



From photography to the convoked image: AI, digital archaeology, and documentary

De la fotografía a la imagen convocada. IA, arqueología digital y documental



Carlos Saldaña-Ramírez

<https://orcid.org/0000-0001-9336-7693>

Universidad Autónoma Metropolitana, México

csaldana@cua.uam.mx

How to cite this article

Saldaña-Ramírez, C. (2026). From photography to the convoked image: AI, digital archaeology, and documentary. *Revista Panamericana de Comunicación*, 8(1), 3817. <https://doi.org/10.21555/rpc.v8i1.3817>

Received: 25 - 04 - 2026

Accepted: 15 - 06 - 2026

Published online: 21 - 06 - 2026

Abstract

Two hundred years after the first photograph attributed to Nicéphore Niépce, visual culture faces a conceptual transformation: images generated by artificial intelligence challenge the categories through which photography has historically been understood. This essay argues that so-called “AI-generated photography” should not be conceived as an extension of photography, but as a convoked image: a visual form produced through the interaction of language, visual archives, statistical models, interfaces, and human imagination. Drawing on Jaron Lanier, Lev Manovich, Vilém Flusser, Jussi Parikka, Sarah Pink, Ian Goodfellow, Alec Radford, Luke Metz, Soumith Chintala, Jonathan Ho, Ajay Jain, and Pieter Abbeel, the article argues that these images are not taken, but requested, calculated, and made to appear. It proposes an anthropology of computational images and an archaeology of the convoked image to understand how AI constructs, reconstructs, and recodes possible worlds, as well as its implications for documentary practice, memory, social imagination, and contemporary artistic creation.

Keywords: Photography; Artificial Intelligence; Convoked image; Documentary; Visual culture; Digital archaeology.

Resumen

A doscientos años de la primera fotografía atribuida a Nicéphore Niépce, la cultura visual enfrenta una transformación conceptual: las imágenes generadas por inteligencia artificial desafían las categorías con las que se ha comprendido históricamente lo fotográfico. Este ensayo propone que la llamada “fotografía generada por IA” no debe entenderse



como una extensión de la fotografía, sino como imagen convocada: una visualidad producida por la interacción entre lenguaje, archivo visual, modelo estadístico, interfaz e imaginación humana. A partir de Jaron Lanier, Lev Manovich, Vilém Flusser, Jussi Parikka, Sarah Pink, Ian Goodfellow, Alec Radford, Luke Metz, Soumith Chintala, Jonathan Ho, Ajay Jain y Pieter Abbeel, se argumenta que estas imágenes no son tomadas, sino solicitadas, calculadas y hechas aparecer. El texto propone una antropología de la imagen computacional y una arqueología de la imagen convocada para comprender cómo la IA construye, reconstruye y recodifica mundos posibles, así como sus implicaciones para el documental, la memoria, la imaginación social y la creación artística contemporánea.

Palabras clave: Fotografía; Inteligencia artificial; Imagen convocada; Documental; Cultura visual; Arqueología digital.

1. Introduction. Two Hundred Years Later and the Crisis of the Index

Two hundred years after the first photograph attributed to Nicéphore Niépce, the question of the nature of the image reopens with unexpected intensity. For much of its history, photography was understood as trace, record, document, evidence, or testimony. Roland Barthes (1981) understood photography through its relation to the “that-has-been,” that is, through an ontological connection with whatever had stood before the camera. Susan Sontag (1977), for her part, noted that to photograph was to appropriate the thing photographed and to establish a particular relationship with the visible world.

Yet to present this scenario as an abrupt rupture exclusive to our own time would call for qualification. It is imperative to recognize that the crisis of the index and the loss of stability of the photographic trace long predate artificial intelligence. Historical practices such as photomontage, staged photography, and, decisively, the “digital turn” of the 1990s had already fractured the equivalence between photography and truth. As Joan Fontcuberta (2016) has extensively argued in his discussion of postphotography, visual culture has spent decades profoundly altering the conditions of production, circulation, and the regime of image credibility.

Even when photography was intervened upon, manipulated, or placed in crisis, it retained a central relation to capture: something, at some moment, had stood before a camera. That relation to the visible world organized not only its technique but also its cultural authority. The appearance of images generated by artificial intelligence, however, introduces a rupture deeper than the transition from analog to digital photography. The digital image could still retain a relation to luminous capture: the sensor replaced the film, but the fundamental operation remained tied to a recorded scene. An image generated by artificial intelligence, by contrast, can dispense with camera, lens, exposure, instant, and direct empirical referent. It is not taken: it is convoked. It does not necessarily record: it synthesizes. It does not depend on a prior scene: it can produce a possible scene.

Here the first conceptual problem of this article arises: why continue to call “photograph” an image that was not photographically produced? The immediate answer might be practical: because it resembles a photograph. But that answer reveals precisely the limit of contemporary visual thought. We continue to name the new with inherited words. We continue to evaluate the generative image by the criteria of the photographic image. We continue to ask whether it “looks real,” whether it “deceives,” whether it “imitates” a camera well, or whether it will “replace” photography. But perhaps these questions are insufficient because they proceed from a subordinating comparison: they force the generative image to appear before a photographic tribunal.

