale. More importantly for the purpose of this essay, the image cannot be reduced to its stillness in a conventional sense but rather its ability to mediate the opposition of aesthetics and logic (2018: 138). Digital technologies render the image explicitly caught in these relations — what Osborne calls a "*distributive* unity of the relations between a materially embedded virtuality and an infinite multiplicity of possible visualisations" (2018: 139) — and this makes explicit the ontological structure of the image. Video imagery makes a good example, as a shift from a structural montage of elements as in the case of film — "the dialectic of movement and stillness" as Laura Mulvey put it (2006: 12) — to a quite different temporal and spatial logic. In this case image and information coexist to present new opportunities for the streaming of historical experience, "a more complex temporality constructed out of the relations between the narrativity of the

story and the interruptive, spatially distributed character of the still or unmoving image, the stasis of which registers the common now of the time of viewing" (Osborne, 2018: 142). This is the historical *contemporaneity* of such works; according to Osborne, "the bringing together of different times (different social times and different historical times) within the disjunctive 'living' unity of the present" (2018: 142). And although it would seem that still photographs (including moving photographs in the form of video) dominate contemporary image-space — with millions of images circulating across networks, and in the 'multiverse' in which events can happen at the same time, streamed through platforms across space-time and collected in vast annotated datasets — there remains considerable ambiguity and unevenness of distribution.

What remains is an indexical relation to an imagined reality and regime of truth that we historically associate with photography (and yet know to be false), and which confirms that nothing is what it seems. In informational space, images are stuck in an interplay of truth and falsehood, and in a confused relation to language (as in the case of OpenAI's ChatGPT and the use of prompts). In short, the computational image is different from what went before as it is composed of information and served through a structure, indeed infrastructure, that is not directly apparent. These images are distributed across the global distribution network of the internet and subject to the politics of globalization which resonates in the inherent tendency for images to be unable to settle in any one place (what Osborne refers to in terms of "migrancy of the image itself" — invoking T. J. Demos's figure of the "migrant image" (2018: 143)). They are subject to the multiplicity of instances inherent to the computational form and its distribution across planetary networks. These images appear to be photographs but are not.

Artificial stillness

In 'On the Concept of History (in Foundation Models)', and taking direct inspiration from Benjamin's essay, Fabian Offert (2023) asks what concept of history is inherent to a specific form of AI, namely OpenAI's CLIP (Contrastive Language-Image Pretraining), released in 2021, and other similar generative models. His starting point is the recognition that any technical object is temporal, and thus any outputs have an inherent relation to past inputs, and he examines the "structuring principles of these

internally consistent outputs, and how […] they relate to the structuring principles humans apply to the past to render it history" (Offert, 2023: 122). The quote previously cited by Osborne (and indeed hesitancy over the value of cross-disciplinary parallels) — "a memory as it flashes up at a moment of danger" — is used to stress how new insights are predicated on immediacy, and at a time of crisis.

Offert's essay contains technical precision, and the term 'foundation model' is clarified as one that is both large and able to be used for downstream tasks, in other words one that depends on the output of a previous task (2023: 124). However, although related, the example of CLIP is different to both DALL-E and Stable Diffusion as it is not generative and does not produce images or texts but rather connects them in latent space. It learns from the context of the image and produces a model based on spatial proximity, and thus can be used to identify images that are similar to each other; given a prompt, it will look for images similar to that prompt. This leads Offert to try to understand its particular concept of history through what he calls 'attribution by proxy' and 'generative attribution' (2023: 125, 128). In the first case, given its ability for image retrieval, CLIP can be a powerful tool to *see* non-obvious connections or *constellations* and produce abstract knowledge or *memory* of a given query. In the second, it can be seen how CLIP informs the training of generative models like DALL-E and Stable Diffusion.