2. Naming the New: Lanier, the Anchor to the Past, and Recent Literature

Jaron Lanier allows this problem to be formulated with clarity. In *You Are Not a Gadget*, Lanier (2010) develops the notion of *lock-in*, that is, the fixing of digital structures that, once naturalized, condition subsequent forms of thought, creation, and experience. For Lanier, software is not neutral: it crystallizes decisions, metaphors, and ways of understanding the world. A digital architecture can open possibilities, but it can also freeze them. Applied to the generative image, this idea allows us to propose that calling an image produced through linguistic instructions “photography with artificial intelligence” constitutes a form of conceptual anchoring. This is not merely a terminological problem. It is a cultural difficulty in imagining the new without subjecting it to the vocabulary of the past.

Digital photography had already suffered this problem. For years it required a surname: “digital.” That surname marked a difference, but also a dependency. Digital photography was evaluated from the standpoint of analog photography: its sharpness, its grain, its manipulation, its archive, its print, its relation to the real. Something similar now occurs with the expression “AI-generated photography.” The qualifier “with AI” attempts to name the novelty while leaving the inherited noun—photography—intact. The result is a hybrid category, useful for everyday circulation but insufficient for critical thought.

To circumvent this *lock-in*, contemporary image theory has begun to populate a vigorous terminological terrain against which this essay must situate itself. Antonio Somaini (2023) uses the term “algorithmic images” to highlight how the latent space of neural networks operates as an invisible meta-archive that shapes our visual culture in unprecedented ways. Joanna Zylińska (2024), for her part, examines this transition toward a “perception machine,” locating the photographic future at a hybrid and conflictive threshold between the limitations of the human eye and the ubiquity of machinic vision. With particular precision for the thesis of this work, Roland Meyer (2025) argues that generative AI captures nothing but instead actively synthesizes “images from images.” For Meyer, the “photographic” is stripped of its ontological density and becomes simply a surface style, a patina that the algorithm can interpolate on demand.

This article proposes another path: to think of these images as “convoked images.” The convoked image is neither a false photograph nor a degraded photograph. It is another form of visibility. Its initial matter is not light but language. Its operation is not capture but generation. Its condition of possibility is not solely the camera, but the model trained on vast visual archives. Its authorship does not reside in a single subject but in a distributed relation among user, interface, model, dataset, platform, and prior visual culture. The convoked image thus forces a displacement of the question. It is no longer enough to ask: is it or is it not photography? The deeper question would be: what kind of world does an image produce when it is born from the interaction of words, data, models, and human imagination?

Every radically new technology passes through a phase of cultural translation. To assimilate it, a society compares it to previous technologies. Cinema was understood through theater, television through radio and cinema, digital photography through chemical photography, and now the generative image through photography. This operation is understandable: thought needs analogies to orient itself. But this can also be dangerous: too strong an analogy can prevent the proper concept from emerging. Lanier (2010) is fundamental to this argument because he allows us to understand how digital structures organize not only technical processes but also forms of imagination. When he speaks of *lock-in*, he does not refer solely to a technology that becomes difficult to modify. He also refers to a form of thought that becomes trapped within the conditions of a prior architecture. His critique of digital culture targets the reduction of the subject, of creativity, and of human experience to formats standardized by platforms, interfaces, and computational systems.

The generative image runs a similar risk. If we insist on calling “photography” everything that visually resembles a photograph, we end up reducing the complexity of the convoked image. We turn it into a variant of the known. We domesticate it. We force it into a genealogy that perhaps no longer fully corresponds to it. In Lanier’s (2010) terms, the problem is not only that technology becomes fixed; the language with which we think it can become fixed as well. This anchoring to the past has critical consequences. If an AI-generated image is thought of as photography, it is then required to answer photographic questions: what did it capture? Where was the photographer? What was the referent? What instant does it preserve? What evidence does it offer? But if that image does not proceed from a capture, these questions fail to explain its nature.

In fact, the convoked image demands other questions: what linguistic instruction originated it? What visual archive conditioned it? What model produced it? What social imaginaries does it reproduce? What desires of the user does it materialize? What biases does it carry? What possible worlds does it inaugurate? A first distinction can be proposed here: photography asks about what stood before the camera; the convoked image asks about what a culture can imagine with machines. This distinction does not seek to diminish the value of photography nor to uncritically celebrate artificial intelligence. It seeks to free both from confusion. Photography retains its power as trace, memory, testimony, and document. The convoked image opens another field: that of the synthetic visualization of possible worlds. The error, then, does not consist in comparing photography and artificial intelligence—the comparison can be fertile. The error consists in thinking that the AI-generated image must be understood, evaluated, or legitimated as photography. As Joan Fontcuberta (2016) has argued in his discussion of post-photography, contemporary visual culture not only multiplies images but also modifies the conditions of their production, circulation, and credibility. The convoked image is inscribed within this mutation, but it demands a category of its own because it does not merely alter photography: it displaces capture as the foundation of visibility.

3. From the Technical Image to the Computational Image: Flusser and Manovich

Vilém Flusser offers a second fundamental support. In *Towards a Philosophy of Photography*, Flusser (2000) proposed that technical images profoundly transform culture because they are no longer simple manual representations of the world but products of programmed apparatuses. For Flusser, the camera is not a transparent tool: it is an apparatus with a program, and the photographer plays with that apparatus, even against it. The photographer controls inputs and outputs but does not fully master the black box that organizes the image’s possibilities. This idea strikingly anticipates the contemporary condition of the generative image. The user of a visual AI also operates a black box. They write an instruction, adjust parameters, select results, discard variations, correct errors, insist. But they do not fully know the internal process by which the model produces the image. They control the input and evaluate the output; between the two unfolds an opaque zone of computation. In this sense, the convoked image extends the Flusserian problem of the apparatus but displaces it toward a more complex black box: no longer the camera programmed to record the world, but the model trained to synthesize it.

There is, however, a decisive difference. In Flusser’s (2000) photography, the apparatus still relates to the visible world through the camera. In the generative image, the apparatus no longer needs to position itself before the world. Its initial world is the archive. Its operative matter is data. Its field of action is probability. Here, Lev Manovich enters. In *The Language of New Media*, Manovich (2001) argues that new media are characterized by principles such as numerical representation, modularity, automation, variability, and transcoding. The image

ceases to be a stable surface and becomes a computational object—variable, manipulable, and programmable. Variability allows a media object to exist in multiple versions, often partially assembled by a computer; transcoding implies that cultural categories are reconfigured by computational logic (Manovich, 2001).

The convoked image radicalizes these principles. It is no longer only a matter of modifying an existing image but of producing a possible image out of relations among language, archive, and model. Numerical representation is transformed into statistical inference. Automation becomes generation. Variability ceases to be a secondary property and becomes the very condition of the image: every generative image is merely one version among many possible ones. Transcoding reaches a new level: a sentence becomes an image; an idea becomes a scene; a memory becomes an atmosphere; a desire becomes a visual composition. Following Manovich (2001), we might say that the generative image is not merely a digital image. It is a computational image in the fullest sense: an image whose existence depends on algorithmic operations that translate culture into data and data into visuality. The consequence is profound: if photography transformed the way we see the world, the convoked image transforms the way we imagine it.

4. A Brief Technical Genealogy: From Statistical Contestation to Diffusion-Based Appearance

So that this reflection does not remain solely on the philosophical plane, it is necessary to situate the emergence of the generative image technically. Ian Goodfellow and his collaborators proposed, in 2014, Generative Adversarial Networks (GANs), a framework in which two models are trained simultaneously: a generator that attempts to produce convincing samples and a discriminator that attempts to distinguish between real data and data produced by the generator (Goodfellow et al., 2014). From the standpoint of visual culture, this architecture is decisive because it shifts the image onto the terrain of statistical verisimilitude. The image no longer needs to be true in the photographic sense; it needs to be convincing within a data distribution. Its realism does not come from having stood before a camera but from having learned patterns from prior images.

Later, Alec Radford, Luke Metz, and Soumith Chintala developed Deep Convolutional Generative Adversarial Networks (DCGANs). Their work showed that these networks could learn hierarchies of visual representation, from parts of objects to complete scenes (Radford et al., 2015). This idea is culturally provocative: the machine does not merely copy images but learns visual structures. It learns what a culture has repeated sufficiently to become a pattern: faces, bodies, landscapes, framings, styles, gestures, proportions, atmospheres. Hence the generative image appears futuristic while also being profoundly anchored in the visual past. Its new worlds are made from the residues of previous images. AI seems to promise the future of the image, but it often does so by statistically recycling the visual past. This tension must be taken seriously. The convoked image is not born from nothing: it is born from an archive. And every archive contains memory, desire, violence, exclusion, repetition, and power.

With diffusion models, this logic takes another form. Jonathan Ho, Ajay Jain, and Pieter Abbeel presented Denoising Diffusion Probabilistic Models, capable of producing high-quality image synthesis through probabilistic processes of progressive noise removal (Ho et al., 2020). The aesthetic difference is important. A photograph is taken; a diffusion image appears. It does not emerge through contact with light but through an iterative operation that converts noise into form. This progressive appearance allows us to think of the generative image as a scene that was not captured but convoked. We might put it this way: photography fixes an instant of the world; the convoked image stabilizes a possibility within a field of probabilities. This

difference does not cancel the aesthetic dimension of the generative image; on the contrary, it situates it within a different regime of appearance. What once depended on an encounter between camera and world now depends on a negotiation among language, model, archive, and visual expectation.

5. The Convoked Image: Word, Archive, Model, and Desire

The convoked image arises when a linguistic instruction activates a computational process capable of producing a visuality. This definition, though apparently simple, implies a radical displacement in the history of images. Photography is born from the encounter between light, apparatus, and sensitive surface. The convoked image is born from the encounter between word, model, and archive. In the first, the world leaves a trace. In the second, the world is inferred from patterns. This difference resonates with Barthes's (1981) reflection on photography as certification of presence. In the convoked image, there is no longer necessarily a "that-has-been"; there might instead be a "this-may-appear." Its regime is not that of proof but that of visual possibility. Its value should therefore not be measured solely by its fidelity to a referent but by the way it organizes relations among language, visual memory, imagination, and culture.