Offert's example takes inspiration from the historical backdrop of Benjamin's essay in terms of subject matter in giving the prompt "a color photo of a fascist parade 1935" to DALL-E, knowing that the politically charged term *Fascism* was banned by OpenAI and therefore misspelling it for effect (although he notes that that trick is no longer possible) (2023: 128). What was returned was an image of what seems to be a Western European city with a mass rally, red flags raised, and smoke rising from a building in the background. What is striking to Offert is that the image is rendered in the style of a historical photograph with sepia tones, in effect reproducing "a (visual) world in which fascism can simply not return because it is safely confined to a black-and-white media prison" (2023: 129). Broadly, there's a version of technological determinism at work here (even if this isn't entirely the case with probabilistic programming) in which medium and history are conflated into a worldview in which fascism is no longer able

to be reimagined. It becomes clear that images produced in this manner are historically determined, too, as they are founded on computational statistics.

Part of the argument here is to put the representational image to one side and instead refer to the statistical properties. An example of this approach, and one that seems to anticipate AI's *aesthetic populism*, is Vitaly Komar and Alexander Melamid's project *The People's Choice* (from 1994–7), which generated the most wanted and least wanted paintings of a given nation (Wypijewski, 1997) based on polling statistics (producing versions of insipid landscapes and abstraction respectively). Although predating a deep learning model of statistics, the project comments on a nation's aesthetic biases and serves to mirror the problem of generalization inherent to machine learning, based as it is on prediction inference. Furthermore, it raises the question as to what extent populism finds its parallel in systems that generalize massive amounts of data and reduce things to their most generalized form by drawing upon a model of history rooted in traditionalism. By extension, do generative images give us an insight into the people's choice today, or indeed produce it, reifying the most wanted and popular? What is produced with generative AI applications like Stable Diffusion seems to do exactly this, and what was the least becomes the most wanted.[8]

Something similar is argued in Hito Steyerl's 'Mean Images' (2023) in which the competing *mean*ings of the term *mean* are developed through a play on words: images are both mean (as in the case of her own "demeaning" portrait rendered through Stable Diffusion), but also the mean of a statistical process (an average of a dataset), produced "through a filter of average internet garbage", as she dismissively puts it (2023: 84). By engaging closely with "the means of mean production" (Steyerl, 2023: 90), one can see how particular social relations emerge from this logic, as it becomes unclear who or what is being trained. This is the case, for instance, in everyday situations such as using an image captcha to verify that the user is human.

The historical roots of this mean-spiritedness are in mathematics and statistics. Generated images or what Steyerl calls "statistical renderings […] shift the focus from photographic indexicality to stochastic discrimination", and "no longer refer to facticity, let alone truth, but to probability" (2023: 82). The concern is the inherent link of statistics on which AI is founded to its application in eugenics and other (final)

solutions based on populist models of algorithmic governmentality. It has been discussed by many other commentators that the images of AI evoke the photographic composites of Francis Galton from the 1880s, not least by Kate Crawford and Trevor Paglan in 'Excavating AI' (2019), who state that the "underlying assumptions of physiognomy seem to have made a comeback with contemporary training sets", demonstrated in their use for tasks such as automated face detection and classification practices. Steyerl further develops this link to statistics in relation to financialization, with image generation a form of gambling in financial markets through which images are "kidnapped" and sold back to the public from which they were derived (like the selling of state assets as part of the process of privatization). She refers to this as "subprime visibility", like subprime speculative finance, which is a kind of poor image, quite literally one with bad credit (and taking inspiration here from Jonathan Beller's notion of 'derivative images' (2021)).[9] The data can be considered to be like debt in this way, based on risk and probability and, like any speculation, prone to vulnerabilities and errors, even collapse, as seen with the financial crash of 2007.

The overall concern, as my argument has developed — and beyond a reduction of value to economics[10] — is to understand how the past comes into contact with the present, and to what extent it interrupts or confirms the temporal process of historicism. For Benjamin, the dialectical convergence of past and present is what provides the politics. To reiterate, a "leap in the open air of history is a dialectical one" (2003: 395), the awareness of which will make the continuum of history explode in such a way as to allow its re-composition. But a fuller understanding of how the present is rendered by generative AI and the worldviews that underpin it seem to indicate a somewhat less radical potential for change and action in the present. In another example, and also mentioned earlier in this essay, Offert uses the prompt "Tank Man, 1989" to produce an image. What is generated is not an act of civil disobedience but merely a soldier proudly looking at a tank (Offert, 2023: 130). The example below (my own attempt using Stable Diffusion with the same prompt) produces a scene in which a seemingly disinterested person walks past a tank. So how has this come about? On a technical level, diffusion models, like the one used here, learn a diffusion process that generates the probability distribution of a given dataset and, unlike a GAN, are not trained on their own outputs. Instead, they follow an iterative process that drowns the