The convoked image is not limited to representing what exists; it can construct worlds we have never seen but can immediately imagine. This capacity changes the cultural function of the image. We are no longer only before representations of the world but before visualizations of possibilities. AI allows us to create cities that do not exist, to reconstruct lost scenes, to imagine climatic futures, to formulate visual memories of unrecorded events, or to project bodies, territories, and communities under speculative conditions. But these possibilities should not be confused with absolute freedom. The convoked image is conditioned by the language that requests it, by the model that interprets it, by the data on which it was trained, by the interface that delimits options, by the platform that imposes restrictions, and by the cultural imaginaries that make the result recognizable. The convoked image seems to open infinite worlds, but those worlds are mediated by prior structures. As Lanier (2010) warns, digital architectures can fix possibilities; as Manovich (2001) shows, computational media translate cultural categories into computational operations.

Here it becomes useful to bring Manovich and Lanier together. With Manovich (2001), we understand that the generative image belongs to the regime of transcoding: a cultural idea is translated into data, and the data is recomposed as visual form. With Lanier (2010), we are warned that this translation can become trapped in inherited architectures that impoverish imagination. The convoked image is, simultaneously, the promise of openness and the risk of repetition. Its power lies in converting the word into a visual device. The *prompt*—or linguistic instruction—is not just any sentence. It is a form of visual writing. It is desire drafted. It is imagination made operational. It is a hypothesis of a world sent to a machine so that it may return transformed into an image.

For this reason, the analysis of a convoked image cannot be limited to the surface. It must ask about the words that made it appear. It must observe what desires the subject formulates, what references it invokes, what styles it requests, what exclusions it poses, what corrections it introduces. It must also observe what the system returns: what it interprets, what it omits, what it exaggerates, what it normalizes, what it stereotypes, what it renders visible, and what it renders impossible. The convoked image is not the exclusive product of the user nor of the machine. It is a negotiation. The subject proposes; the model responds. The subject adjusts; the model reinterprets. The subject insists; the interface limits. The subject selects; the archive reappears. In that circuit, the final image is merely the visible result of a chain of decisions, translations, and resistances. This leads to a central claim of this article: artificial intelligence does not only

generate images; it generates ways of imagining. The subject who convokes an image is also transformed by the images they receive. They learn to write for the machine, learn to desire within its possibilities, learn to adapt their imagination to probable outcomes. The machine, for its part, does not imagine in the human sense, but it returns visual configurations capable of affecting human imagination. Thus, the convoked image not only represents possible worlds: it trains the sensibility with which those worlds are thought, desired, and evaluated.

6. Toward an Anthropology of the Computational Image

If AI-generated images do not proceed from a capture but from an interaction among language, archive, model, and human intention, then their analysis demands another framework. Traditional photographic criticism does not suffice. It is not enough to ask about composition, lighting, referent, or authorship. Nor does technical analysis of the model suffice. What is needed is a perspective capable of observing how these images participate in the contemporary production of subjectivity, memory, and social imagination. I propose calling this perspective an “anthropology of the computational image.”

This proposal engages with digital ethnography, particularly with the work of Sarah Pink and her collaborators, who argue that research on the digital must attend to practices, experiences, relations, materialities, contexts, and forms of life mediated by technologies (Pink et al., 2016). From this perspective, AI-generated images should not be analyzed solely as visual objects but as cultural practices: processes in which subjects write, test, adjust, interpret, compare, discard, save, and share. The convoked image can be analyzed according to this logic: not as an isolated object but as a cultural practice. A person does not merely generate an image; they write, test, adjust, interpret, compare, discard, save, share. The process reveals a form of visual desire. The instruction shows what the subject wants to see, what words they use to imagine, what they understand by beauty, realism, future, pain, body, city, memory, or truth.

But the subject does not fully control the process. AI returns images that also affect them. It proposes forms, styles, compositions, and atmospheres. It teaches them what kind of image is more probable. It suggests paths. It disappoints them. It surprises them. It conditions them. In this exchange, the subject does not only produce images: they are also produced by them. This relation allows us to think of the convoked image as an anthropological practice of computational life. Contemporary societies do not merely live with machines; they imagine with them. They do not only store data; they produce memory through systems that recode archives. They do not only create images; they negotiate with algorithmic infrastructures that modulate the visible. The convoked image, then, should not be analyzed only for what it shows but for what it reveals about an era: its fantasies, its fears, its stereotypes, its desires for the future, its forms of nostalgia, its criteria of reality, its thresholds of surprise, and its modes of relating to technique.

This view approaches what Parikka (2012) understands as media archaeology: a way of examining layers, temporalities, residues, and continuities between old and new technologies. In this sense, an image generated of a future public hospital, an imaginary migrant border, a flooded Latin American city, or a community reconstructed from memory does not prove that those worlds existed. But it can reveal how a society imagines illness, displacement, disaster, loss, or hope. The convoked image does not always document events; it can document imaginaries. This is one of its most important powers for contemporary visual culture: if photography documented what stood before the camera, the convoked image documents the ways in which a culture learns to imagine with machines.

7. Where Does the Convoked Image Reside? Brain, Perception, and Machinic Sensibility

A seemingly simple question opens another zone of discussion: if it has long been said that the cerebral hemispheres perform distinct functions—one associated with logic, calculation, or language, and the other with emotion, intuition, or imagination—where, then, does the convoked image reside? The question is suggestive, but it requires an initial precision. The popular division between a “logical” hemisphere and an “emotional” hemisphere has been widely questioned by contemporary neuroscience. Cerebral functions are not distributed so simply. Although relevant lateralizations exist—for example, certain linguistic processes tend to be more associated with the left hemisphere and certain modes of spatial attention with the right—perception, imagination, emotion, memory, and aesthetic judgment involve distributed networks. The image does not occur in one half of the brain: it occurs in an architecture of relations.

Contemporary neuroaesthetics has insisted precisely that aesthetic experience involves perceptual, affective, cognitive, and evaluative processes that interact with one another (Chatterjee & Vartanian, 2014; Pearce et al., 2016). To deepen this distributed mechanism, it is essential to incorporate the foundational perspective of Anjan Chatterjee (2014) in *The Aesthetic Brain*, who demonstrates how the human brain evolved to desire beauty and to process artistic manifestations through modular systems that link the visual perceptual cortices with the pleasure centers of the limbic system and with semantic evaluation networks. This precision, far from weakening the metaphor, makes it more interesting. The convoked image does not belong to one hemisphere because it does not belong to a single regime of experience either. It is not merely calculation nor merely emotion. It is not only language or only visuality. It is not only technical operation or only aesthetic experience. It resides precisely in the zone of crossing between what modern culture has tended to separate: reason and imagination, word and image, algorithm and affect, analysis and desire.

Antonio Damasio proves useful for thinking through this integration. In *Descartes' Error*, Damasio (1994) questions the radical separation between reason and emotion, showing that affective processes participate in decision-making and in the rational orientation of behavior. If emotion is not the enemy of thought but a constitutive dimension of cognition, then the experience of an image cannot be reduced to its formal analysis or to its sensory impact. Seeing an image involves feeling, recognizing, comparing, remembering, anticipating, and deciding. The convoked image intensifies this integration. To generate it, the subject must formulate an instruction: they organize words, concepts, styles, atmospheres, references, and expectations. At this first moment, the image does not yet exist as visuality; it exists as linguistic desire. But when the model returns an image, the process shifts toward perception, affect, and aesthetic judgment: the subject decides whether what appeared corresponds to what was imagined, whether it moves them, whether it surprises them, whether it disturbs them, whether it disappoints them. The convoked image is born, then, from an oscillation between language and sensation.

Iain McGilchrist offers a complementary path, provided he is read carefully and not as a repetition of the simplistic hemisphere myth. In *The Master and His Emissary*, McGilchrist (2009) proposes that the difference between hemispheres does not lie simply in “which” function each performs but in “how” they attend to the world. One tends to operate in a more focal, instrumental, and analytical manner; the other, in a more contextual, relational, and open manner. Beyond the debates his thesis has generated, this distinction is fertile as a critical metaphor for thinking about the convoked image: AI converts an analytical instruction into a contextual appearance. It asks for verbal precision but produces visual ambiguity. It requires a

command but returns an atmosphere. From this perspective, the convoked image inhabits an intermediate space: between the focal attention of the *prompt* and the relational openness of the image. The user attempts to control with words; the machine responds with a visuality that exceeds those words. One writes “a migrant woman walking in the rain in a futuristic city,” but the result will not simply be the sum of those terms. Gestures, colors, textures, distances, climates, faces, wounds, clichés, cinematic resonances, and cultural sediments will appear that were not explicitly present in the instruction. The convoked image overflows the language that originates it.

This determines the way we see it, enjoy it, and analyze it. A photograph usually invites us to ask about the relation between image and event: what happened? Who was there? What was decided to be shown? What was left out of the frame? The convoked image forces another question upon us: how was an idea transformed into visual experience? Its aesthetic pleasure derives not only from recognizing a scene but from witnessing the appearance of something that previously existed only as verbal possibility, cultural memory, or imaginary desire. The neuroscience of visual imagination also helps sustain this shift. Pearson (2019) argues that visual imagination cannot be understood as a simple isolated process but as a complex cognitive activity linked to perception, memory, and voluntary control. The convoked image activates this complex network in a particular way because it requires comparing three simultaneous levels: what was requested, what appeared, and what that appearance makes us feel or think. Analyzing a convoked image thus means studying an experience of double translation. First, the subject translates an imagination into language. Then, the system translates that language into visuality. Finally, the subject translates the image once more into affect, interpretation, and meaning. In that circulation, something more than an image is produced: a form of visual subjectivity is produced. The question “where does the convoked image reside?” can then be answered conceptually: it resides in the interface between brain, language, archive, and machine. It resides in the circuit that connects a human intention with an algorithmic infrastructure and with a sensory experience. It resides in the tension between the calculation that produces it and the affect that receives it. It resides, above all, in the space where human imagination begins to negotiate with systems capable of returning images of what it has not yet seen. The convoked image does not belong to the hemisphere of reason nor to the hemisphere of emotion. It belongs to a new ecology of perception, in which thinking, feeling, writing, calculating, and looking become inseparable operations.

8. Archaeology of the Convoked Image

To analyze a convoked image, it is not enough to look at the final image. One must excavate its conditions of appearance. Jussi Parikka is especially useful here. Media archaeology is not limited to studying old technologies but analyzes the relation between old and new media, between past and present, among archives, technical layers, memory, and digital culture (Parikka, 2012). The convoked image can be thought of as a stratified object: a visual surface that concentrates layers of language, archive, model, interface, desire, and power. I therefore propose an archaeology of the convoked image.