training data with noise, and then noise is gradually removed, effectively removing "the noise of reality" associated with photographic representation. All that remains is "a rendition of correlated averages" rather than an image of an actually existing event (Steyerl, 2023: 84). This consists of forward and reverse processes, and a sampling procedure which sounds uncannily dialectical, producing something like an image at a standstill but not one latently dynamic or action-generating beyond its surface appearance. Multiple probabilistic events might have been generated but imagination is compromised into a form of aesthetic populism.

Fig. 3: Image generated by the prompt "Tank Man, 1989" (after Fabian Offert), using Stable Diffusion.

Collapse

In line with Offert, any conclusion (if such a thing were possible) would suggest that generative AI forecloses political potential (2023: 131). Rather than allowing for the continuum of history to *explode*, it seems to *implode* into its own sense of immutability, leaving historical events without context and, as such, their re-composition seemingly pointless beyond entertainment value. In other words, *generative AI has no concept of history* — just like history is a mere reflection of 'real history' and has no history of its own (a reference to Marx and Engels's *The German Ideology*, in which ideology is understood to be an illusion produced by those in power). There is no conceptual or indeed political foundation for the model of history that unfolds, only computational statistics and probability functions. And yet, large image-language models like Stable Diffusion, like official history if we are to follow Benjamin, are also prone to collapse. In 'The Curse of Recursion' (Shumailov et al., 2023), also mentioned by Steyerl, there is an increased

expectation that model-generated material from the internet will increasingly be folded into the resulting models, with its context — already remote — further weakened through a process of recursion. This is not the end of history as such, but what the paper describes as "model collapse" (or entropy if we invoke statistics again) (Shumailov et al., 2023: 3). What unfolds is a technical argument which establishes that model collapse is a degenerative learning process that affects learned generative models, in which "generated data end up polluting the training set of the next generation of models; being trained on polluted data, they then mis-perceive reality" (Shumailov et al., 2023: 3). Models do not forget what went before but build upon previous falsifications, constructing a model of the world that is progressively remote from reality and reinforcing their own populist worldviews at scale (somewhat akin to fascism).

The "state of emergency" to which Benjamin once referred (2003: 392) remains the rule and not the exception as part of informational capitalism in terms of labour conditions, extractive practices, climate collapse, right wing populism, and the colonial wars in Ukraine and Palestine at the time of writing. Given these grave circumstances, the proposition to bring about a real state of emergency to oppose fascism as the historical norm seems to hold as much relevance now as it did then. The suggestion of impending collapse is an important point not only in terms of learning and the ontological structure of the image, but for its allegorical potential, too, toward a disjunctive model of history that corresponds with what has been discussed throughout this essay, and for what might be redeemed from the catastrophe in front of us. What can be seen in these ruins is not so much representational nor photographic but latent traces of material relations and infrastructures that render historical experience in compromised form.

Acknowledgments

References

Azar, M., G. Cox, L. Impett (2021) 'Introduction: Ways of Machine Seeing', *AI & Society* 36: 1093-1104. https://doi.org/10.1007/s00146-020-01124-6.

Azoulay, A.A. (2019) *Potential History: Unlearning Imperialism*. London: Verso.

Barthes, R. (1984) *Camera Lucida: Reflections on Photography*. London: Fontana.

Beller, J. (2021) *The World Computer: Derivative Conditions of Racial Capitalism*. Durham, NC: Duke University Press.

Benjamin, W. (2003) 'On the Concept of History', in H. Eiland and M.W. Jennings (eds.) *Selected Writings, Volume 4, 1938–1940*. Cambridge, MA: Belknap Press of Harvard University Press.

Benjamin, W. (1999) *The Arcades Project*. Cambridge, MA: Belknap Press of Harvard University Press.

Bowden, C. (2018) 'Forking in Time: Blockchains and a Political Economy of Absolute Succession', *APRJA* 7(1): 141-149. https://doi.org/10.7146/aprja.v7i1.116061.

Buck-Morss, S. (1995) *The Dialectics of Seeing: Walter Benjamin and The Arcades Project*. Cambridge, MA.: The MIT Press.

Chávez Heras, D. and T. Blanke (2020) 'On Machine Vision and Photographic Imagination', *AI & Society* 36: 1153-1165. https://doi.org/10.1007/s00146-020-01091-y.

Cox, G. and J. Lund (2021) 'Time.now', in N. B. Thylstrup, D. Agostinho, A. Ring, C. D'Ignazio and C. Veel (eds.) *Uncertain Archives: Critical Keywords for Big Data*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/12236.001.0001.

Crawford, K. and T. Paglen (2019) 'Excavating AI: The Politics of Images in Machine Learning Training Sets', *Excavating AI*, 19 September. Available at: https://www.excavating.ai/ (Accessed: 28 March 2024).

Dewdney, A. and K. Sluis (2023) *The Networked Image in Post-Digital Culture*. London: Routledge.

Ernst, W. (2017) *The Delayed Present: Media-Induced Temper(e)alities & Techno-traumatic Irritations of 'the Contemporary'*. Berlin: Sternberg Press.

Eshun, K. (2003) 'Further Considerations on Afrofuturism', *CR: The New Centennial Review* 3(2): 287-302. https://www.jstor.org/stable/41949397.

The Invisible Committee (2017) *Now*. Berlin: Sternberg Press.

The Last Angel of History (1995). Directed by J. Akomfrah and written by E. George, Black Audio Film Collective, C4/ZDF, London. https://vimeo.com/322377580.

Lorey, I. (2022) 'Benjamin: Leaps on Now-Time', in *Democracy in the Political Present: A Queer Feminist Theory*. London: Verso, pp.59-76.

Lund, J. (2019) *Anachrony, Contemporaneity, and Historical Imagination*. Berlin: Sternberg Press.

Mackenzie, A. and A. Munster (2019) 'Platform Seeing: Image Ensembles and Their Invisualities', *Theory, Culture & Society* 36(5): 3-22. https://doi.org/10.1177/0263276419847508.

Mitchell, W. J. T. (2006) *What Do Pictures Really Want? The Lives and Loves of Images*. Chicago: The University of Chicago Press.

Mulvey, L. (2006) *Death 24x Frames a Second: Stillness and the Moving Image*. London: Reaktion Books.

Offert, F. (2021) 'Latent Deep Space: Generative Adversarial Networks (GANs) in the Sciences', *Media+Environment* 3(2). https://doi.org/10.1525/001c.29905.

Offert, F. (2023) 'On the Concept of History (in Foundation Models)', *The Interdisciplinary Journal of Image Sciences* 37(1): 121-134.

Osborne, P. (2018) 'The Distributed Image', in *The Postconceptual Condition: Critical Essays*. London: Verso, pp.135-145.

Osborne, P. (2019) 'The Image Is the Subject: Once More on the Temporalities of Image and Act', in R. Görling, B. Gronau and L. Schwarte (eds.) *Aesthetics of Standstill*. Berlin: Sternberg Press, pp.124-137.

Samuel, A.L. (1959) 'Some Studies in Machine Learning Using the Game of Checkers', *IBM Journal of Research and Development* 3(3): 210-229.

Schwarte, L. (2019) 'Art in Times of Political Stagnation', in R. Görling, B. Gronau and L. Schwarte (eds.) *Aesthetics of Standstill*. Berlin: Sternberg Press, pp.90-104.

Shumailov, I., Z. Shumaylov, Y. Zhao, Y. Gal, N. Papernot and R. Anderson (2023) 'The Curse of Recursion: Training on Generated Data Makes Models Forget', *arXiv*. https://arxiv.org/abs/2305.17493v2.