This archaeology would not seek buried material remains but operative and cultural layers: the words that originated the image, the aesthetic references convoked, the model used, the visual archives that condition its possibility, the cultural biases it reproduces, the discarded iterations, the relation between human desire and algorithmic response, the distance between what was requested and what was obtained, and the way in which the final image reorganizes the subject’s imagination. This archaeology allows us to understand that the generative image is not an innocent surface. It is a visual sediment. Accumulated within it are art history, photogra-

phy, advertising, cinema, social media, databases, colonial fantasies, gender imaginaries, racial clichés, technological aspirations, and individual desires. In this, the convoked image connects with Azoulay's (2008) concerns about the political dimension of photography, though it displaces the problem: it is no longer only about the civil contract surrounding a captured image, but about the political, technical, and cultural conditions of a generated image.

Faced with a photograph, classical criticism could ask: what happened before the camera? Faced with a convoked image, the question changes: what instructions, archives, and models made this appearance possible? The difference is decisive. In photography, analysis can be oriented toward the recorded scene, the point of view, the framing, the historical context, the act of capture, and the subsequent circulation of the image. In the convoked image, analysis must follow a different chain: instruction, interpretation, generation, selection, circulation, and reading. The event is not necessarily prior to the image; sometimes it appears as a result of it. This condition does not render the convoked image irrelevant. On the contrary, it makes it culturally dense. An image that does not record an event can nonetheless reveal a structure of social imagination. It can show how fear of the future, nostalgia for a lost past, the fantasy of progress, symbolic violence against certain bodies, or the aestheticization of catastrophe are visually encoded. For this reason, the archaeology of the convoked image must ask not only what is seen but how it was made visible; not only what it represents but what conditions made that representation possible; not only what world appears but what world had to be calculated, inherited, or desired for that image to emerge. The generative image is not a window onto the world but an excavation of the computational imaginary of an era.

9. Astonishment, Saturation, and the Fatigue of Novelty

The speed with which artificial intelligence has entered everyday life has produced an additional difficulty: we barely begin to understand one tool when another seems to surpass it. Novelty becomes obsolete before it becomes experience. Contemporary visual culture lives in a paradox: never before have we seen so many astonishing images, and yet, at the same time, it may never have been so difficult to be astonished. Walter Benjamin had already observed that technologies of reproduction alter an era's regime of perception. Technical reproducibility transformed the relation to aura, to authenticity, and to aesthetic experience (Benjamin, 2003). In the present, technical generativity introduces another mutation: it is not only that images are mass-reproduced; visual worlds are produced on demand. Abundance is no longer only of copies but of possibilities.

This abundance produces fatigue. Among generations formed within the permanent circulation of digital images, visual novelty tends to become naturalized quickly. Nothing seems sufficiently new because every image already appears inscribed within an infinite chain of variations, filters, memes, simulations, and technical updates. Nothing seems fully real because every image can be intervened upon. Nothing seems sufficient because another version can always be generated. This loss of astonishment should not be understood merely as a subjective incapacity. It is also a historical condition of perception on platforms. The constant flow of images reduces the duration of attention. Permanent comparison diminishes the intensity of experience. The infinite possibility of variation weakens the value of a singular image.

At this point, Sontag's (1977) reflection on photographic saturation acquires a new dimension: it is no longer only a matter of a world flooded with photographs but of a world flooded with possible images. Here the convoked image faces an aesthetic and cultural challenge: it is not enough to produce technically impressive images. The question is no longer whether an AI can generate something that looks real, but whether it can produce a meaningful experience. And that experience does not depend solely on the model but on the artist, the context, the

intention, the relation to audiences, and the way the image is inscribed within a narrative. The artist remains an artist with any tool, but no tool is neutral. Every tool opens a field of possibilities and closes off others. The camera taught us to see the world as a frame. Editing taught us to think of time as relation. Software taught us to intervene in the image as data. Artificial intelligence is teaching us to imagine as an interlocution with generative systems. The challenge consists in not confusing technical capacity with artistic power. The convoked image will not be relevant because it can produce anything, but because certain artists, documentarians, designers, photographers, researchers, and communities learn to convoke necessary images: images that do not merely dazzle but think; that do not merely imitate but reveal; that do not merely simulate worlds but allow us to understand them differently.

10. Documentaries and Artificial Intelligence: Testimony Without a Camera?

The question of the documentary then arises in all its complexity. Can a documentary be made with artificial intelligence? The answer cannot be simply technical. It is not enough to say yes because AI can produce plausible images. Nor is it enough to say no because those images were not captured. A documentary does not depend solely on the camera; it depends on an ethical relation with the historical world. Bill Nichols has insisted that a documentary is organized as a form of representation of the historical world (Nichols, 2017). Ariella Azoulay, for her part, has problematized photography as a civil, political, and ethical relation among subjects implicated in the photographic act (Azoulay, 2008). From these perspectives, the problem is not only the appearance of reality but the pact that the image establishes with the subjects, events, and memories it convokes.