Soon W. and M. Tyżlik-Carver (2023) 'Unerasable Images', in M. Devries, M. Tyżlik-Carver, W. Soon and G. Beiguelman (eds.) *Boundary Images*. Minneapolis/London: University of Minnesota Press/Meson Press, pp. 21-60.

Stable Diffusion. 'Stable Diffusion Online'. Available at: https://stable diffusionweb.com/#ai-image-generator (Accessed: 28 March 2024).

Steyerl, H. (2009) 'In Defense of the Poor Image', *e-flux* 10(November). https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/.

Steyerl, H. (2023) 'Mean Images', *New Left Review* 140/141: 82-97. https://newleft review.org/issues/ii140/articles/hito-steyerl-mean-images.

Wypijewski, J. (ed.) (1997) *Painting by Numbers: Komar and Melamid's Scientific Guide to Art*. New York: Farrar Straus Giroux.

Notes

[1] This relation between seeing and knowing is something I have previously explored with reference to John Berger's *Ways of Seeing*. If the relations between what we see and what we know are never settled, as Berger put it, then how are these relations further unsettled by developments in computer vision systems (Azar et al., 2020)?

[2] In the case of historical materialism this interplay of forces leads to the eventual improvement of social reality, a latent force ready to be awakened by revolutionary consciousness — what Benjamin describes as the synthesis of *dream* and *awakening* consciousness (Benjamin, 1999: 463).

[3] Full quote: "There was once, as we know, an automaton constructed in such a way that it could respond to every move by a chess player with a countermove that would ensure the winning of the game. A puppet in Turkish attire and with a hookah in its mouth sat before a chessboard placed on a large table. A system of mirrors created the illusion that this table was transparent from all sides. Actually, a hunchback dwarf [sic] — a master at chess — sat inside and guided the puppet's hand by means of strings. One can imagine a philosophical counterpart to this apparatus. The puppet, called 'historical materialism,' is to win all the time" (Benjamin, 2003: 389).

[4] Action-less action is key to Daoism and the paradox of *Wu Wei* (Chinese: "nonaction"; literally, "no action"), which doesn't mean not acting but effortless action or actionless action such that difficult tasks can be carried out with skill and efficiency.

[5] See also Winnie Soon's *Unerasable Images* (2018–2019), https://siusoon.net/projects/unerasable images, and Winnie Soon and Magdalena Tyżlik-Carver, 'Unerasable Images' (2023: 21-60).

[6] Full prompt: "There is a picture by Klee called *Angelus Novus*. It shows an angel who seems about to move away from something he stares at. His eyes are wide, his mouth is open, his wings are spread. This is how the angel of history must look. His face is turned toward the past. Where a chain of events appears before *us*, *he* sees one single catastrophe, which keeps piling wreckage upon wreckage and hurls it at his feet. The angel would like to stay, awaken the dead, and make whole what has been smashed. But a storm is blowing in from Paradise and has got caught in his wings; it is so strong that the angel can no longer close them. The storm drives him irresistibly into the future, to which his back is turned, while the pile of debris before him grows toward the sky. What we call progress is *this* storm" (Benjamin, 2003: 392).

[7] Benjamin's 'Toward the Critique of Violence', written in 1921, examines the justification of violence in different contexts given its operations in law and the state.

[8] Given more space, there could be an extended discussion here around what images themselves want, with reference to W. J. T. Mitchell's *What Do Pictures Really Want?* (2006).

[9] These ideas, closely related to those in Steyerl's 'Mean Images', were presented as a keynote address on 31 October 2023 as part of 'Critical AI in the Art Museum: Practices & Politics', hosted by ANU, http://criticalai.art/.

[10] It is also worth noting that in the case of machine learning, the costs are high, not only in financial but in environmental terms, as the infrastructure consists of massive, energy-heavy, top-down cloud architectures and server farms.

Geoff Cox is Professor of Art and Computational Culture at London South Bank University, UK, where he is co-director of the Centre for the Study of the Networked Image, and also Adjunct at Aarhus University, Denmark.

Email: coxg8@lsbu.ac.uk