A generative image can participate in a documentary practice if it does not attempt to supplant evidence or falsify testimony. It can serve to visualize absent memories, reconstruct unphotographed scenes, protect identities, represent traumatic experiences, elaborate speculative worlds, or expand interactive narratives. But it must declare its procedure. It must make its generative condition visible. It must avoid presenting itself as proof when it is in fact interpretation, reconstruction, or metaphor. AI does not eliminate the ethical problem of documentary; it intensifies it. If documentary photography already involved decisions of framing, power, representation, and responsibility—problems extensively discussed by Sontag (1977), Azoulay (2008), and Nichols (2017)—the convoked image adds new layers: who formulates the instruction? On what data was the model trained? What visual stereotypes does it reproduce? What relation does it maintain with the represented subjects? What does it disclose to the viewer about its process of generation? The most fertile question, then, would not be: can artificial intelligence make a documentary? But rather: under what ethical, aesthetic, and methodological conditions can a convoked image participate in a documentary experience?

The convoked image can be documentary not when it imitates photography, but when it establishes a responsible relation with a memory, a testimony, a community, or a historical experience. As Roland Meyer (2025) rightly argues, generative artificial intelligence operates through the recursive synthesis of “images from images,” which reduces the “photographic” to a mere superficial quality or an interpolable aesthetic style. In this way, the convoked image within a documentary framework acquires value not by attempting to simulate a false indexical record of light, but by honestly assuming its dimension as code. Its value lies not in passing itself off as evidence but in expanding the ways of figuring what was not recorded, what cannot be shown directly, or what needs to be imagined in order to be understood.

A particularly relevant possibility opens here for interactive documentary. If traditional documentary organized a relation among image, editing, and viewer, interactive documentary

introduced an experience of navigation, participation, and traversal. The convoked image can add a further layer: the possibility of constructing visualities from memories, accounts, archives, and user decisions. But this possibility must be treated with care. A generative visuality within a documentary cannot function as technological decoration; it must form part of an ethical and narrative strategy. A documentary made with AI should not ask only what images it can generate but what relation it produces with those who watch, with those who narrate, and with what is being remembered. In this sense, the convoked image can contribute to a new form of documentary not because it replaces the camera, but because it allows us to think of memory as a zone of reconstruction, imagination, and shared responsibility.

11. Limitations and Research Agenda

Despite the thoroughness of the ontological and conceptual framework proposed in this essay, certain theoretical-methodological limitations must be noted in order to circumscribe its scope. First, given the extreme speed at which machine-learning architectures mutate, the analytical categories tied to the convoked image must remain in a state of constant adaptive revision. This text concentrates primarily on the static dimension of probabilistic synthesis; however, the imminent standardization of text-to-video generation and the automated generation of real-time three-dimensional spatial environments will pose challenges of consistency and verisimilitude that will exceed the legacies of the still image analyzed here. Second, by privileging an approach grounded in the philosophy of the image, critical theory, and media archaeology, the study lacks a quantitative or qualitative empirical framework that would assess the cognitive reception and specific psychological responses of real audiences and cultural industries exposed to these visual flows.

From these delimitations, a predictive research agenda can be outlined, structured around three essential directions for the academic field of communication and visual culture:

1. **Political economy of the latent space:** Investigate the corporate mechanisms of technological centralization, the extraction of visual data without consent (*data scraping*), and the geopolitical labor dynamics of data labeling that sustain foundational AI models.
2. **Semiotics of the linguistic instruction:** Develop critical methodologies specific to the analysis of *prompting* as a new hybrid textual-visual genre, examining how algorithmic grammars impose standardized forms of desiring and imagining through corporate alignment filters and biases.
3. **Applied cases in co-creative documentary:** Empirically evaluate contemporary documentary projects that employ generative AI in collaboration with vulnerable communities or victims of historical trauma, analyzing the impact of the convoked image when explicitly used to fill gaps in state archives or censored records.

12. Toward a Revolution of the Image and of Artists

The present moment can be understood as a cultural limbo. Images generated by artificial intelligence already circulate massively, but we still lack solid criteria for evaluating them. We know how to recognize their spectacular quality, but we do not always know how to think through their status. We know how to admire their realism, but we do not always know how to interpret their conditions. We know how to fear their falsifying uses, but we are only beginning to understand their poetic, critical, and documentary possibilities. This limbo is not an empty pause. It is the space that precedes a revolution of the image. Earlier categories still carry weight, but they no longer suffice. Photography remains fundamental, but it can no

longer absorb every form of contemporary visibility. The digital image remains the technical environment of our culture, but the generative image introduces a different logic: the image is no longer only captured, edited, or shared; now it is convoked.

To break the anchor to the past does not mean abandoning the history of photography. It means ceasing to demand that the convoked image be what it is not. It means recognizing that we face another form of visual production, with its own powers and dangers. It means constructing new criteria for an image born of language, of the archive, of the model, and of human imagination. The convoked image should not be naively celebrated. It is traversed by biases, corporate dependencies, technical opacities, data extractivism, cultural appropriations, and risks of disinformation. But nor should it be reduced to threat or simulacrum. Within it opens a field of aesthetic, documentary, and anthropological experimentation that is only just beginning. Perhaps the decisive question is not whether AI will replace photography, but what new forms of sensibility will appear once we stop asking it to do so. Photography does not disappear; it changes place. It retains its force as trace, evidence, and memory of the world. The convoked image, by contrast, opens the space of possible worlds: not what was there, but what can be imagined, reconstructed, disputed, or warned of. Between the two there is neither mere continuity nor absolute rupture. There is a fertile tension. Photography taught us that the world could leave a trace. The convoked image forces us to ask what happens when imagination itself becomes technically generable. Therein lies the challenge for artists, documentarians, and image researchers: not to use AI to repeat the past with greater speed, but to open visual forms that do not yet have a name. To create will not simply mean producing images; it will mean building the conditions for those images to matter.

13. Conclusions

The image generated by artificial intelligence demands a conceptual transformation. Calling it photography may be convenient, but it limits our understanding of it. Photography is historically founded on a relation to light, the camera, the referent, and the event. The convoked image, by contrast, arises from the interaction among language, visual archive, statistical model, interface, and human imagination. It is not taken: it is convoked. It does not necessarily record a prior world: it can make a possible world appear.

Following Lanier (2010), it can be affirmed that insisting on inherited categories produces a conceptual anchoring to the past. Following Manovich (2001), the convoked image can be understood as a radicalization of the computational image, of automation, of variability, and of transcoding. With Flusser (2000), it can be thought of as an operation within a black box. With Goodfellow et al. (2014), Radford et al. (2015), and Ho et al. (2020), we recognize a technical genealogy that transforms the image into statistical verisimilitude and probabilistic appearance. With Pink et al. (2016) and Parikka (2012), the possibility opens of studying it as a cultural practice and as a stratified object, amenable to an anthropology of the computational image and an archaeology of its conditions of appearance.

The convoked image does not replace photography. Nor does it negate it. It forces photography to redefine its boundaries. At the same time, it demands new critical criteria for thinking about truth, memory, documentary, creativity, and subjectivity in an era in which words can become visual worlds.

If photography documented what stood before the camera, the convoked image documents the ways in which a culture learns to imagine with machines. In that difference lies both the danger and the promise of the contemporary visual revolution.

14. Conflicts of Interest

The author declares that there is no conflict of interest, financial or otherwise, that could have influenced the results or the interpretation of the findings presented in this article.

15. Ethical Compliance

The research underlying this article is theoretical-documentary in nature and does not involve experimentation with human subjects, collection of personal data, or intervention with participants. It therefore did not require evaluation by an institutional ethics committee. The work was carried out in accordance with the general ethical principles of academic research, including respect for authorship rights and the correct attribution of the sources consulted.

16. Funding

This work did not receive specific funding from public-sector agencies, commercial entities, or nonprofit organizations.

17. Open Science Objects

The article does not generate primary datasets, as it is a theoretical-documentary review.

18. References

- Azoulay, A. A. (2008). *The civil contract of photography*. Zone Books.
- Barthes, R. (1981). *Camera lucida: Reflections on photography*. Hill and Wang.
- Benjamin, W. (2003). *La obra de arte en la época de su reproductibilidad técnica*. Ítaca.
- Chatterjee, A. (2014). *The aesthetic brain: How we evolved to desire beauty and enjoy art*. Oxford University Press.
- Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences*, 18(7), 370–375. <https://doi.org/10.1016/j.tics.2014.03.003>
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. Putnam.
- Flusser, V. (2000). *Towards a philosophy of photography*. Reaktion Books.
- Fontcuberta, J. (2016). *La furia de las imágenes: Notas sobre la postfotografía*. Galaxia Gutenberg.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *arXiv*. <https://doi.org/10.48550/arXiv.1406.2661>
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *arXiv*. <https://doi.org/10.48550/arXiv.2006.11239>
- Lanier, J. (2010). *You are not a gadget: A manifesto*. Alfred A. Knopf.
- Manovich, L. (2001). *The language of new media*. MIT Press.
- McGilchrist, I. (2009). *The master and his emissary: The divided brain and the making of the Western world*. Yale University Press.

- Meyer, R. (2025). Images from Images: Generative AI and the Reconfiguration of the 'Photographic'. *Photography and Culture*, 18(3-4), 281–290. <https://doi.org/10.1080/17514517.2025.2522517>
- Nichols, B. (2017). *Introduction to documentary* (3rd ed.). Indiana University Press.
- Parikka, J. (2012). *What is media archaeology?* Polity.
- Pearce, M. T., Zaidel, D. W., Vartanian, O., Skov, M., Leder, H., Chatterjee, A., & Nadal, M. (2016). Neuroaesthetics: The cognitive neuroscience of aesthetic experience. *Perspectives on Psychological Science*, 11(2), 265–279. <https://doi.org/10.1177/1745691615621274>
- Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, 20, 624–634. <https://doi.org/10.1038/s41583-019-0202-9>
- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital ethnography: Principles and practice*. SAGE.
- Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv*. <https://doi.org/10.48550/arXiv.1511.06434>
- Somaini, A. (2023). Algorithmic images: Artificial intelligence and visual culture. *Grey Room*, 93, 74–115.
- Sontag, S. (1977). *On photography*. Farrar, Straus and Giroux.
- Zylinska, J. (2024). *The perception machine: Our photographic future between the eye and AI*. MIT Press.

